Three Essays in Economic Inequality

Jangyoun Lee

Faculty of Economics
University of Cambridge

This dissertation is submitted for the degree of
Doctor of Philosophy

Clare Hall September 2020

To my loving wife, Soohye, and to my precious two children, Dongha and Sunha . . .

Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Acknowledgements and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution. It does not exceed the prescribed word limit of 60,000 words.

Jangyoun Lee
September 2020

Acknowledgements

I am deeply grateful to my supervisor, Ha-joon Chang, for a tremendous amount of support, continued guidance and encouragement. He is a great role model in showing me how to be an honest and devoted researcher. Throughout my time as a research student at the University of Cambridge, I have benefited from many insightful discussions with Ha-joon, receiving constructive feedbacks that have undoubtedly improved the quality of my research.

I am truly indebted to my undergraduate mentor at the Korea University, Chang-jin Kim, without whose encouragement I would have never started my career as an economist at the Bank of Korea fifteen years ago and my PhD journey at a very late age. I am extremely thankful to Kang-kook Lee who saw the potential value of my research and encouraged me a lot during the most struggling time of my research period. I am also grateful to Sascha Becker, Wiji Arulampalam and Marija Vukotic for their invaluable advice and support throughout my MSc studies at the University of Warwick.

I would like to thank Tiago Cavalcanti, Petra Geraats, Elisa Faraglia, Kamiar Mohaddes and participants at the Cambridge PhD Macro workshop for insightful discussions and helpful advice. Many thanks go to the administrative and IT staff at the Faculty of Economics, particularly to Silvana Dean, Louise Cross and Jake Dyer. I am also grateful to my college tutor, Trudi Tate, and the administrative staff, Rebecca Williams, for their full support during my postgraduate life at Clare Hall.

My research has profited greatly from conversation with conference and workshop participants at a number of locations, including the 8th Society for the Study of Economic Inequality (ECINEQ) Meeting (Paris), the 95th Western Economic Association International (WEAI) Annual Conference (Denver), the 2020 Korean Economic Review (KER) International Conference (Seoul) and the 16th Economic Policies Research Unit (EPRU) Summer School (Braga). Especially, I am grateful to the programme committee of the 2021 Royal Economic Society (RES) Annual Conference (Belfast) for acceptance of the second chapter in this thesis. I was also honoured to be presented and received the Best Graduate Student Paper Award at the 2019 Korean Labour Income and Panel Study (KLIPS) Annual Conference (Seoul) for the third chapter in this thesis.

I gratefully acknowledge the financial support from the Cambridge Commonwealth, European and International Trust, the Bank of Korea and the Clare Hall, without whom my MPhil and PhD studies would not have been possible.

My time as a postgraduate student in Cambridge has been made enjoyable by the network of friends around me in the Faculty of Economics, Ha-joon's reading group and beyond. I will be forever thankful of the support I have received from Myungun Kim, Zeina Hasna, Shuyi Ge, Gustavo Nicolas Paez, Seunghyun Maeng, Taehee Oh, Pedro Mendes Loureiro, Mustafa Metin Başbay, Mateus Lino Labrunie, Shin Kubo, Jinki Lee, Jinhyun Yoo and Sanghyun Park. I have benefited from many helpful discussions with them through receiving constructive feedbacks that has undoubtedly given a myriad of creative ideas.

I would like to thank many fellows at the Bank of Korea, especially Heung-mo Lee, Hyung-joon Lim, Sung-sik Byun, Jong-ku Kang and Tae-kyung Kim. Their support has been important throughout my career. A special thanks to Dong-joon Ahn, Sung-ha Park, Won-joon Park, Chang-hyun Park, Seung-cheol Lee, Seok-mo Nam, and Ji-hyun Bae, for always listening to me, entertaining me and encouraging me.

A special feeling of gratitude to my loving parents, Myung-hyun Lee and Ok-ja Ma, to my siblings, Yeon-hee and Eui-yuan, to my mother-in-law, Young-sook Kim, to my sister-in-law, Soo-eun Kim, and to my lifelong friend, Hee-kwon Chung. Most of all, I am truly indebted to my beloved family, Soohye, Dongha and Sunha, who have sincerely encouraged and supported me. I totally dedicate this thesis to them.

Abstract

This thesis contains three chapters. The first two chapters provide the theoretical and empirical analysis of fast-growing inequality over recent decades, and the last chapter examines the employment effect of raising the minimum wage, a policy commonly used to support low-wage workers to mitigate increasing inequality.

The first chapter investigates why the upsurge of top income shares has coincided with economic slowdowns in the United States after the late 1970s. Based on theoretical and empirical backgrounds, I argue that a fast-growing unearned income from 'wealth residual', defined as the unexplained increase in wealth that is not accompanied by any increase in real output, lies behind them. To demonstrate this hypothesis, after measuring wealth residual from the national accounts and the associated statistics, I perform a set of panel vector autoregressive models with heterogeneous dynamics using a comprehensive dataset of the United States at the state level. The estimation results highlight that the rapid growth of wealth residual during the last four decades, influenced by the government's regressive policies on rent-related activities, has generated a co-evolution of fast-growing inequality and economic slowdowns. My further policy analysis reveals that a sharp increase in wealth tax or housing supply itself is not likely to solve the problem; on the contrary, industrial policies that can divert excess rent-related finance into productive investment must take precedence.

The second chapter investigates why some people 'stay' wealthy while others 'remain' poor despite the rapid expansion of aggregate wealth, and to what extent the government can affect wealth inequality. Based on the empirical evidence of persistently heterogeneous returns on wealth across households, I introduce the 'rentier premium', defined as the gap between the return on asset ownership and the return on capital used for the production of real output in a perfectly competitive market, into the standard heterogeneous agents dynamic stochastic general equilibrium model. I also examine the effect of two different types of government on wealth distribution and welfare gains. This study concludes that the rise in rentier premium and the government's regressive policies have acted as key drivers in the co-evolution of rising wealth inequality and declining labour shares in the United States since the 1980s.

The third chapter investigates how much higher the minimum wage can safely rise

without leading to employment losses. The unprecedented impeachment of South Korea's former president in mid-2017 and the subsequent large rise in the minimum wage allows us to investigate the non-linear employment effects of increases in the minimum wage on low-paid workers. To demonstrate these effects, I use two-step generalized method of moments, difference-in-differences and regression discontinuity designs based on individual-level panel data over the period of 2009–2018. The estimation results show that a 14.9% rise in the real minimum wage in 2018 seems to hit the tipping point, which is 5.5% in my estimation. This unexpected double-digit growth in the minimum wage has led to a reduction in the number of hours worked by low-wage workers, compared with modest increases over a decade under business-friendly governments. Paradoxically, the large increase in the minimum wage has had an unintended positive consequence of promoting better work-life balance in South Korea – a place that is notorious for over-working. We ask: Can increases in the minimum wage actually help low-income workers? It depends.

Table of contents

List of figures

List of tables

Chapter 1

Wealth Residual Hypothesis: Behind Rising Inequality and Falling Growth

This paper investigates why the upsurge of top income shares has coincided with economic slowdowns in the United States after the late 1970s. Based on theoretical and empirical backgrounds, I argue that a fast-growing unearned income from 'wealth residual' – the unexplained increase in wealth that is not accompanied by any increase in real output – lies behind them. To demonstrate this hypothesis, after measuring wealth residual from the national accounts and the associated statistics, I perform a set of panel vector autoregressive models with heterogeneous dynamics using a comprehensive dataset of the United States at the state level. The estimation results highlight that the rapid growth of wealth residual during the last four decades, influenced by the government's regressive policies on rent-related activities, has generated a co-evolution of fast-growing inequality and economic slowdowns. My further policy analysis reveals that a sharp increase in wealth tax or housing supply itself is not likely to solve the problem; on the contrary, industrial policies that can divert excess rent-related finance into productive investment must take precedence.

Keywords: wealth residual; inequality; growth; panel vector autoregressive model
JEL: D31; C33; O43

1.1 Introduction

"The rent of land, therefore, considered as the price paid for the use of the land, is naturally a monopoly price. It is not at all proportioned to what the landlord may have laid out upon the improvement of the land, or to what he can afford to take; but to what the farmer can afford to give." - Smith (1776), Wealth of Nations (p.162)

In the middle of the twentieth century, it came to be believed that economic growth would bring about an increase in aggregate wealth and higher living standards for society as a whole. Actually, during the post-World War 2 period, there existed some evidence supporting this belief, especially for the US, often referred to as 'the Golden Age of American Capitalism' (Marglin and Schor, 1992). Since the late 1970s, even though the aggregate wealth in the US has rapidly expanded, the trend of incomes towards greater equality of incomes has been suddenly reversed (Atkinson, 2015). For instance, real gross domestic product (GDP) per capita of the US grew at an average annual rate of 1.7 percent during 1978-2015 – much lower than the 2.7 percent seen from 1959-1977 – while the share of the top 1% income was about 22.0 percent in 2015, compared to only 9.0 percent in 1977. The 1978-2015 period has been also characterised by slowdowns in corporate investment. Net fixed capital formation in the non-financial corporate sector was, on average, 6.5 percent of gross value added and 4.5 percent of net capital stock during 1959-1977. In contrast, these numbers fell respectively to 5.1 percent and 3.2 percent during 1978-2015.

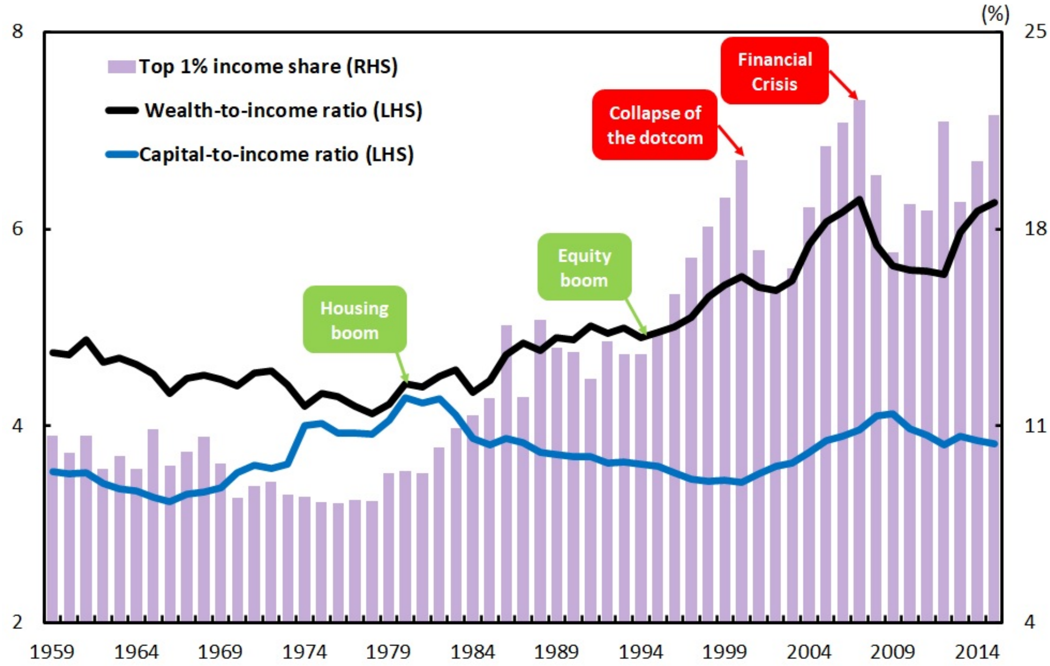
Why has the upsurge of top income shares coincided with economic slowdowns in the US after the late 1970s? I find a clue from wealth-to-income ratios that have significantly risen in comparison to capital-to-income ratio since 1978[1]. As in Figure 1.1, the wealth-to-income ratio remained stable at approximately 4.5 until 1978, but the ratio substantially increased up to around 6.3 in 2015, mainly due to increases in the price of equities and real estate. In contrast, capital-to-income ratio relatively remains flat. Concurrently, there has been the upsurge of income in the richest 1% group during the same period. In this respect, this paper attempts to investigate the gap between wealth and capital, referred to as 'wealth residual' – the unexplained increase in wealth without a corresponding increase in investment (Stiglitz, 2015a).[2] I then examine the role of wealth residual in the concurrence of rising top income inequality (hereafter, "inequality") and falling growth rates of real income per capita

[1] In this paper, as in Saez and Zucman (2016), I set up the year 1978 as a watershed because the share of the top 1% income in 1978 was the lowest point in history and wealth-to-income ratios had been stable until 1978. Also, in 1978, the US government reduced the maximum rate on long-term capital gains from its historical peak of 39.875% to 28% for the first time after the World War 2.

[2] Wealth and capital are conceptually different (e.g. Fagereng et al., 2020, 2019; Rognlie, 2015; Stiglitz, 2015a). The former reflects control over resources, while the latter is a key input into production processes. The

(hereafter, “growth”) during the last four decades.

Fig. 1.1 Wealth-to-income, capital-to-income, and share of the top 1% income



Notes: Wealth, the market value of stocks, bonds, housing, pensions, and business assets minus liabilities held by households and non-profit organisations; capital, the current-cost net stock of fixed assets; income, net national income
Sources: The author utilised the Federal Reserve Economic Data and the World Inequality Database

A growing body of recent studies (e.g. Brun and Gonzalez, 2017; Eggertsson et al., 2018; Fagereng et al., 2019; Rognlie, 2015; Stiglitz, 2015a) suggests that the decoupling of wealth and capital poses a challenges for the general theory of an economy driven by the underlying laws of supply and demand, focusing on labour and capital only. In standard models where firms accumulate capital and households buy corporate securities, the aggregate market value of these two variables is expected to move in the same direction. That is, capital is the part of firms’ net worth that yields the flow of future dividends capitalised in the market value

difference can be directly tracked down from the series of the national accounts: profit is the flow of income accruing to the owners of the produced capital stock necessary for production, while economic rent is the flow of income accruing to the owners of certain financial assets or tangible non-produced assets like land (Alvaredo et al., 2020). Thus, wealth is a more encompassing term than capital because it not only includes assets produced by ‘humans’, such as the capital stock of buildings, equipment, infrastructure, but also includes assets that are not productively employed to generate new income, such as some financial claims or land. The detailed description of wealth residual is illustrated in Section 3.

of corporate securities. So, in the absence of frictions, these two aggregates are equal and their ratio (the so-called "Tobin's Q") should equal one in equilibrium. However, historical values of Tobin's Q during the post-1978 period have neither been constant nor equal one, implying that wealth has been largely accumulated through 'persistent capital gains' in the stock and housing markets rather than through savings (Eggertsson et al., 2018; Fagereng et al., 2019; Gutierrez and Philippon, 2017). This general theory also finds it difficult to generate results that fit the stylised macroeconomic facts of the post-1978 period (Atkinson, 2015; Eggertsson et al., 2018; Stiglitz, 2015a).[3]

In a series of innovative papers, Stiglitz (2015a,b,c,d) tries to resolve this puzzle by introducing a new concept of wealth residual that is not related to the production of real output but is associated with asset ownership, giving rise to economic rents[4] from land and monopoly. On top of his theoretical perspective, this paper is, I believe, the first to present the estimate of wealth residual for the US by using the national accounts and the associated statistics: wealth residual increased by 59.3 times between 1978 and 2018, while capital increased by only 8.7 times. As a result, the ratio of wealth residual to total wealth in 2018 was 39.7 percent, much larger than 8.8 percent in 1978. More importantly, due to a privileged access to wealth residual, the owners of wealth residual can get additional wealth gains at the expense of their counterparts through rentier premium – the returns over and above the returns on capital in a competitive market. As a result, while owners of equities and land benefit greatly from the rise of its value and from increasing returns on wealth, households whose income depends mostly on labour suffer from weak capital formation.

The important source of the disparity between the growth of wealth and that of capital is land (Fagereng et al., 2019; Piketty, 2014; Rognlie, 2015; Stiglitz, 2015a,e). More importantly, much of the increase in aggregate wealth is an increase in the *value* of land but rarely related to an increase in the *(productivity-adjusted) amount* of land. That is, the increase in the value of land is not the result of the creation of more land; it is merely the increase in the *price* of existing land. Since the appreciation of land value reflects a desire for its positional value (e.g. access to high-quality education or scenic views), as firstly explored in Hirsch (1977), attempts to acquire land can only benefit one agent at the expense of another agent, resulting in a zero-sum game without any increase in the productive capacity of the

[3]Eggertsson et al. (2018) describe the five puzzling macroeconomic trends that are inconsistent with the canonical Kaldor's stylised facts: (i) an increase in the wealth-to-income ratio, with a stagnating capital-to-income ratio, (ii) an increase in Tobin's Q to a level above one, (iii) a decrease in the real rate of interest, while the measured average return on capital is relatively constant, (iv) a decrease in both the labour share and the capital share, and (v) a decrease in investment-to-output ratio despite low borrowing costs.

[4]Economic rents are defined as 'that part of the payment to an owner of resources over and above that which those resources could command in any alternative use' (Buchanan, 1980), that is, 'the receipt in excess of the owner's opportunity cost of the resources' (Chang, 1994).

economy – referred to as ‘scarcity rent’ that are evidently capitalised in non-financial assets (e.g. Rognlie, 2015; Stiglitz, 2015a). With a similar mechanism, another crucial source of wealth residual is rent-related finance – those financial assets that do not contribute to the increase in (productive) capital or land – referred to as ‘monopoly rents’ that are unobtrusively capitalised in financial assets (e.g. Brun and Gonzalez, 2017; Lusardi et al., 2017; Philippon and Reshef, 2012; Stiglitz, 2015d).

Based on these theoretical and empirical backgrounds, this paper presents the wealth residual hypothesis that a fast-growing unearned income[5] from wealth residual is a key driver of the concurrence of surging inequality and sluggish growth over the last four decades.[6] To demonstrate this hypothesis, I perform a set of panel vector autoregressive (PVAR) models with heterogeneous dynamics developed by Pedroni (2013), which control for state-fixed effects and allow for full heterogeneity of dynamics across states. This technique results in a distribution of impulse response functions, permitting a much more robust inference than those that rely on average estimates and assume slopes are homogeneous (Goes, 2017).

To make it possible to use this technique, I put together a much larger and more comprehensive panel dataset of the US at the state level on wealth residual, inequality, growth, and relevant environmental and policy variables, such as human capital, innovation, tax, regulations, labour bargaining power, and political regimes, than it has hitherto been available. More importantly, the use of the US state-level data is much superior to the use of cross-country data in assessing the empirical relationship among economic variables because cross-state heterogeneity is much smaller than cross-country heterogeneity. Accordingly, this method allows us to avoid inherent problems in postulating a common parametric structure for cross-country samples that include countries at different stages of economic development and/or with different characteristics (e.g., size, specialisation). Finally, it not only mitigates the low-quality and low comparability issues inherent in the use of cross-country data but also helps us alleviate significant problems of endogeneity, omitted variables, and measurement errors often generated from cross-country analyses (Banerjee and Duflo, 2003).

The main empirical findings in this study are as follows. First, the relationship between growth and inequality had structurally changed by the beginning of 1978. Unlike the 1959-

[5]Unearned increment, coined by Mill (1821), is an increase in the value of land or any property without expenditure of any kind on the part of the proprietor. He proposed taxing it so that it benefits every member of a society. Mill’s concept was refined and developed by George (1879) as unearned income that refers to income gained through ownership of land and monopoly.

[6]Admittedly, higher inequality may hurt growth through various channels. First, higher inequality can aggravate the adverse effects of credit constraints on human capital formation and hence lead to lower growth when human capital is an engine of growth (e.g. Galor and Moav, 2004). Second, inequality leads to weak aggregate demand because those at the bottom spend a larger spend fraction of their income than those at the top (e.g. Auclert and Rognlie, 2018). In this respect, this study builds on these ongoing studies to find uncovered forces driving the concurrence of rising inequality and falling growth.

1978 samples that witnessed the traditional trade-off between growth and inequality, higher inequality is detrimental to growth over the period of 1978-2015. Second, I extend my analysis by including a measure of wealth residual to test the wealth residual hypothesis. This study reveals that there exists a wealth residual mechanism within the post-1978 growth-inequality relationship: after wealth residual increases, growth and inequality became worse at the same time. This is mainly because while owners of wealth residual benefit greatly from the rise of its value and from increasing equity returns, households whose income depends mostly on labour suffer from the stagnation of capital formation. These results are robust when I use different model specifications and alternative measures of inequality (e.g. Gini index, Atkinson index, top 10% shares, and top 0.1% shares instead of top 1% income shares) and wealth residual (e.g. land price index instead of the share of land in the market value of real estate). Finally, this study incorporates government's rent-related policies into the trivariate PVAR model. I conclude that the government must enact stricter regulation of rent-related activities (e.g. loan-to-value ratio) to dampen the intensity of positional competition or increase affordable housing to reduce absolute poverty. Furthermore, using discrete choice models, this study finds that a labour-friendly government tends to win elections over a rentier-friendly one when it has maintained both wealth residual and inequality at low levels during its term in the office.

Overall, this study concludes that the rapid increase in wealth residual, influenced by the government's regressive policies on rent-related activities, has generated a co-evolution of fast-growing inequality and economic slowdowns. This is because the increase in wealth residual has rarely contributed to the expansion of productive capacity of the economy, partly due to intensified positional competition. That is, attempts to acquire wealth residual can only benefit one agent at the expense of another agent, resulting in a zero-sum game without any increase in the productive capacity of the economy (the so-called "exploitation mechanism"– Stiglitz (2015a)). In other words, the force of the exploitation mechanism seems to have crowded out the accumulation of capital in the US since the late 1970s.[7]

The rest of the paper is organised as follows. Section 2 reviews the related literature, while Section 3 presents the theoretical model to spell out its empirical implications. In Section 4, I describe an empirical strategy that includes data, methodology, and model specifications. Section 5 presents empirical results and checks robustness. Section 6 discusses the policy implications of our findings. The last section concludes.

[7] As Smith (1776) predicted much earlier, our conclusion implies that an economy which uses the revenues gleaned from the extraction of its resources to fuel property boom might show an increase in aggregate wealth, but its future productive capacities may well be significantly diminished, and instead, inequality will grow.

1.2 Literature Review

The starting point of this paper is Piketty (2014), who concludes that the upsurge of wealth-to-income ratios was not accompanied by decreasing returns on capital, yielding the over-accumulation of capital and growing top income inequality. But debates arose regarding several of his assumptions – especially, the elasticity of substitution between capital and labour (e.g. Chirinko, 2008; Chirinko and Mallick, 2017; Rowthorn, 2014) and the equation of the concept of capital and that of wealth (e.g. Eggertsson et al., 2018; Stiglitz, 2015a,e). In this respect, this paper is closely connected to Stiglitz (2015a), who argues that wealth and capital are distinctly different concepts. If wealth is equated with capital, the increase in capital should be associated with a decline in the return to capital and an increase in real wages. However, this hypothesis is contradicted by the macroeconomic data of the post-1978 US economy. To address this puzzle, Stiglitz (2015a,b,c,d) proposes a theory that the disparity between wealth and capital (i.e., wealth residual) might be partially due to an increase in scarcity rents, evidently capitalised in non-financial assets (e.g., land), and monopoly rents, unobtrusively capitalised in financial assets. I further develop his theory by empirically analysing the effect of wealth residual on the existing growth-inequality relationship by using the first-ever estimate of wealth residual of the US.

This study also resonates with the recent literature on puzzling trends of the post-1978 US macroeconomic data that has overturned the Kaldor (1961)'s famous stylised facts, such as constant interest rates, a constant labour share, and a constant capital-to-income ratio. Notably, Piketty and Zucman (2014) document evidence that wealth-to-income ratios have risen over the past forty years, along with an increase in Tobin's Q. They decompose the increase in wealth into two components – a saving component and a valuation component. The saving component can be described as the equation $\beta = \frac{s}{g}$, where β is the ratio of the wealth-to-income ratio in the long run, s is the saving rate, and g is the growth rate. There can be valuation effects in the short run as the price of capital goods (i.e., Tobin's Q) increases. In line with this argument, Fagereng et al. (2019), using Norwegian administrative panel data on income and wealth, show that saving rates including capital gains ("gross saving rates") increase markedly with wealth, while saving rates net of capital gains ("net saving rates") are approximately constant across the wealth distribution. It implies that wealthier households own assets that experience 'persistent capital gains' which they hold onto instead of selling them off to consume. They term this phenomenon "saving by holding".

Meanwhile, Eggertsson et al. (2018) argue that an increase in the wealth-to-income ratio, an increase in Tobin's Q, and a divergence between the marginal and the average returns on capital can be explained by an increase in monopoly rents. Similarly, Karabarbounis and

Neiman (2013) and Barkai (2020) provide evidence that the capital share and the labour share for the US income declined at the same time; instead there is an increase in the mark-ups or what Karabarbounis and Neiman (2018) call factor-less income – income which accrues neither to labour nor to capital. In contrast, Rognlie (2015) demonstrates the dominant role of scarcity rents from real estate in explaining these post-Kaldor's stylised facts.

This study is also closely related to the empirical literature concerning the relationship between growth and inequality. Since Kuznets (1955) posited an inverted U-shaped relationship between growth and inequality in the industrialised countries, a number of studies have tried to establish the relationship between the two variables. However, the results are discordant, depending on the main channel stressed (e.g. credit-market imperfections, political economy, socio-political instability, and differences in marginal propensity to save between different income groups) besides the functional form that each author has used.

We can classify these studies into three broad categories. The first category of authors has argued that inequality leads to higher savings because the rich save proportionately more than the poor, subsequently leading to an increase in investment and growth rates (Forbes, 2000; Galor and Tsiddon, 1997; Partridge, 1997). In particular, the recent work of Cavalcanti and Giannitsarou (2017), using the concept of network cohesion, a summary statistics that arises naturally in dynamic models of endogenous perpetual growth with network externalities, has confirmed the conventional trade-off between economic growth and inequality for a given network structure. More specifically, when the network's cohesion is low, the more likely it is to have high growth and inequality in the long run.

In contrast, the second category of studies states that inequality hurts growth because it leads to redistributive pressures, either through the median voter who enacts redistributive taxes (Alesina and Rodrik, 1994; Persson and Tabellini, 1994), or through the generation of socio-political instability, risk of violent conflict, and the pursuit of rent-related activities (Acemoglu and Robinson, 2000; Alesina and Perotti, 1996; Benabou, 1996; Perotti, 1996), and reduced capacity to invest when capital markets are imperfect and set-up costs are large (Aghion et al., 1999; Galor and Zeira, 1993). Meanwhile, the recent work of Brueckner and Lederman (2018), using instrumental variables regressions, show that inequality has a significant negative effect on transitional growth in high income countries during 1960–2007 possibly through human capital channels due to financial constraints.

The final category of studies suggests a potential non-linear relationship between growth and inequality. Notably, Barro (2000) shows that the relationship is negative among poor countries but positive among rich countries, after controlling measurement errors using the three-stage least squares estimators. Banerjee and Duflo (2003) criticize the functional form assumptions made in previous studies and argue that the growth rate is an inverted U-shaped function of net changes in inequality. Halter et al. (2014), using system generalized method of moments (GMM) estimators based on a panel of 106 countries during the 1965-2005 period, find that higher inequality helps economic performance in the short term but reduces the growth rate farther in the long term. The recent work of Costantini and Paradiso (2018), using a penalised principal component estimator based on panel data of the US at the state level over the period of 1960-2015, suggests the existence of an S-shaped relationship between growth and inequality.

In this regard, this study contributes to the literature that have attempted to establish the non-monotonic relationship between growth and inequality. However, unlike previous studies that mostly focus on a single-equation estimation, this study attempts to capture the interactive feedback dynamics of the growth-inequality relationship.

Finally, this paper is methodologically related to the literature that uses PVAR techniques in studying the empirical relationship between growth and inequality. Emergence of availability of high-quality cross-state panel data of the US has only recently made it possible to use this method to overcome the limits of cross-country estimation. There have been a few studies of the growth-inequality relationship by using this technique. Atems and Jones (2015), using a panel of 51 US states spanning 1930-2005, show that higher inequality not only reduces the growth rate of real per capita personal income but also the long-run level of real per capita personal income. My study distinguishes itself from Atems and Jones (2015) in highlighting the channel of wealth residual in the growth-inequality nexus that is missing in their studies.

1.3 Wealth Residual Hypothesis

Wealth and capital are conceptually different (e.g. Brun and Gonzalez, 2017; Eggertsson et al., 2018; Fagereng et al., 2020, 2019; Rognlie, 2015; Stiglitz, 2015a). The former reflects control over resources, while the latter is a key input into production processes.[8] The difference can be directly tracked down from the series of the national accounts: profit is the flow of income accruing to the owners of the produced capital stock necessary for production, while economic rent is the flow of income accruing to the owners of certain financial assets or tangible non-produced assets like land (Alvaredo et al., 2020). Thus, wealth is a more encompassing term than capital because it not only includes assets produced by 'humans',

[8] The concepts of wealth and capital had been extensively discussed in the literature (e.g. Auerbach, 1989; Boskin, 1988; Bradford, 1989; Eisner, 1988; Goldsmith, 1982; Hicks, 1975; Peek, 1986).

such as the capital stock of buildings, equipment, infrastructure, but also includes assets that are not productively employed to generate new income, such as some financial claims or land. Such a monetary definition of wealth is subject to the changes of financial market valuations, which can make the overall value of the stock of wealth diverge substantially from the value of the capital stock as measured by its replacement cost. In particular, wealth has been largely accumulated through ‘persistent capital gains’ in the stock and housing markets rather than through productive savings (Eggertsson et al., 2018; Fagereng et al., 2019; Gutiérrez and Philippon, 2017).[9]

According to my estimation using the national accounts and the associated statistics, much of the increase in wealth that can be observed from the 1980s onwards does not correspond to a rise in productive capital. More and more money was lent to investors who mainly did not use it to create new businesses or make productive investments in existing businesses, but to speculate in already existing assets thereby pushing up asset prices (e.g. Almeida et al., 2016; Gruber and Kamin, 2017; Gutierrez and Philippon, 2017). For instance, suppose that valuable real estate is owned mostly by the wealthy. If this real estate becomes more valuable, the wealth of the country increase, but wealth also becomes more unequally distributed. But simply because the price of land has gone up does not mean that the US economies have become more “productive".

1.3.1 Definition of wealth residual

Before using the data, I formally define wealth residual to explore how it relates to concepts of income and wealth in the national accounts of the US. The primary goal of this section is to highlight a close theoretical connection between wealth residual and inequality, and growth.

An agent i born at time ρ, with a lifespan of M years, optimises her lifetime utility $U^i(c^i_\rho, c^i_{\rho+1}, ..., c^i_{\rho+M-1})$, where $U^i(\cdot)$ is concave and non-decreasing. At each time t, the agent chooses between consumption c^i_t and purchasing two types of assets[10]. The first type is non-financial assets (e.g., real estate), H^i_{t+1}, which is purchased at price q_t, depreciates at

[9]Fagereng et al. (2019) argue that the majority of movements in household wealth is due to asset price movements (i.e. capital gains or losses) rather than net saving. Households appear to treat capital gains differently from other forms of income and consume very little of these even if they are persistent.

[10]This assumption reflects the definition of household wealth (net worth, assets minus liabilities) in the Financial Accounts. First, household assets consist of non-financial assets and financial assets. Non-financial assets include real estate and consumer durables. Financial assets include checkable deposits and currency, time deposits and short-term investments, money market fund shares, debt securities, loans, corporate equities and mutual fund shares, life insurance reserves, pension entitlements, equity in non-corporate business, and

rate δ, and yields a rental rate ρ_t. The second type is financial assets (e.g. equities), S^i_{t+1}, which is purchased at price X_t and pays a dividend of d_t every year. The agent works l^i_t hours and receives an hourly wage of ω_t. The agent holds zero asset at t=ρ and leaves no bequests. Her budget constraint is given by

$$c^i_t + q_t H^i_{t+1} + S^i_{t+1} X_t = q_{t-1}(1-\delta)H^i_t + S^i_t(X_{t-1} + d_t) + \rho_t H^i_t + \omega_t l^i_t \tag{1.1}$$

where this budget constraint naturally implies that there are two distinct sources of income: labour income and capital income.

Definition 1. An agent i's total income, Y^i_t, equals labour income plus net rental income from non-financial assets plus dividends from financial assets.

$$Y^i_t = \underbrace{\omega_t l^i_t}_{\text{labour income}} + \underbrace{(\rho_t - \delta q_{t-1})H^i_t + S^i_t d_t}_{\text{(net) capital income}}. \tag{1.2}$$

If we sum equation (1.2) over all US residents, $\Sigma_{i \in US} Y^i_t$, Definition 1 coincides with the Bureau Economic Analysis (BEA)'s definition of national income, which approximately equals the national product (see Fox and McCully (2018) for more details).

Definition 2. An agent i's wealth gains, WG^i_t, equal the change in the price of the non-depreciated portion of the non-financial assets plus that of financial assets:

$$WG^i_t \equiv (1-\delta)H^i_t(q_t - q_{t-1}) + S^i_t(X_t - X_{t-1}). \tag{1.3}$$

Since wealth gains are not reflected in the current BEA's definition of national income, I incorporate these missing components into the budget constraint in the same way as other types of capital income. I define it as 'wealth-gains-augmented income'.[11]

Definition 3. At t, each agent i's wealth-gains-augmented income is heterogeneous with regard to her asset ownership. I classify each agent into either the rentiers ($i \in R$) or the work-

miscellaneous assets. Second, liabilities consist of loans (e.g. home mortgages, consumer credit, depository institution loans, and other loans and advances) and deferred & unpaid life insurance premiums.

[11]This concept is close to the Haig-Simons income suggested by Fagereng et al. (2019) and Robbins (2019). While the author calculates it as a concept of national wealth (i.e., an agent can directly own either capital or financial assets or both), I calculate wealth-gains-augmented income as a concept of household wealth (i.e., an agent cannot directly own capital). I also regard the use of wealth gains is more appropriate than the use of capital gains because while the market value of assets can be deviated from its replacement-cost, the value of capital is assumed to be the same as its replacement-cost in competitive markets.

ers ($i \in W$).[12] More specifically, the rentiers are able to accumulate assets and work at the same time, while the workers earn their income entirely from labour input. Furthermore, due to their monopoly of access to wealth residual, the rentier class can receive wealth gains (or losses) due to asset price changes while earning additional capital income at the expense of the workers due to the change in the rental rate ρ_t, of non-financial assets (e.g. housing rent).[13]

$$Y_t^i = \begin{cases} \omega_t l_t^i + (\rho_t - \delta q_{t-1})H_t^{i\in R} + S_t^{i\in R} d_t + WG_t^{i\in R} & \text{if } i \text{ is a rentier} \\ \omega_t l_t^i - \rho_t H_t^{j\in R} & \text{if } i \text{ is a worker.} \end{cases} \tag{1.4}$$

Equation (1.4) implies that the dynamics of wealth accumulation depend on difference in asset ownership ($i \in R$ or $i \in W$), heterogeneous rates of return on non-financial wealth, ρ_t, and financial wealth, d_t, and different degree of gains (or losses) from wealth residual, WG_t.

Definition 4. Aggregate household wealth, W_t, equals the market value of aggregate non-financial assets and aggregate financial assets, minus aggregate liabilities. Since, to a large extent, households' financial position ultimately represents a claim on non-financial assets of the whole economy[14], aggregate household wealth equals capital, K_t, plus wealth residual, WR_t, which is defined as the unexplained increase in wealth that is not accompanied by any increase in real output (Stiglitz, 2015a), such that

$$W_t = \Sigma_{i\in US} W_t^i = \underbrace{\Sigma_{i\in US} q_t^i H_t^i + \Sigma_{i\in US} S_t^i X_t^i}_{\text{All households}} = \underbrace{K_t + WR_t}_{\text{Whole economy}} . \tag{1.5}$$

Equation (1.5) implies that wealth residual is the increase in the market value of aggregate household wealth over and above its replacement-cost. In other words, wealth residual equals aggregate household wealth minus capital, $WR_t = W_t - K_t$.

[12] Historical data of the Survey of Consumer Finances (SCFs) compiled by the Federal Reserve has confirmed this: about 50% of the US households have owned almost zero or negative wealth during 1962-2016.

[13] A house is a long-lived asset that produces a flow of housing services. The market price of these services is the annual rent and the rent-to-house-price ratio is analogous to a dividend yield. However, unlike a stock or a bond, a change in housing rent result in a zero-sum game between landlords and tenants.

[14] Since financial wealth nets to zero across sectors in an aggregate sense, aggregate household wealth at a particular time can be represented as the sum of all of its non-financial assets (Holmquist and McIntosh, 2015). That is, at the end of the day, households own all non-financial assets (i.e., capital) of the other sectors through their holdings of equity shares or debt securities.

1.3.2 Measuring wealth residual

I measure wealth residual by combining the Financial Accounts of the Federal Reserve (FED) and the Fixed Assets Accounts of the Bureau of Economic Analysis (BEA). The level of aggregation is the national level, comprising all US residents. First, household wealth[15] is calculated by the market value of aggregate assets, net of their liabilities, from the Financial Accounts Table B.1 (FED code: FL152090005). Second, capital is calculated by the current-cost net stock of fixed assets[16] from the Fixed Assets Table 1.1 (BEA code: K1PTOTL1ES000). Accordingly, as seen in equation (1.5), wealth residual is naturally calculated by the gap between household wealth and capital.[17]

Figure 1.2 shows the result for measuring wealth residual for the US. The most notable result is a large increase in wealth residual after the late 1970s. In 2018, roughly 40% of household wealth was not related to production of real output but was instead associated with asset ownership, which was much larger than 8.8% in 1978.

Looking at broad trends in these time series, we can roughly divide them into two eras. In the first era, from 1959 to 1977, there were almost constant amount of wealth residual at around 0.8 trillion dollars, while capital was kept increasing. In the second era, between 1978 and 2018, there are rapid increases in wealth residual up to 41.4 trillion dollars in 2018, mainly due to housing and equity booms. Until 1977 the path of wealth followed that of capital, but wealth suddenly diverged from capital since then and has not come back yet. In this respect, the ownership of wealth residual may be a source of excessive returns (i.e. the returns over and above the return on capital in a competitive market) that drives the upsurge of wealth in the richest group.

Meanwhile, from equation (1.5), we can infer that the upsurge of wealth residual is being driven by either land, $q_t H_t$[18], or rent-related finance (i.e., some financial assets that is not linked to capital or land), $S_t X_t$, or both. First, I estimate a portion of land in wealth

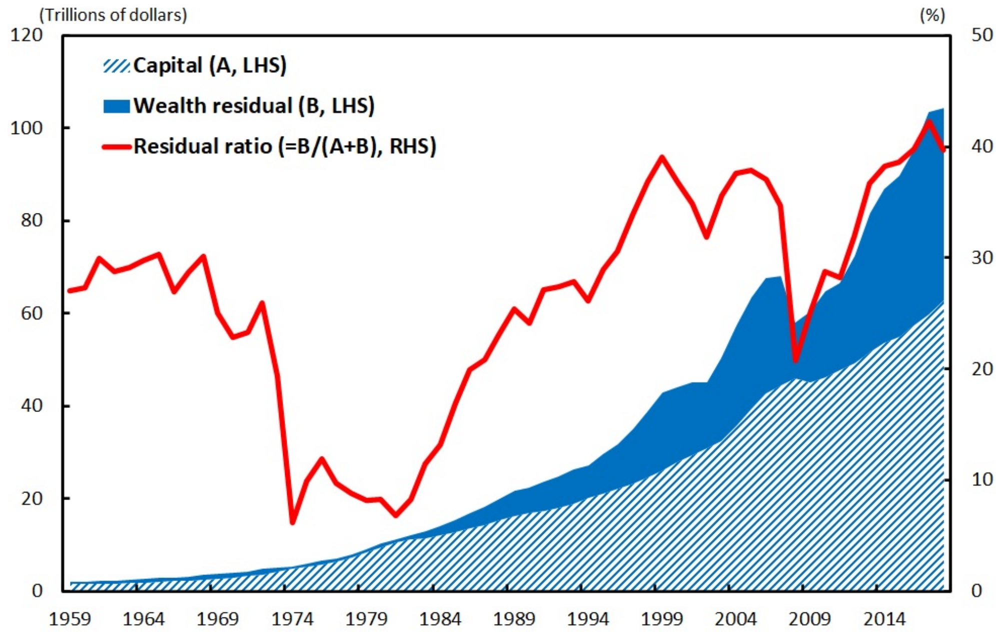
[15]For some tables, the Financial Accounts incorporate households' holdings and non-profit institutions' holdings. Since non-profit institutions held only 5.7% of the combined wealth in 2018, this paper refers to the combined wealth as simply household wealth.

[16]According to the BEA, buildings, machines, software, and even intangible capital or the intellectual property product (e.g. creation of a song) can be fixed assets, if they are used for a year or more in the production of goods or services. Houses and apartments are also included, even if the owner lives there.

[17]While a measure of capital relies on very precise assumptions on depreciation of historic capital goods (e.g. Collard-Wexler and De Loecker, 2016), wealth measured by market values may be close to a true indicator. So, the limitation of the wealth residual measure is related to how well the BEA measures capital. However, precise evidence on the significance of the measurement errors in BEA capital stock is missing until now, leaving us with an open empirical question.

[18]According to the Financial Accounts, non-financial assets consist of real estate and consumer durables. Since consumer durables are solely calculated by the replacement-cost value, there must be no wealth residual for consumer durables.

Fig. 1.2 Wealth residual, 1959-2018



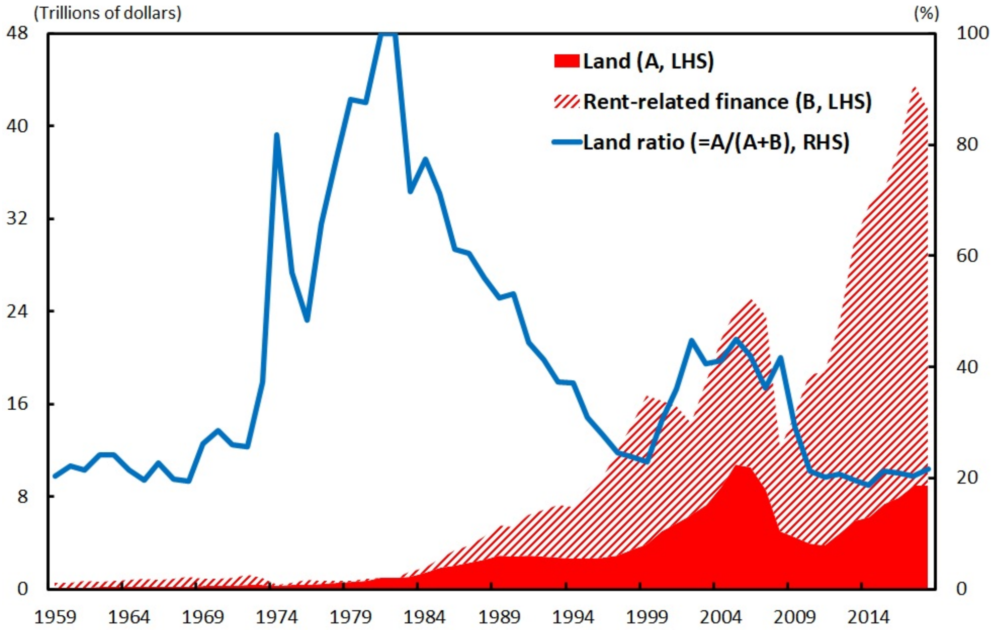
Source: The author's own estimation based on data from the FED and the BEA

residual by using a concept of Davis and Heathcote (2007), who see that a house can be seen as a bundle comprising a reproducible tangible structure and a non-reproducible plot of land.[19] Since the structure can be priced explicitly as the replacement-cost, the residuals are assumed to be the value of a home location – referred to as 'scarcity rents' that are evidently capitalised in non-financial assets, as argued by Stiglitz (2015a) and Rognlie (2015). Thus, I calculate wealth residual values for land by subtracting the replacement-cost value of structures from the market value of real estate from the Financial Accounts Table B.101 (FED code: LM155035005, LM155012605, and LM165013665). Second, rent-related finance is naturally calculated as the gap between wealth residual and land – referred to as 'monopoly rents' that are unobtrusively capitalised in financial assets, as argued by Stiglitz (2015d) and Brun and Gonzalez (2017).

Figure 1.3 shows the result for decomposition of wealth residual for the US. It depicts that by far the largest component of wealth residual is from rent-related finance, while land was a crucial source of wealth residual until the mid-2000s.

[19]If housing were simply a manufactured good, the price of housing would be determined by construction costs, and housing prices would increase at roughly the same rate as the price of other goods. But since housing has a valuable location component that is in limited supply, an increase in the demand for housing can link directly to increases in the value of good locations. Thus, for those who own many homes, rising house prices represents potential wealth gains which increases net wealth, but, for those who do not own their home, it means the higher rents in the rental market or the need for a higher propensity to save in order to buy housing, as it becomes necessary to make a larger deposit for a mortgage.

Fig. 1.3 Decomposition of wealth residual, 1959-2018



Source: The author's own estimation based on data from the FED and the BEA

1.3.3 Who gains (or loses) from wealth residual?

A question naturally arises from the above data: who gains or loses from the rapidly growing wealth residual? I attempt to find a clue from the Distributional National Accounts compiled by Piketty et al. (2018).[20] By combining Table E2, E2b, E2c, and E3 of this statistic, I estimate the shares of wealth (i.e., assets minus liabilities) and the average wealth by wealth group for each asset composition for the US households. All our data are expressed in constant 2018 dollars to correct for inflation, using the national income deflator.

Table 1.1 shows the result of estimated percentage of wealth held by the bottom 50%, the middle 40%, the top 10%, the top 1%, the top 0.1%, and the top 0.01% for each asset composition – equities, non-corporate business assets, real estate, and other assets (deposits, bonds, pensions, etc.). In 2015, the top 10% hold 71.9% of aggregate wealth, while the bottom 50% hold only 0.2% despite having vastly more people. Moreover, the top 10% own 88.6% and 72.0% of equities and non-corporate business assets, while the bottom 50% own only 1.4% and 1.9%, respectively. When we restricted to real estate, it looks quite different,

[20] Data files and replication files are available at http://gabriel-zucman.eu/usdina/.

with the top 10% holding significantly less as a share (58.8%), while the middle 40% are doing better than those for equities or non-corporate business assets. If we limit ourselves to the very wealthy, the top 1% hold about 61.7% and the top 0.01% (only 30 thousand individuals) own almost 20% of the whole equities.

Table 1.1 Share of wealth held by wealth group for each asset composition, 2015

(%)

	Equities	Business	Real estate	Others	Total
Bottom 50%	1.4	1.9	0.5	-0.7	0.2
Middle 40%	9.9	26.0	40.7	30.4	27.9
Top 10%	88.6	72.0	58.8	70.3	71.9
Top 1%	61.7	38.7	16.9	34.8	37.1
Top 0.1%	37.5	16.8	4.5	17.7	18.9
Top 0.01%	19.9	6.4	0.9	9.4	9.5

Notes: Equities, corporate equities including S-corporation and money market shares; Business, non-corporate business assets including sole proprietorships, farms, partnerships, and intellectual property products; Real estate, owner- and tenant- occupied housing, net of mortgage debt; and Others, fixed income claims (i.e., currency, deposits, bonds, and other interest-paying assets, net of non-mortgage debts), pensions, and insurances.
Source: The author's own calculation based on the Distributional National Accounts

Table 1.2 shows the estimated average wealth (in constant 2018 dollars) held by wealth group for each asset composition. In 2015, on average, the bottom 50% own only 1.5 thousand dollars, while the top 10% own 2.3 million dollars, approximately 1,533 times larger than the bottom 50%. If we look at the super-wealthy, the top 0.01% hold 304.4 million dollars on average[21], about 202,335 times larger than the bottom 50%. For each asset composition, equities, non-corporate business assets, and real estate owned by the top 10% were much larger than the bottom 50%, approximately 312 times, 185 times, and 576 times, respectively. If we see the super-wealthy, equities, non-corporate business assets, and real estate held by the top 0.01% were extremely larger than the bottom 50%, approximately 69,905 times, 16,401 times, and 8,550 times, respectively. These results support the view that the rapid growth of wealth residual, as we have seen in Figure 1.2 and Figure 1.3, was clearly beneficial to the top wealthy, especially in equities and non-corporate business assets for the top 0.01% and real estate for the top 10%, as in line with Kuhn et al. (2020).

Table 1.3 describes the estimated annual growth rates of average wealth by wealth group

[21] According to the Forbes, in 2019, there are 2,153 billionaires in the US. The wealthiest person throughout the world is the founder and chief executive officer of Amazon, Jeff Bezos who owns 131 billion dollars (see https://www.forbes.com/billionaires/#6f15298d251c).

Table 1.2 Average wealth held by wealth group for each asset composition, 2015

(Thousands of dollars)

	Equities	Business	Real estate	Others	Total
Bottom 50% (A)	1.8	1.4	0.6	-2.3	1.5
Middle 40%	15.7	23.0	63.4	121.5	223.7
Top 10% (B)	560.5	254.6	366.7	1,123.7	2,305.4
Top 1% (C)	3,902.6	1,369.2	1,052.0	5,566.2	11,889.9
Top 0.1% (D)	23,695.8	5,927.9	2,806.4	28,284.1	60,714.2
Top 0.01% (E)	125,585.9	22,563.5	5,447.2	150,765.9	304,362.6
B/A	312.0	185.1	575.5	†	1,532.6
C/A	2,172.3	995.2	1,651.2	†	7,904.2
D/A	13,189.9	4,308.7	4,404.8	†	40,361.8
E/A	69,905.3	16,400.5	8,549.7	†	202,335.1

Notes: See notes in Table 1.1. † implies meaningless values.
Source: The author's own calculation based on the Distributional National Accounts

for each asset composition. I divide the 1962-2015 period, a time range imposed by the availability of data, into two periods: pre-1978 versus post-1978. During the pre-1978 period, the middle 40% and the top 10% showed an increase of 1.5% and 0.2%, respectively, while the top 0.01% and the bottom 50% saw their net worth shrink by -0.9% and -0.7%, respectively. In contrast, during the post-1978 period, even though all wealth group experience wealth growth except for the bottom 50%, the wealthiest group see their wealth increase the fastest, mainly due to gains from equities and non-corporate business assets. These results imply that each individual's returns on wealth is heterogeneous with regard to her asset composition, even correlated with the amount of wealth that she owns, as empirically shown by the recent studies (Fagereng et al. (2016, 2020, 2019)– Norway; Bach et al. (2016)– Sweden; Cao and Luo (2017)– US).[22]

Finally, we can ask ourselves why unequal ownership of wealth residual emerges. As argued by a number of studies (e.g. Lusardi et al., 2017; Pagano, 2014; Stiglitz, 2015a,b,c,d,e), it is closely related to a monopoly of access (i.e., entry barrier) to various types of property. For instance, once we assume that not everyone has equal access to knowledge, knowing something slightly before others do can yield large returns (the so-called "insider trading" or "asymmetric information"). Accordingly, a few individuals would benefit from superiority in

[22] By using the Panel Study of Income Dynamics recompiled by Cao and Luo (2017), Lee (2020) shows that, during the 1984-2011 period, households at the top 10% and the top 5% earn an annualised return (including capital gains) above 39% and 70%, respectively, while there are 50% of households who earn a low return close to zero or below.

Table 1.3 Annual growth rates of wealth for each asset composition

(%)

	Pre-1978: 1962-1977					Post-1978: 1978-2015				
	Total	Equities	Business	Real estate	Others	Total	Real estate	Equities	Business	Others
Bottom 50%	-0.7	-0.7	-1.5	2.4	-3.0	-1.7	1.6	5.2	-1.2	†
Middle 40%	1.5	-0.5	0.0	1.2	3.2	2.0	2.5	1.5	0.8	3.1
Top 10%	0.2	-3.2	1.5	1.3	0.5	3.0	5.0	1.0	1.4	3.8
Top 1%	-0.5	-3.8	3.2	2.0	-0.7	4.2	5.7	2.2	0.9	5.7
Top 0.1%	-0.8	-3.9	6.2	2.3	0.2	5.7	6.5	3.9	0.7	7.4
Top 0.01%	-0.9	-3.7	8.6	2.5	1.4	7.0	7.4	4.7	0.2	8.9

Notes: See notes in Table 1.1. † implies meaningless values.
Source: The author's own calculation based on the Distributional National Accounts

financial knowledge, as financial markets have become more complex.[23] Furthermore, given that financial firms that offer these complex products require high minimum investments, as wealth increases, an individual's ability to absorb risk also increases, enabling the wealthy to take on more high-risk and high-return products (e.g. Saez and Zucman, 2016).[24] According to Shakhnov (2014), PEF typically has horizons of 10-13 years with the high minimum required commitment (at least a median of $1 million for funds of $100 million) but at most 5% of US households can afford to invest in it. Of course, this long-term horizon and high entry costs which erect high entry barrier are compensated for by substantially higher returns. For instance, the annualised return on PEF (9.2 percent) was much higher than those for stock (3.2 percent), T-bond (0.4 percent) and than the rate of inflation (2.4 percent) during 1997-2011.

More fundamentally, the growing influence of corporate lobbying to defend unproductive rents (e.g. Drutman, 2015; Esteban and Ray, 2006), such as lobbying for deregulations for activities with high entry barriers or for tax cuts for the wealthy, can create a *persistent* monopoly of access to wealth residual. According to the standard economic theory, competitive forces should limit outsized profits, but, if the government does not ensure that markets

[23] Some people may argue that the higher returns earned by the wealthy seem to be mainly driven by exceptional investment skills rather than by privileged access to private information. However, we can refute this argument once we assume that there is no 'entry barrier' to these complex financial activities other than 'natural' differences in people's intelligence in perfect capital market where clever poor people can borrow money to be educated in these things. In reality, most global banks, such as Barclays, BNP Paribas, and Citibank, have a separate business unit with dedicated teams of client advisors and product specialists exclusively for the wealthy. They provide a wide range of investment opportunities, such as private equity finance (PEF).

[24] Sheikh and Qiao (2009), based on ten years of monthly returns from 1998 to 2008, show that the mean and standard deviation of PEF (13.09%, 23.36%p) and hedge funds (8.12%, 6.49%p) are much higher than bonds (5.59%, 3.58%p).

are competitive, there can be an opportunity for large rents. For example, lobbying spending of the US corporations during 1998-2017, obtained from the Center for Responsive Politics (CRP), confirms this. Until the global financial crisis in 2008, the financial sector was the top spender on lobbying and has invested almost $8 billion in political influence purchasing over the last two decade. Also, the 'lobbying intensity' that is defined as the ratio of the portion of lobbying spending to that of sectoral GDP in the financial sector (0.74) is much greater than that of the non-financial sector (0.58), including manufacturing, retail trade and wholesale trade. Therefore, the resulting deregulations on financial industries have provided fertile ground for pursuing persistent rent-related activities.[25]

1.3.4 Wealth residual hypothesis

Based on Stiglitz (2015d)'s dynamic model that incorporates a concept of wealth residual, this paper presents three hypotheses that will be empirically examined in Section 4, 5, and 6.

Hypothesis 1: A fast-growing unearned income from wealth residual attributes to the co-evolution of growing inequality and falling growth.

Assume that there is wealth residual in fixed supply that has the characteristic of positional goods (e.g. land), $\bar{T}$, that does not affect production of conventional goods. All the wealth of the economy is held by the rentiers, as we have already seen in Table 1.1. The demand for wealth residual by the rentiers is given as $M(a,p)$. In equilibrium,

$$M(a,p) = p\bar{T} \tag{1.6}$$

where $a = K + p\bar{T}$ and p is price of wealth residual. Therefore from equation (1.6), p can be solved as a function of a. K can then be rearranged as

$$K = a - p(a)\bar{T}. \tag{1.7}$$

[25] The rise in top incomes in the financial sector has confirmed this. An empirical study by Philippon and Reshef (2012) shows that in the past two decades workers in the financial industry have received a 50% of wage-premium with respect to workers in other sectors. Furthermore, the wage-premium for top executive compensation in finance reaches 250%. According to their estimates, financial sector compensations have been about 40 percent higher than the level that would have been expected under perfect competition.

Lastly, differentiating equation (1.7) with regard to a yields:

$$\frac{dK}{da} = 1 - p'\bar{T} = 1 - \frac{M_a}{1 - M_p} < 1 \tag{1.8}$$

where M_a is the wealth elasticity of the demand for wealth residual and M_p is the price elasticity of the demand for wealth residual.

Equation (1.8) gives the following implication. If M_a is sufficiently large and M_p is sufficiently small, an increase in a may be associated with a decrease in K (i.e., $\frac{dK}{da} < 0$). These conditions for M_a and M_p can be interpreted as '*congestion and crowding*', which were firstly presented by Hirsch (1977).[26] More specifically, the emergence of congestion and crowding for wealth residual can create the increase in 'scarcity rents', leading to the upsurge of wealth residual (i.e., it will look like arms races).[27]

Hence, when the increase in wealth is largely due to an increase in the value of wealth residual, the economy can experience growing inequality and falling growth at the same time. This is because since the rentiers hold all wealth residual, the increase in the value of wealth residual affects those cohorts only. In contrast, the workers, who do not own any wealth residual, are only affected to the extent that the growth in wealth residual values crowds out capital accumulation, leading to lower growth. Moreover, if we add a spill-over effect to the workers, an increase in scarcity rents might disadvantage the workers further.[28]

Hypothesis 2: On bubble paths, wealth increases as capital diminishes.

In equilibrium, the return to holding wealth residual and the return to holding capital must be the same. Since wealth residual is assumed to be non-productive, its entire return is its wealth gain, $\frac{\dot{p}}{p}$. Then the equilibrium condition is given as

$$\frac{\dot{p}}{p} = F_K - \delta \tag{1.9}$$

[26]Hirsch (1977) coined the concept of 'positional goods', explaining that the positional economy is composed of all aspects of goods, services, work positions and other social relationships that are either scarce in some absolute or socially imposed sense or subject to congestion and crowding through more extensive use.

[27]For instance, let us imagine that rich individuals compete for houses in certain fashionable parts of New York. As the wealthy get wealthier, they compete more vigorously for this real estate, and the prices of those assets with fixed supply increase, without any increase in real output.

[28]For instance, in an economy that landlords rent out some of their land to workers at a rental price of pF_K, the increase in pF_K reduces workers' consumption and savings at the same time.

where F_K is the return to capital and δ is the depreciation rate. Assume that the rentiers save a fixed fraction, s, of full net income including wealth gains for simplicity[29]:

$$\dot{K} + \dot{p}\bar{T} = s(F_K K - \delta K + \dot{p}\bar{T}). \tag{1.10}$$

Combining equation (1.9) and equation (1.10) yields,

$$\dot{K} = s(F_K K - \delta K) - (1-s)\dot{p}\bar{T} = (sK - (1-s)p\bar{T})(F_K - \delta). \tag{1.11}$$

Accordingly, we can obtain a pair of differential equations that describe the dynamics of this economy, as depicted in Figure A.1. The steady states, K^* and p^*, are calculated by the solution to the loci $F_K = \delta$ and $p = \frac{sK}{(1-s)}$. Note that any value of p along $K = K^*$ is an equilibrium since $\dot{K} = 0$ when $K = K^*$. More specifically, to the right of $K = K^*$, p is decreasing and to the left it is increasing. Above the $\dot{K} = 0$ locus, but to the left of K^*, K is decreasing, while above the $\dot{K} = 0$ locus, to the right of K^*, K is increasing. Conversely, below the $\dot{K} = 0$ locus, but to the left of K^*, K is increasing, while below the $\dot{K} = 0$ locus, to the right of K^*, K is decreasing.

K^* in combination of any value $p < p^* \equiv \frac{sK^*}{(1-s)}$ is a stable equilibrium, while K^* in combination of any value of $p \geq p*$ is an unstable equilibrium. The saddle point trajectory EE^* divides the bottom quadrant into a convergent and non-convergent region. Below EE^*, paths converge to $K = K^*$. In contrast, above EE^*, they diverge (path 1). As a trajectory below the $\dot{K} = 0$ locus and to the left of K^* approaches K^*, the slope is $\frac{dp}{dK} \approx \frac{p}{sK-(1-s)p}$ which is finite below the locus $p = \frac{sK}{1-s}$. Hence, trajectories hit the vertical axis, at which point they remain in the steady state (path 2). We can similarly show that if $K_0 > K^*$, K will also hits the vertical axis; but if the initial value of $p > \frac{sK}{1-s}$, K will initially increase before decreasing to K^* (path 3).

In summary, there are an infinity of stable equilibria, in all of which the level of income is the same in the equilibrium but in which there can be markedly different values of wealth residual $p\bar{T}$. That is, p is in this sense fully indeterminate. But if $K < K^*$ and the initial p is too high, the economy can suffer from asset bubbles. Those bubbles can lead to an increase in household wealth along with possible adverse effects on the accumulation of productive capital.

[29] Stiglitz (2015d) demonstrates that even in a generalised savings function $s(K, p, \dot{p})$, the results are similar. However, since the aggregate savings may differ depending on the composition of wealth, the generalised model can further show the increase in the value of wealth residual endogenously crowds out capital accumulation.

Hypothesis 3: The government's regressive policies on rent-related activities are associated with the increase in wealth residual.

The first policy is taxation on wealth residual. The demand for wealth residual depends on the price of wealth residual, p, and its opportunity cost, the return on capital r. If wealth residual is taxed, the cost of owning wealth residual, μ, becomes

$$\mu = r + \rho^T p - (1 - \rho^{wg})\dot{p} \tag{1.12}$$

where ρ^T and ρ^{wg} are the tax rate on scarcity rents and wealth gains, respectively, associated with wealth residual. Then the demand for wealth residual can be illustrated as $M(a, p, \mu)$. So incorporating taxation on wealth residual into equation (1.7) yields,

$$K = a - p(a, \rho^T, \rho^{wg}). \tag{1.13}$$

Since $p = M(a, p, \mu)$, given that the cost elasticity of the demand for wealth residual $M_\mu < 0$,

$$\frac{da}{d\rho^T} = \frac{dp}{d\rho^T}\bar{T} = \frac{pM_\mu \bar{T}}{1 - M_a - M_p - \rho^T M_\mu} < 0 \tag{1.14}$$

if expectation about wealth gains are assumed to be fixed with a natural stability condition that ensures that the denominator is positive.

Similarly, if we assume $(1 - \rho^{wg})\frac{\partial \dot{p}}{\partial \rho^{wg}} \approx 0$,

$$\frac{da}{d\rho^{wg}} = \frac{dp}{d\rho^{wg}}\bar{T} = \frac{\dot{p}M_\mu \bar{T}}{1 - M_a - M_p} < 0. \tag{1.15}$$

These two conditions imply that, at any given K, the higher ρ^T and ρ^{wg} are associated with the lower wealth residual because the higher tax rate on wealth residual (including wealth gains) can increase the user cost of wealth residual and, thus, help align capital and wealth.

The second policy is credit supply. Assume that there are three distinct classes of individuals in this economy - the workers, the capitalists who save out of profits and invest only in capital goods without access to credit, and the rentiers who own wealth residual. In this economy, a bank provides credit based on collateral. The mechanism is simple: when the price of wealth residual, p, goes up, the bank is willing to lend more. If the bank is willing to lend more, the price of wealth residual will go up further. The demand for wealth residual is given by

$$M(a^R, c, p) = p\bar{T} = a^R + c \tag{1.16}$$

where c is the available amount of credit provided by a bank and a^R is the wealth of the rentiers, which is the value of wealth residual minus what they owe in credit ($a^R = p\bar{T} - c$). Rearranging equation (1.16) yields

$$M(a^R, c, \frac{a^R + c}{\bar{T}}) = a^R + c. \tag{1.17}$$

We can solve equation (1.17) for $p = \Phi(c)$. Hence the wealth of the rentiers is entirely driven by the provision of credit

$$a^R = p\bar{T} - c = \bar{T}\Phi(c) - c. \tag{1.18}$$

To close this model, we need an additional equation that describes capital accumulation. Following Stiglitz (2015d), I take the simplest version of Kaldor (1957) where capitalists save a fraction of their income, s_p, putting their money into capital goods as

$$\dot{K} = s_p rK - \delta K. \tag{1.19}$$

In steady state,

$$F_{K^*} = \frac{\delta}{s_p}. \tag{1.20}$$

From equation (1.20), we can see that the provision of additional credit has no effect on capital in equilibrium. In contrast, combining equation (1.17) and the sum of the wealth of the rentiers and the capitalists, $a = a^R + K$, yields,

$$\frac{da}{dc} = \frac{da^R}{dc} = \frac{-(1 - M_c - M_p/\bar{T})}{1 - M_{a^R} - M_p/\bar{T}} = \bar{T}\Phi_c - 1 = \frac{M_c - M_p}{1 - M_{a^R} - M_p} > 0 \tag{1.21}$$

where this result implies that an increase in credit supply by a central bank (i.e. monetary expansion) can increase wealth through an increase in the price of wealth residual without any effect on the capital formation. Since it is only the rentiers who own wealth residual with their monopoly of access to credit, all of the increase in wealth goes to the rentiers. That is, monopoly rents capitalised in rent-related finance leads to the upsurge of wealth residual.

The last policy is financial regulation (e.g. loan-to-value ratio) set by the financial authority. If we assume that a bank only provides credit with wealth residual as collateral but provides it at zero interest rate, so that owners of wealth residual borrow as much as they can, the financial authority limits the amount of credit that is made available. As more credit is provided, the price of wealth residual will be bid up. In equilibrium,

$$c = \alpha p \bar{T} \tag{1.22}$$

where α is the collateral requirement. Therefore the regulation on the collateral requirement directly affects the amount of credit supply (i.e., $\frac{dc}{d\alpha} > 0$) and thus the price of wealth residual since $\frac{da}{d\alpha} > 0$ in equation (1.21).

1.4 Empirical Strategy

1.4.1 Data

To test the wealth residual hypothesis, I consider a US sample consisting of 51 states (including Washington, D.C.), with annual data spanning the 1978-2015 period (that is, N=51 and T=38), a time range imposed to establish a balanced panel. The use of the US state-level data is much superior to the use of cross-country data in assessing the empirical relationship among economic variables. First, cross-state heterogeneity in a country (the US, in this case) is much smaller than cross-country heterogeneity. Accordingly, it allows us to avoid inherent problems in postulating a common parametric structure for cross-country samples that lie in different stages of economic conditions and are also characterised by other heterogeneous structural differences. Also, using the US state-level data not only mitigates low-quality and comparability issues from the use of cross-country data but also helps alleviate significant problems of endogeneity, omitted variables, and measurement errors often generated from cross-country analysis (Banerjee and Duflo, 2003; Barro, 2000). Finally, the use of the large cross section (N) and time series (T) allow us to segregate the data into various groups of interest without the issues of sample representativeness or degrees of freedom. Table 1.4 reports the summary statistics of the variables used in this study.

First, for main variables, I use the growth rate of real per capita personal income in constant 2012 US dollar, taken from the BEA, as a measure of growth (denoted by 'Y')[30]. Personal income refers to all earnings from wages, investments, and profit sharing from

[30] There are two possible measures of growth at state-level from the BEA: personal income and regional GDP. To compare the results of the pre-1978 period and those of the post-1978 period in Section 4 and Section 5, I choose personal income since regional GDP is only available from 1977.

businesses received by all individuals in a country. Second, I measure inequality ('INQ') using the share of the top 1% income (pre-tax national income), drawn from the updated Frank-Sommeiller-Price Series from the US State-Level Income Inequality Database (Frank, 2009)[31], which provides a panel of distributional indices at the state level for the long period. Third, as the measure of wealth residual ('WR'), I use the share of land in the market value of real estate, assuming that rent-related finance expands in proportion to the increase in the value of land, as in equation (1.22).[32] I calculate the value of land by subtracting the replacement-cost value of structures from the market value of real estate, taken from the Lincoln Institute of Land Policy (LILP), as elaborated on this in Section 1.3.2.

Next, to control the important factors behind the growth-inequality relationship, I consider each of the following indicators. First, human capital, measured by the educational attainment embodied in a worker, is a major determinant of the worker's lifetime earnings (e.g. Barro and Lee, 2013; Lee and Lee, 2018). For the measure of human capital (denoted by 'EDU'), I use the college attainment (i.e., the total number of college graduates divided by the total state population) that Frank (2009) constructed. Second, the evolution of top income inequality partly relates to innovation (e.g. Aghion et al., 2019). Indeed, if the increase in top income inequality has been pervasive across occupations, it has particularly affected occupations that appear to be closely related to innovation, such as entrepreneurs, engineers, scientists, as well as managers. For the measure of innovation ('INO'), I use the flow number of utility patents (i.e., patents for invention) per million of inhabitants, drawn from the United States Patent and Trademark Office (USPTO).

To control for a possible spurious correlation between wealth residual and top income inequality (or growth) through rent-related policies,[33] I also add the following variables: (i) taxation, measured by the maximum rate of tax on ordinary income (including short-term capital gains) ('TAXw') and long-term capital gains ('TAXk') drawn from the National Bureau of Economic Research[34]; (ii) structural change ('SC'), measured by the employment share of manufacturing in private industry drawn from the BEA; (iii) credit supply ('MOR'),

[31] Pre-tax national income is the sum of all personal income flows accruing to labour and capital, after taking into account the operation of private and public pensions, as well as disability and unemployment insurance, but before taking into account other taxes or transfers.

[32] In specific, I have considered that a change in the price of land would be different in each state while a change in the price of rent-related finance would be identical across the whole country. Admittedly, to be more robust, we need to disaggregate the Financial Accounts of the US by every state. However, this goes beyond the scope of our study and is left to future research.

[33] For instance, as lower taxes in Texas lead to both higher top incomes and higher wealth residual because the rich in California are fleeing to Texas.

[34] While the actual income tax rate on a taxpayer is endogenous (i.e. it depends upon his income), variation in state tax laws across states is exogenous to individual decisions in labour supply and asset purchase (see (Feenberg and Coutts, 1993, updated from the authors up to 2018)) or https://users.nber.org/∼taxsim/state-rates/).

Table 1.4 Summary statistics

	Unit	Obs.	Mean	Std.	Min.	Max.
A. Main variables						
Income growth, Y	%	1,938	5.1	3.4	-9.8	29.9
Income inequality, INQ	%	1,938	14.7	4.7	4.8	36.1
Wealth residual, WR	%	1,938	23.3	18.1	5.0	81.9
B. Control variables						
Human capital, EDU	%	1,938	15.7	5.2	6.4	45.9
Innovation, INO		1,938	212.2	176.3	17.9	1,363.8
Tax on ordinary income, TAXw	%	1,938	43.7	7.9	28.0	75.9
Tax on capital gains, TAXk	%	1,938	24.9	5.5	15.0	37.0
Structural change, SC	%	1,938	17.9	8.5	.3	41.6
Credit supply, MOR	%	1,938	8.1	2.9	3.6	17.9
Financial regulation, LTV	%	1,938	76.7	4.3	25.7	96.7
Housing supply, HOU	%	1,938	67.3	7.1	34.6	81.3
Government size, GE	%	1,938	13.7	2.4	7.1	26.9
Welfare, WE	%	1,938	4.4	1.3	1.6	10.3
Labour bargaining power, UNI	%	1,938	14.2	6.9	2.0	38.3
C. Robustness check						
Gini, INQa	%	1,938	56.3	5.1	43.9	71.1
Atkinson, INQb	%	1,938	24.8	4.6	16.3	41.1
Top 10% share, INQc	%	1,938	40.4	5.8	23.2	62.2
Top 0.1% share, INQd	%	1,938	6.4	3.1	1.1	23.4
Land price, WRa	%	1,938	88.5	114.5	.0	1,668.1

Notes: Each variable denotes the following data: Y, the growth rate of real per capita personal income (2012 US dollar); INQ, the share of the top 1% income; WR, the share of land in the market value of real estate; EDU, college attainment; INO, the flow number of patents per million of inhabitants; TAXw, the maximum rate of tax on ordinary income (including short-term capital gains); TAXk, the maximum rate of tax on long-term capital gains; SC, the employment share of manufacturing in private industry; MOR, the average mortgage rate; LTV, the loan-to-value ratio; HOU, the homeownership rate; GE, the state government expenditure as a percentage of GDP; WE, the state government expenditure on health care and welfare as a percentage of GDP; UNI, the union membership density; INQa, the Gini index; INQb, the Atkinson index; INQc, the share of the top 10% income; INQd, the share of the top 0.1% income; and WRa, the land price index (2007=100).

measured by the average mortgage rate taken from the Federal Housing Finance Agency (FHFA); (iv) financial regulation ('LTV'), measured by the loan-to-value ratio taken from the FHFA; and (v) housing supply ('HOR'), measured by the homeownership rate (i.e. the percentage of homes that are occupied by the owner) drawn from the Census Bureau (CB).

For additional environmental variables, I also add (vi) government size ('GE'), measured by the state government expenditure as a percentage of GDP taken from the CB; (vii) welfare ('WE'), measured by the state government expenditure on health care and welfare as a percentage of GDP taken from the CB; and (viii) labour bargaining power ('UNI'), measured by the union membership density drawn from the BLS.

Lastly, for my robustness analysis in Section 5.3, I use the following variables as alternative measures for inequality and wealth residual: Gini index (denoted by "INQa"), Atkinson index ('INQb'), the share of the top 10% income ('INQc') and the share of the top 0.1% income ('INQd') from (Frank, 2009, updated from the author up to 2015), and land price index (2007=100, denoted by 'WRa') from the LILP.

1.4.2 Empirical evidence from state-level data

Figuratively speaking, there are 51 labs in the US. For the past four decades, each state has been heading in different directions, creating social experiments that reveal whether the US works better as a low-tax, low-regulation place in which government makes little provision for its citizens (e.g. Texas), or as a high-tax, high services, highly regulated one in which leans heavily on its affluent residents to fund a social-safety net (e.g. California).[35] These experiments allow us to conduct a preliminary inspection of the panel data at state-level because the observed dynamics should at least provide some evidence of such an underlying relationship between wealth residual and inequality, and growth in a large enough sample during the post-1978 period, if the wealth residual hypothesis is correct.

Figures A.2, A.3, and A.4 compare the heterogeneous changes in the economic outcomes of wealth residual, inequality, and growth across states during the pre- and post-1978 periods as a result of different political experiments.

As described in Figure A.2, the average value of land per housing unit – average depreciated value of housing minus average cost of the structure of housing – as a proxy for wealth residual, has significantly increased in all states, but shows remarkably different patterns with regard to their locations. In 1978, the average values of land in most states were very low at below one thousand dollars except in Hawaii (55.9 thousand dollars), Delaware (14.6 thousand dollars), and Alaska (14.2 thousand dollars). At that time, we can say that house prices fully reflected construction costs, meaning that as a house was no different from a manufactured good. However, after four decades, the average values of land in most states soared up, although there are huge differentials among states. For example, the average

[35]The Economist, "Special Report on California & Texas", 2019.6.20.

value of land in Washington, D.C. has increased from 9.2 thousand dollars in 1978 to 723.2 thousand dollars in 2015, while the average value in Alaska has remained stable at around 15.0 thousand dollars.

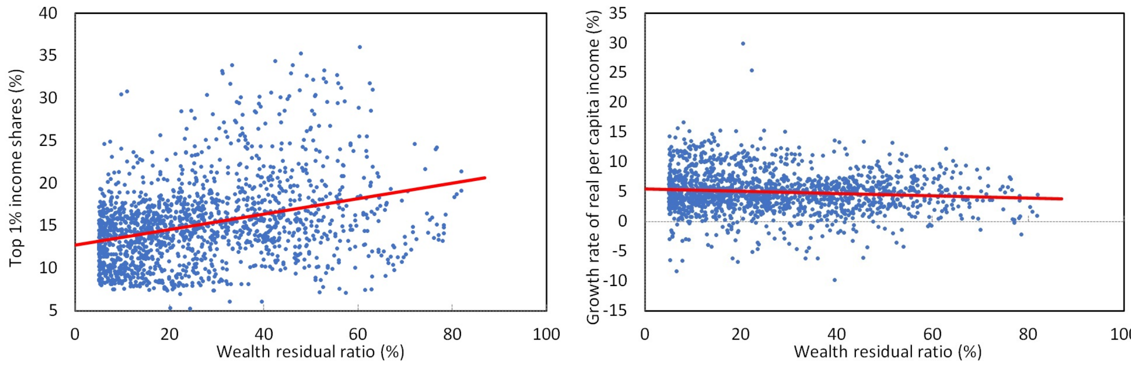
Meanwhile, as depicted in Figure A.3, inequality became worse in all states, but it showed relatively more heterogeneous co-movement between states. The average share of the top 1% income in 1978 was 8.6 percent, ranging from 4.8 percent in Alaska to 10.7 percent in Florida, while the average share was 18.1 percent in 2015, ranging from 11.4 percent in Alaska to 32.9 percent in New York.

In contrast, as shown in Figure A.4, the growth rates of real per capita personal income sharply declined in all states, showing relatively more homogeneous co-movement between states. Real per capita personal income grew 7.0 percent from 1959 to 1978, ranging from on average 5.8 percent in Delaware to 8.0 percent in North Dakota, while it grew 4.8 percent from 1978 to 2015, ranging from 4.0 percent in Alaska to 5.4 percent in Massachusetts.

Basic stylised facts can provide some preliminary insights regarding whether the data support the wealth residual hypothesis. Figure 1.4 plots the contemporaneous correlations between wealth residual and inequality and between wealth residual and growth, based on pooling samples of 51 states over the period of 1978-2015. Such basic correlations show evidence of the significantly positive relationship between wealth residual and inequality (correlation coefficient = 0.33): the extraordinary growth in the top 1% incomes has coincided with the upsurge of wealth residual. So, if the existence of a positive relationship between wealth residual and inequality persists with more refined estimation techniques, this could strongly support the wealth residual hypothesis that the sudden increase in wealth residual has led to higher inequality since the late 1970s. In contrast, Figure 1.4b gives evidence of the weakly negative relationship between wealth residual and growth (correlation coefficient = -0.10).

Therefore, in conducting more refined estimation using the same dataset, I attempt to carefully control for other important determinants of inequality and growth, chosen from the relevant literature (e.g. Banerjee and Duflo, 2003; Barro, 2000; Forbes, 2000; Perotti, 1996) to minimise omitted-variables and measurement-error problems. More specifically, for control variables, I consider human capital, innovation, taxation, structural change, credit supply, financial regulation, housing supply, government size, welfare, and labour bargaining power, as explained in Section 4.1. Also, to capture the Kuznets (1955) inverted-U curve for the relationship between inequality and the level of national income, I control for per capita GDP and its square in the inequality specification. The following equations represents the

Fig. 1.4 Correlations between wealth residual and inequality, and growth



(a) Wealth residual and inequality (b) Wealth residual and growth

Note: Wealth residual ratio refers to share of land in the market value of real estate

empirical framework:

$$\begin{aligned} INQ_{i,t} &= \alpha_0 + \alpha_1 WR_{i,t-1} + \alpha_2 GDP_{i,t-1} + \alpha_3 GDP^2_{i,t-1} + Z'_{i,t-1}\gamma + \lambda_i + \theta_t + \varepsilon_{i,t} \\ Y_{i,t} &= \beta_0 + \beta_1 WR_{i,t-1} + Z'_{i,t-1}\delta + \lambda_i + \theta_t + \varepsilon_{i,t} \end{aligned} \tag{1.23}$$

where INQ, a measure of inequality; Y, growth; WR, wealth residual; GDP, the logarithm of real GDP per capita; Z, a vector of control variables; λ, state-fixed effects; θ, time-fixed effects; and ε, the error term.

By including state- and time-fixed effects, we can control for permanent cross-state differences in the dependent variable (INQ or Y) and macroeconomic changes, which might affect INQ or Y in any period, respectively. I estimate the two equations in (1.23) by taking out period averages of variables since this fixed effects (FE) estimation allows us to control for possible time-invariant omitted-variable bias. To mitigate the concerns that inequality (or growth) dynamics feeds back to wealth residual (i.e., reverse causality), the relevant explanatory variables are measured at $t-1$.[36]

Tables 1.5 and 1.6 show estimation results. Column 1 in each table reports the result of random effects estimation ('RE') that unobserved variables are assumed to be uncorrelated with all explanatory variables[37], while column 2 represents the result of FE estimation without time-fixed effects. Column 3 presents the results of our basic specification, while

[36]To remove the possible endogeneity of our wealth residual measure, we need to consider instrumental variable estimation techniques; however, it is practically difficult to find a set of fully convincing exogenous instruments. Subsequent research on this issue will be required.

[37]I have performed the Hausman test (see Cameron and Trivedi (2010) for detailed information). It shows strong rejection of the null hypothesis that RE provides consistent estimates.

columns 4, 5, and 6 add channels of human capital, innovation, and both, respectively. The estimates consistently show that α_1 are positive and statistically significant in the inequality specification while β_1 are negative and statistically significant in the growth specification, when controlling for other explanatory variables. Accordingly, these estimates seem to strongly support our wealth residual hypothesis.

Table 1.5 Results of panel regressions– inequality

	(1) RE	(2) FE	(3) FE	(4) FE	(5) FE	(6) FE
WR	0.021 (0.016)	0.021 (0.017)	0.033** (0.015)	0.034** (0.015)	0.032** (0.015)	0.033** (0.015)
EDU				-0.044 (0.078)		-0.064 (0.072)
INO					0.003*** (0.001)	0.003*** (0.001)
GDP	-52.915 (69.336)	-88.527 (64.701)	-110.170* (59.628)	-110.587* (60.846)	-120.331** (56.179)	-121.218** (57.791)
GDP squared	2.808 (3.316)	4.569 (3.081)	5.237* (2.841)	5.259* (2.902)	5.677** (2.678)	5.722** (2.756)
State-fixed	No	Yes	Yes	Yes	Yes	Yes
Time-fixed	No	No	Yes	Yes	Yes	Yes
N	1,887	1,887	1,887	1,887	1,887	1,887
R^2	0.33	0.26	0.38	0.37	0.37	0.36

Notes: Variable description is given in 1.4. The variables that represent taxation, structural change, credit supply, financial regulation, housing supply, government size, welfare, and labour bargaining power were controlled. Robust standard errors are presented in parentheses. ***, **, and * indicate 0.01, 0.05, and 0.1 levels of significance, respectively.

Furthermore, the estimated coefficients on *EDU* and *INO* reveal that innovation, measured by the number of patents per capita, is positively and significantly correlated with rising top income inequality, as in line with Aghion et al. (2019), while human capital, measured by college attainment, has a role of dampening top income inequality, as in accordance with Barro and Lee (2013), when controlling for other explanatory variables. This suggests

Table 1.6 Results of panel regressions– growth

	(1) RE	(2) FE	(3) FE	(4) FE	(5) FE	(6) FE
WR	-0.016***	-0.030***	-0.050***	-0.050***	-0.050***	-0.050***
	(0.005)	(0.007)	(0.009)	(0.009)	(0.009)	(0.009)
EDU				-0.008		-0.008
				(0.040)		(0.040)
INO					0.000	0.000
					(0.000)	(0.000)
State-fixed	No	Yes	Yes	Yes	Yes	Yes
Time-fixed	No	No	Yes	Yes	Yes	Yes
N	1,887	1,887	1,887	1,887	1,887	1,887
R^2	0.26	0.16	0.57	0.56	0.57	0.56

Notes: Variable description is given in 1.4. The variables that represent taxation, structural change, credit supply, financial regulation, housing supply, government size, welfare, and labour bargaining power were controlled. Robust standard errors are presented in parentheses. ***, **, and * indicate 0.01, 0.05, and 0.1 levels of significance, respectively.

that the recent revolution of top income inequality partly related to innovation, even though education has contributed to reducing the degree of inequality. In contrast, the associations between growth and innovation, and human capital are not significant.

A common feature of empirical studies on the growth-inequality relationship is that they typically rely on single-equation methods, such as ordinary least squares, instrumental variables (for controlling endogeneity issues), or non-parametric methods (for allowing arbitrary non-linear relationship). However, structural models (e.g. general equilibrium frameworks) have shown that growth and inequality are linked through various channels (e.g. education, credit market, politics, etc.), and their relationship may generate the feedback dynamics. Accordingly, a single-equation estimation may fail to properly estimate the genuine relationship between growth and inequality, since these methods cannot reflect these interactive feedback dynamics. In this regard, I use a PVAR technique, which are more suitable to capture the dynamics and feedbacks of endogenous variables of interest.

1.4.3 Methodology

To begin with, I conduct a series of tests for unit root (or stationarity) in our panel dataset. The LLC test (Levin et al., 2002), the HT test (Harris and Tzavalis, 1999), the IPS test (Im et al., 2003), and the Fisher-type Augmented Dickey-Fuller (ADF) test (Choi, 2001) have as the null hypothesis that all the panels contain a unit root, whereas the Hadri (2000) test has as the null hypothesis that all the panels are stationary. Under the condition that our panel data are strongly balanced, I perform all five tests because each of these tests has the advantages and disadvantages (see Pesaran (2012) for the details). As suggested by Levin et al. (2002), all the tests are carried out on demeaned data to mitigate the impact of cross-sectional dependence. Lag lengths were chosen using the Akaike information criteria (AIC). Our inferences are based on a 5% level of sinificance.

Table A.1 presents the results. It shows that the first four tests, namely, the LLC, the HT, the IPS and the ADF, reject the null hypothesis of a unit root for each series in my panel, and the Hadri test does not reject the null hypothesis that there is no unit root in any series in my panel. Since our panel data are stationary, we can proceed to the estimation of the PVAR model with the variables in levels.

Next, I use a Pedroni (2013)'s structural PVAR technique that accommodates state-specific effects. This method allows the dynamics to be fully heterogeneous amongst states and decomposes the dynamics between different responses to idiosyncratic and common shocks.[38] Furthermore, this model provides us with a useful empirical methodology to study the relationship between growth and inequality, and wealth residual because the PVAR approach combines the traditional VAR approach, treating all the variables in the system as endogenous, and the panel-data approach allowing for unobserved individual heterogeneity by introducing fixed effects, resulting in an improved consistency of the estimation (Love and Zicchino, 2006).[39].

[38] Goes (2017) argues that this technique is a much more informative way of interpreting results (e.g. averages, medians and interquartile ranges) than traditional PVAR analysis. For instance, had I calculated average impulse responses from parameters estimated with traditional dynamic panels, I would have no way of knowing how many states in the sample have dynamics that are similar to the average dynamics, as the underlying assumption is that those parameters are equal for all states. Effectively, knowing exactly how many states in the sample present certain dynamics provides for much more robust inference than simply relying on average estimates. In addition to that, if individual dynamics are heterogeneous, aggregating or pooling slopes can lead to biased estimates (Pesaran and Smith, 1995), making individual regressions for each group member preferable.

[39] Although it is often said that all variables are treated as endogenous in a PVAR model, this is not strictly correct. When performing macroeconomic analysis there are clearly cases where variables are best thought of as strictly exogenous, and that should determine how the VAR is formulated. Perhaps the clearest example of this would be in the context of a small open economy, where the foreign variables would be expected to affect the domestic ones but not conversely, i.e. the foreign variables would be determined by their own lag values

In that context, for each state i = [1, ..., M]$'$ of a balanced panel, let $x_{i,t}$ be a vector of n endogenous variables with state-specific time dimensions t = [1, ..., T]$'$. To control for state-specific effects, I demean the data, $x^*_{i,t} = x_{i,t} - \bar{x}_i$, where $\bar{x}_i \equiv T_i^{-1} \sum_{t=1}^{T} x_{i,t}$ for all i. Then the model can be described as:

$$B_i x^*_{i,t} = A_i(L) x^*_{i,t-1} + e_{i,t} \tag{1.24}$$

where B_i, a matrix of contemporaneous coefficients; $x^*_{i,t}$, an n-dimensional vector of demeaned and stacked endogenous variables; $A_i(L) \equiv \sum_{j=0}^{J_i} A^i_j L^j$, a polynomial of lagged coefficient with state-specific lag-lengths J_i[40]; and a matrix of coefficient A^i_J; and $e_{i,t}$, a vector of stacked residuals.

To allow for heterogeneous dynamics, I estimate and identify reduced-form VARs for each state i:

$$\begin{aligned} B_1 x^*_{1,t} &= A_1(L) x^*_{1,t-1} + e_{1,t} \\ &\vdots \\ B_M x^*_{M,t} &= A_M(L) x^*_{M,t-1} + e_{M,t}. \end{aligned} \tag{1.25}$$

Then I estimate another auxiliary VAR to recover common dynamics. Common dynamics, $\bar{x}_t \equiv M^{-1} \sum_{i=1}^{M} x^*_{i,t}$ for $\forall t$, are captured by averages across states for each period:

$$\bar{B}_i \bar{x}^*_t = \bar{A}(L) \bar{x}^*_{i,t-1} + \bar{e}_t. \tag{1.26}$$

After transforming the reduced-form residuals in equation (1.24) and (1.25) into their structural equivalents ($u_{i,t} = B_i^{-1} e_{i,t}$ and $\bar{u}_{i,t} = \bar{B}_i^{-1} \bar{e}_{i,t}$, respectively), I run n-by-M ordinary least squares regressions to decompose the shocks into two parts:

$$\begin{aligned} u_{1,t} &= \Lambda_1 \bar{u}_t + \tilde{u}_{1,t} \\ &\vdots \\ u_{M,t} &= \Lambda_M \bar{u}_t + \tilde{u}_{M,t} \end{aligned} \tag{1.27}$$

where $u_{i,t}$, composite shocks; $\bar{u}_t$, common shocks; $\tilde{u}_{i,t}$, idiosyncratic shocks; and Λ_i, n-by-n diagonal matrices with state-specific coefficients, denoting the relative importance of

and not those of domestic variables. Nevertheless, in my estimation, it is hard to say that all control variables in my model are strictly exogenous. So I ruled out a PVAR model with exogenous variables.

[40] I choose J_i based on a criteria of Lagrange Multiplier to assure that residuals approximate white noises. The correct lag length selection is essential for a PVAR technique since having too short legs fails to capture the system's dynamics, leading to an omitted-variable-bias problem, while having too many lags causes a loss of degrees of freedom, resulting in over-parameterisation.

common shocks for each state. Note that $\tilde{u}_{i,t}$ vectors are truly idiosyncratic by construction since they are orthogonal to the shocks derived from the average dynamics shared by all states.

Subsequently, I recover the matrices of composite responses to structural shocks, $R_i(L)$, for each state, by using the Lutkepohl (2007) methods, which are shown in the vector moving average representations of M structural VARs:

$$\begin{aligned} x^*_{1,t} &= R_1(L)u_{1,t} \\ &\vdots \\ x^*_{M,t} &= R_M(L)u_{M,t}. \end{aligned} \tag{1.28}$$

Then, I use the loading matrices estimated in equation (1.27) to decompose the composite responses into state-specific responses to common shocks and responses to idiosyncratic shocks:

$$\begin{aligned} R_1(L) &= \Lambda_1 R_1(L) + (I - \Lambda_1\Lambda_1')R_1(L) \\ &\vdots \\ R_M(L) &= \Lambda_M R_M(L) + (I - \Lambda_M\Lambda_M')R_M(L). \end{aligned} \tag{1.29}$$

Let $R_i(L) = \bar{R}_i(L) + \tilde{R}_i(L)$, where $\bar{R}_i(L) \equiv \Lambda_i R_i(L)$ and $\tilde{R}_i(L) \equiv (I - \Lambda_i\Lambda_i')R_i(L)$. I then use the cross-sectional distribution of $R_i(L)$, $\bar{R}_i(L)$ and $\tilde{R}_i(L)$ to describe some properties of the collection of impulse response functions (IRFs) calculated, such as their averages, medians, and interquartile ranges. After recovering the point estimates of all the IRFs, I calculate standard errors of medians through a resampling simulation that repeats all the steps above 500 times.

Finally, I assume a recursive structure on the contemporaneous relationship among the variables (i.e., Cholesky decomposition) for my PVAR model.[41] First of all, since the measure of growth is used in the calculation of the measure of inequality, inequality is likely to respond contemporaneously to growth, while growth is likely to respond sluggishly to inequality, as seen in past empirical studies (e.g. Kim and Nelson, 1999; Sims and Zha, 2006). The remaining order of the variables is based on two assumptions: (i) growth is contemporaneously exogenous to wealth residual and (ii) wealth residual is contemporaneously exogenous to inequality. These assumptions are made because real per capita personal income is a flow variable, whereas wealth residual is a stock variable, and inequality is

[41]This assumption implies that variables listed earlier in the order contemporaneously impact the other variables, while variables later in the order impact those listed earlier only with a lag. Consequently, a variable listed in the first order is considered to be most contemporaneously exogenous.

a result of the distribution of the aggregate wealth. That is, wealth residual is likely to respond contemporaneously to growth, while wealth residual is likely to respond sluggishly to inequality. In this study, I explore the robustness of my results by varying this identifying assumption on ordering, showing that the identifying assumption on the order of the endogenous variables does not affect the results considerably. Lastly, it is natural to place the order of the government policy variables last since the government determines its policies after observing various economic indexes including growth, wealth residual, and inequality in reality.

Based on the methodology and data described above, I construct three specifications. The first specification is a bivariate PVAR model that identifies the relationship between growth and inequality: $x_{i,t} \equiv [\mathrm{Y}_{i,t}, \mathrm{INQ}_{i,t}]'$, where Y is the growth rate of real per capita personal income and INQ is the share of the top 1% income. The second specification is a trivariate PVAR model that identifies the impact of wealth residual on the growth-inequality relationship: $x_{i,t} \equiv [\mathrm{Y}_{i,t}, \mathrm{WR}_{i,t}, \mathrm{INQ}_{i,t}]'$, where WR is the share of land in the market value of real estate. Lastly, I extend my analysis by introducing several measures of the government's rent-related policies into the trivariate model to investigate the effects of the policies on the channel of wealth residual: $x_{i,t} \equiv [\mathrm{Y}_{i,t}, \mathrm{WR}_{i,t}, \mathrm{INQ}_{i,t}, \mathrm{GP}_{i,t}]'$, where GP stands for the government's rent-related policy.

1.5 Results

In this section, I present the impulse response functions (IRFs) estimated from the three model specifications described in the previous section.[42] Each figure displays $n \times n$ matrix of IRFs over 20 years from a percentage-point shock, where n is the number of variables, with 5% confidence intervals calculated from a resampling simulation with 500 repetitions ('Median composite IRFs')[43] or with the interquartile ranges calculated from the distribution of IRFs across 51 states ('Heterogeneous composite IRFs'). I also check whether my empirical findings are robust (i) to different ordering of the endogenous variables and (ii) to alternative measures of the key variables.

[42] I am sincerely grateful to Carlos Goes for sharing the E-views code of estimating Pedroni (2013)'s structural PVAR models with heterogeneous dynamics.

[43] In a resampling simulation method, the confidence intervals do not represent the uncertainty around a point estimate for any particular state but rather for the median estimate of all states because the distribution might change for each repetition.

1.5.1 The post-1978 growth-inequality relationship

From Figure 1.1, we can conjecture that there is a turning point for the trend of the wealth-to-income ratio and the share of the top 1% income at the beginning of 1978.[44] Stiglitz (2015a) argues that most of increases in the wealth-to-income ratio since the late 1970s reflect wealth residual that is not related to productive capital but is associated with asset ownership, giving rise to economic rents. To examine this, I split the whole post-World War 2 sample into two subsamples: 1959-1977 and 1978-2015. By comparing the IRFs across these two subsamples, this study has found that movements in the growth rate of real per capita income and the top 1% income share indicate a remarkable change in the relationship between growth and inequality at the beginning of 1978.

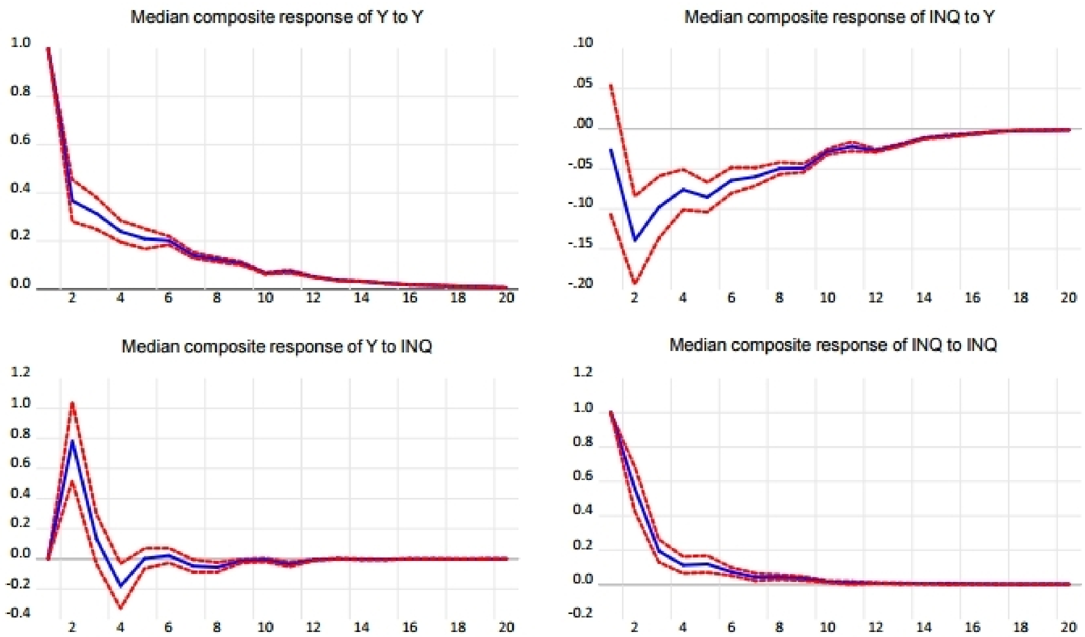
Consider first Figure 1.5a that shows the marginal IRFs for the 1959-1977 period. First, in the response of the growth rate of real per capita personal income (hereafter, "growth") to its own shock, growth increases sharply on impact but substantially decreases in the second year and then slowly goes back to the initial level in twenty years. Similarly, in the response of the share of the top 1% income (hereafter, "inequality") to its own shock, inequality increases sharply on impact but significantly decreases until the third year and then slowly returns to the initial level in ten years.

Next, the effect of a growth shock on inequality is *negative and long-lived*. More specifically, in response to a percentage-point increase in the growth rate of real per capita personal income, the share of the top 1% income falls (i.e., higher growth $\rightarrow$ lower inequality) and stays below its steady-state level for twenty years, with a trough response of -0.15 percentage points in the second year. In contrast, the effect of an inequality shock on growth is *positive but short-lived*. To be more specific, in response to a percentage-point increase in the share of the top 1% income, the growth rate of real per capita personal income increases (i.e., higher inequality $\rightarrow$ higher growth) and remains above its steady-state level for only three years, with a peak response of 0.8 percentage points in the second year.

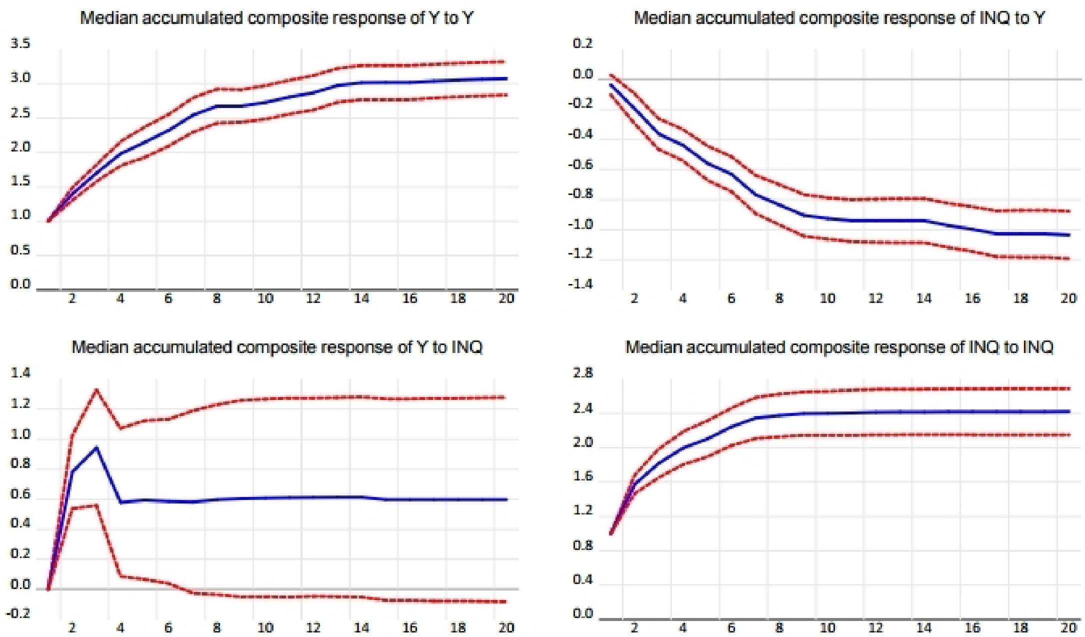
Furthermore, the cumulative IRFs in Figure 1.5b show that, although the effect of an inequality shock on growth is short-lived, there exists a permanent level-up effect of inequality on growth. They also show that, since the effect of a growth shock on inequality is long-lived, there naturally exist a permanent level-down effect of growth on top income

[44] Many studies have documented evidence of structural breaks at around 1980 in terms of the volatility of real GDP growth (e.g. Kim and Nelson, 1999; McConnell and Perez-Quiros, 2000) or inequality (e.g. Atkinson, 2015; Cutler and Katz, 1992). In this paper, I regard the year 1978 as a watershed because the share of the top 1% income was the lowest at that time and wealth-to-income ratios were overall stable until then. There was also the transition of the tax policy paradigm in 1978 – the US reduced the maximum rate on long-term capital gains from its historical peak of 39.875% to 28% for the first time after the World War 2.

Fig. 1.5 Impulse responses of the bivariate model, 1959-1977



(a) Marginal impulse responses



(b) Cumulative impulse responses

Notes: Each variable denotes the following data: Y, growth rate of real per capita personal income; INQ, share of the top 1% income. Dotted lines denote 5% confidence intervals calculated from a resampling simulation with 500 repetitions.

inequality. Accordingly, these results support the traditional views that higher growth leads to lower inequality but higher inequality is beneficial to growth (e.g. Forbes, 2000; Galor and Tsiddon, 1997; Partridge, 1997).

Now consider Figure 1.6a that represents the marginal IRFs of the 1978-2015 period, in which the story is totally reversed. First, the duration of a growth shock is shrunk to one-fifth of the 1959-1977 sample, while that of an inequality shock is extended to double of the 1959-1977 sample. Second, the effect of a growth shock on inequality is *positive and short-lived*. In particular, in response to a percentage-point increase in the growth rate of real per capita personal income, the share of the top 1% income rises (i.e., higher growth $\rightarrow$ higher inequality) and stays above its steady-state level for only three years, with a peak response of 0.5 percentage points on impact. In contrast, the effect of an inequality shock on growth is *negative and long-lived*. More specifically, in response to a percentage-point increase in the share of the top 1% income, the growth rate of real per capita personal income falls (i.e., higher inequality $\rightarrow$ lower growth) and remains below its steady-state level for more than twenty years, with a trough response of -0.4 percentage points in the third year.

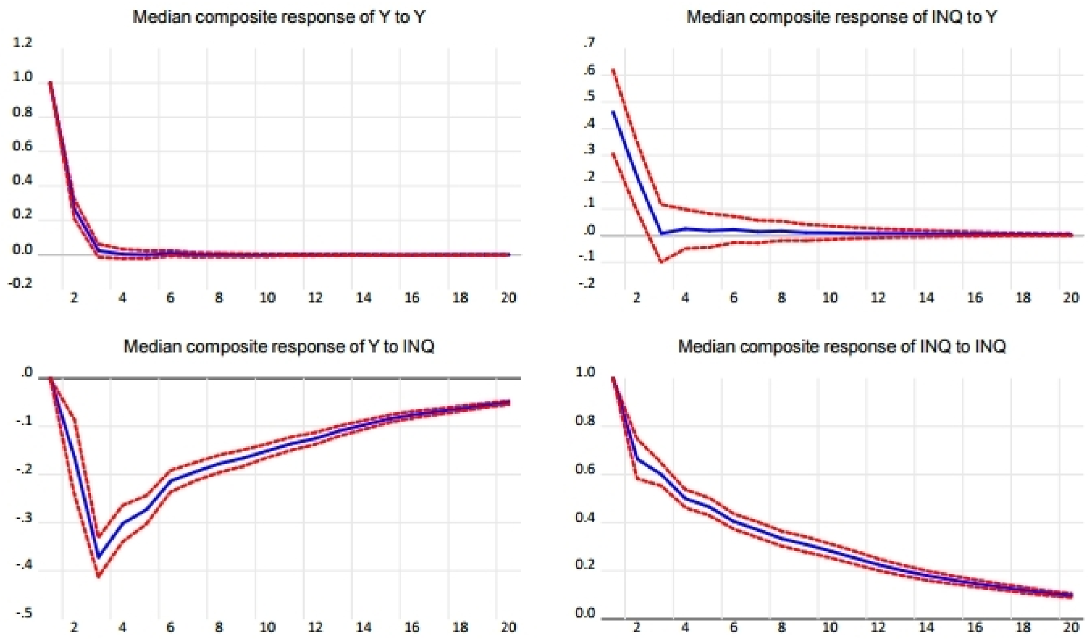
Furthermore, the cumulative IRFs in Figure 1.6b describe that there exist a permanent growth-hindering effect of inequality and a permanent inequality-enhancing effect of growth, supporting the recent views that higher inequality may hurt growth and higher growth would not bring lower inequality, unlike in the pre-1978 period (e.g. Acemoglu and Robinson, 2000; Atems and Jones, 2015; Brueckner and Lederman, 2018).

Overall, the empirical findings from the bivariate model suggest that the relationship between growth and inequality had completely changed since 1978. We used to think of there being a trade-off: we could achieve more equality but only at the expense of overall economic performance. However, the post-1978 evidence shows that greater equality and improved economic performance can be complements.

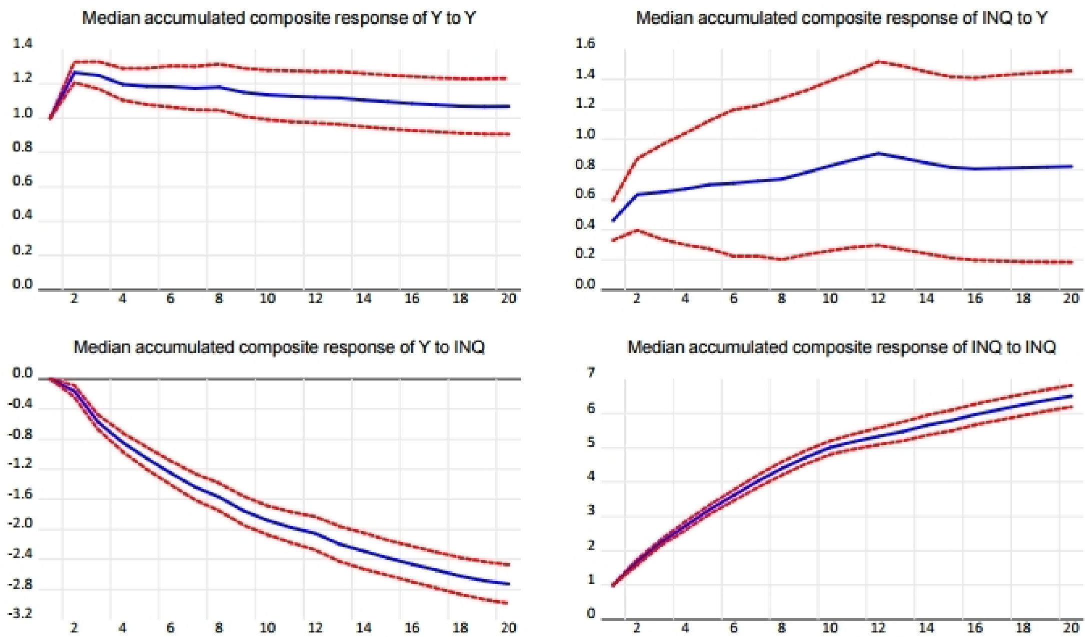
1.5.2 Wealth residual channels

A question naturally arises from the above results: through what channels the growth-inequality has structurally changed since the late 1970s? To answer this, I extend my analysis by estimating a three-variable PVAR model that includes a measure of wealth residual for two purposes. The first one is that, conceptionally, it allows us to test the wealth residual hypothesis that is theoretically presented in Section 3.4. The second one is that, methodologically, the bivariate PVAR model possibly suffers from an omitted-variable-bias problem. That is, there might be other macroeconomic shocks besides an inequality shock, simultaneously

Fig. 1.6 Impulse responses of the bivariate model, 1978-2015



(a) Marginal impulse responses



(b) Cumulative impulse responses

Notes: Each variable denotes the following data: Y, growth rate of real per capita personal income; INQ, share of the top 1% income. Dotted lines denote 5% confidence intervals calculated from a resampling simulation with 500 repetitions.

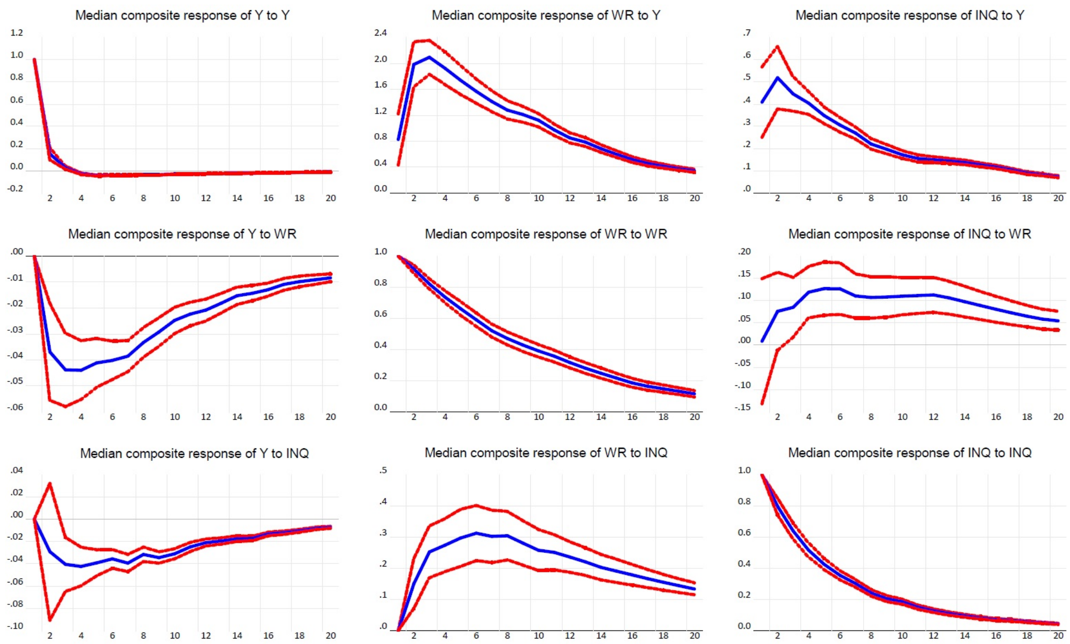
affecting growth and inequality.

Consider first Figure 1.7a that shows the marginal IRFs for the 1978-2015 period. Though the inclusion of wealth residual does not change the underlying growth-inequality relationship, we can identify the effect of wealth residual on the growth-inequality relationship. First, the effect of an wealth residual shock on inequality is *positive and long-lived.* More specifically, in response to a percentage-point increase in the share of land value in the market value of real estate, inequality significantly increases (i.e., more wealth residual $\rightarrow$ higher inequality) and stays above its steady-state level for more than twenty years, with a peak of 0.10 percentage points in the fifth year. In contrast, the effect of an wealth residual shock on growth is *negative and long-lived.* More specifically, in response to a percentage-point increase in the share of land value in the market value of real estate, the growth rate of real per capita personal income decreases (i.e., more wealth residual $\rightarrow$ lower growth) and remains below its steady-state level for more than twenty years, with a trough of -0.5 percentage points in the third year.

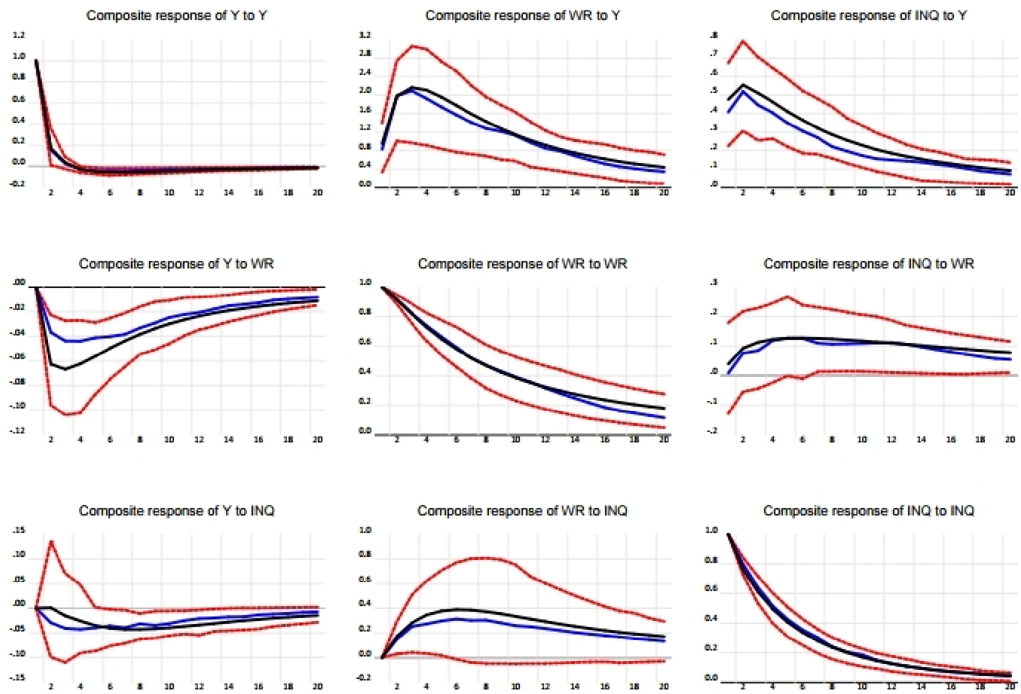
Now consider Figure 1.7b that shows heterogeneous marginal IRFs for the 1978-2015 period with the median, averages, and interquartile ranges calculated from the distribution of IRFs across 51 states. It shows that, even though individual states' dynamics are heterogeneous, the results do not change the sign of point estimates for at least 75 percent of all the states. To be more specific, a percentage-point shock to wealth residual leads to a decrease in the growth rates until the third year but the extents of the decreases vary from -0.02 to -1.1 percentage points. Similarly, a percentage-point wealth residual shock leads to an increase in inequality until the fifth year but the extents of the increases vary from 0 to 0.25 percentage points.

Overall, the empirical findings in the trivariate model refute the traditional channel that higher inequality gives more resources to people who have higher marginal propensity to save (so-called 'the rich'), leading to an increase in capital accumulation and thereby higher growth rates. On the contrary, my empirical results support the view that higher growth during the post-1978 period only benefits the top income earners at the expense of the bottom income earners through the sudden increase in unearned income, closely related to the appreciation of asset values, as argued in the recent studies (e.g. Atkinson, 2015; Eggertsson et al., 2018; Lee, 2020; Piketty, 2014; Rognlie, 2015; Stiglitz, 2015a). Especially, this unearned income has a property of positional competition: an attempt to acquire wealth residual can only benefit the rentiers (those who already owned assets) at the expense of their renters (those who does not own any assets), resulting in a zero-sum game without any increase in the productive capacity of the economy. For society as a whole, it has gen-

Fig. 1.7 Marginal impulse responses of the trivariate model



(a) Median composite responses and confidence intervals



(b) Heterogeneous composite responses across sample

Notes: Each variable denotes the following data: Y, growth rate of real per capita personal income; INQ, share of the top 1% income; and WR, share of land in the market value of real estate. Blue and black lines stand for the median and average estimates, respectively. Dotted lines denote 5% confidence intervals calculated from a resampling simulation with 500 repetitions in the upper graph, while interquartile ranges calculated from the distribution of IRFs of the 51 states in the lower graph.

erated a co-evolution of fast-growing inequality and falling growth over the past four decades.

1.5.3 Robustness check

I check if the previous findings are robust (i) to the identification assumption on the order of contemporaneously exogenous variables and (ii) to the measurement of inequality and wealth residual. I present all alternative results of the marginal IRFs in Figure A.5, A.6, and A.7.

First, I consider alternative identifying assumptions. The most critical order variation is to assume WR as contemporaneously exogenous to Y and INQ, that is [WR, Y, INQ], to check the robustness under the opposite assumption on contemporaneous relation to the baseline trivariate model. I also consider the model of [Y, INQ, WR], where INQ is assumed to be contemporaneously exogenous to WR.

Figure A.8 summarises the results of the IRFs for the models of [WR, Y, INQ] and [Y, INQ, WR]. While the model of [WR, Y, INQ] shows a little lower trough response of growth to wealth residual (-0.05 vs. -0.04 in the baseline) and a little lower peak response of inequality to wealth residual (0.05 vs. 0.1 in the baseline), the model of [Y, INQ, WR] represents quite a lower trough response of growth to wealth residual (-0.2 vs. -0.04 in the baseline) and quite a higher peak response of inequality to wealth residual (0.25 vs. 0.1 in the baseline). However, such differences do not change the overall interpretation of the results of the baseline model.

Figure A.9 summarises the results of the IRFs using alternative measures of inequality and wealth residual. First, it shows the IRFs, replacing the share of the top 1% income with other typical measures of inequality (Gini index, Atkinson index, share of the top 10% income, and share of the top 0.1% income). The consistency in the direction of responses, using different measures for inequality, suggests that the relationship among growth, inequality, and wealth residual is qualitatively and quantitatively similar. Second, I use land price index as a proxy for wealth residual. Figure A.9a describes that the trough response of growth to a percentage-point shock in an alternative wealth residual measure is larger than the baseline model (-0.006 vs. -0.04 in the baseline). Similarly, Figure A.9b describes that peak response of inequality to an alternative wealth residual measure is also higher than the baseline (0.3 vs. 0.1 in the baseline). However, the direction of response of growth and inequality is not changed and statistically significant. Overall, such differences do not change the interpretation of the results of the baseline model.[45]

1.6 Policy Implications

In this section, I attempt to answer the following questions: (i) what kind of economic policies have contributed to the upsurge in wealth residual?, (ii) given the two-party political system of the US in reality, how has a political regime shift by an election from rentier-friendly to labour-friendly affected wealth residual channels?, and conversely, (iii) whether has the degree of wealth residual influenced on the political regime change? To address these questions, I use the techniques of PVARs, panel regression models, discrete choice models that are suitable to each question based on a comprehensive panel dataset at state level.

1.6.1 Drivers of wealth residual increase

We can conjecture that the large share of rent-related activities and the consequent rise of the share of the top income can undermine the marginal productivity theory of income distribution. This suggests that institutional factors play an important role in influencing the distribution of income (Alvaredo et al., 2013). In this respect, this study incorporates the government's rent-related policies into the trivariate PVAR model. Figure 1.8a, 1.8b, 1.8c, 1.8d, 1.8e, and 1.8f show the marginal IRFs of the quadvariate models for the 1978-2015 period by adding proxies for taxation (TAX), structural change (SC), credit supply (MOR), financial regulation (LTV), and housing supply (HOU), respectively.

Taxation The last row in Figure 1.8a shows the marginal IRFs to a percentage-point shock to the maximum tax rate on long-term capital gains (hereafter, "tax"). In the first column, the effect of a tax shock on growth is insignificant in the short run but *weakly negative in the long run*: in response to a percentage-point increase in the maximum tax rate on capital gains, the growth rate of real per capita personal income falls a little from the ninth to the fifteenth year (i.e. higher tax $\rightarrow$ weakly lower growth rates in the long run), supporting the view that an increase in taxes at the top do not result in substantial decreases in the growth

[45] The magnitude of responses with regard to alternative measures reflects their different statistical properties (e.g. mean, variance).

rates (e.g. Piketty et al., 2014; Stiglitz, 2015a). In the second column, the effect of a tax shock on wealth residual is *positive and long-lived*: in response to a percentage-point increase in the maximum tax rate on capital gains, the share of land value significantly increase (i.e. higher tax $\rightarrow$ higher wealth residual) and stay above its steady-state level for over twenty years, with a peak response of 0.06 percentage point in the fifth year. The third column shows that the effect of a tax shock on inequality is insignificant. Overall, these results imply that a progressive tax on capital gains alone may not be an effective way in achieving higher growth rates and lower inequality at the same time. This is partly because, even though the government decides to increase the tax rate on capital gains, the increasing tax burden on asset owners is likely to result in increased rents on tenants or reduced incentive in investments in education or innovation (e.g. Becker and Tomes, 1994; Stiglitz, 2015a), when there exists a limited supply of good quality rental properties.[46]

Structural change The last row in Figure 1.8b describes the marginal IRFs to a percentage-point shock to the employment share of manufacturing in private industry (hereafter, "structural change"). In the first column, the effect of an structural-change shock on growth is an *inverted-hump shape*: in response to a percentage-point increase in the share of manufacturing, the growth rate of real per capita personal income falls up to the fifth year but after that, it significantly increases (i.e., higher manufacturing share $\rightarrow$ lower growth rates in the short run but higher growth rates in the long run). In the second column, the effect of a structural shock on wealth residual is *negative and long-lived*: in response to a percentage-point increase in the share of manufacturing, the share of land value significantly decrease (i.e., higher manufacturing share $\rightarrow$ lower wealth residual), supporting the view that the upsurge of wealth residual is associated with the decline in productive activities – the surging financial supply to unproductive activities has crowded out productive investment, as argued by Stiglitz (2015a), Brun and Gonzalez (2017), and Rognlie (2015). In the third column, the effect of a structural change shock on inequality is *negative and long-lived*: in response to a percentage-point increase in the share of manufacturing, the share of the top 1% income decreases (i.e., higher manufacturing share $\rightarrow$ lower inequality). Overall, these results imply that the fall in employment in manufacturing industries is closely related to the recent sluggish growth rates and soaring inequality (e.g., the decline of Detroit).

[46]If the equilibrium rate of real interest is endogenous there can be tax shifting. For instance, Becker and Tomes (1994) argue that progressive income taxes may actually lead to an increase in inequality because of the reduced investments in children. Moreover, Stiglitz (2015d) shows that the before tax rate can increase, so much so that the effect of the capital taxation, if not carefully designed, could increases inequality. Instead, a tax on wealth residual (the so-called "Georgian tax") can actually lead to reducing inequality.

Credit supply One of the main drives of the growth in wealth residual is credit availability (Stiglitz, 2015a,d). The last row in Figure 1.8c shows the marginal IRFs to a percentage-point shock to the average interest rate on mortgages (hereafter, "credit supply"). In the first column, the effect of a credit supply shock on growth is an *inverted-hump shape*: in response to a percentage-point increase in the average interest rate on mortgages, the growth rate of real per capita personal income significantly falls up to -0.20 percentage points in the second year and then turns to increase at the beginning of the sixth year (i.e. higher credit supply $\rightarrow$ higher growth in the short run but lower growth in the long run), staying above its steady-state level for over twenty years. In the second column, the effect of a credit supply shock on wealth residual is *negative and long-lived*: in response to a percentage-point increase in the average interest rate on mortgages, the share of land value significantly decreases (i.e. higher credit supply $\rightarrow$ higher wealth residual) and stays below its steady-state level for over twenty years, with a trough response of -1.0 percentage point in the eighth year. In the third column, the effect of a credit supply shock on inequality is *negative and long-lived*: in response to a percentage-point increase in the average interest rate on mortgages, the share of the top 1% income persistently decreases (i.e. higher credit supply $\rightarrow$ higher inequality), with a trough of -0.5 percentage point in the fourth year. Since the demand for wealth residual depends itself on the availability of credit, if a favoured few get access to credit, then their wealth increases relative to those without such access. Therefore adjusting interest rates on mortgages can be an effective policy instrument in controlling wealth residual and its impact on inequality without hurting growth in the long run.

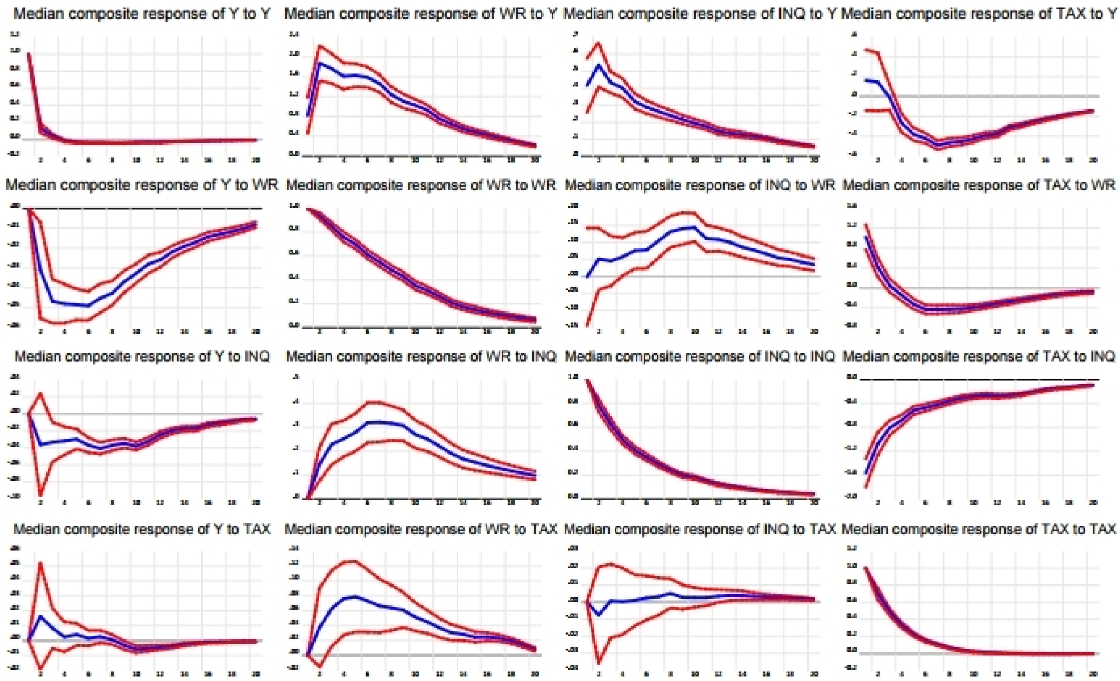
Financial regulation The last row in Figure 1.8d shows the marginal IRFs to a percentage-point shock to the loan-to-value ratio (hereafter, "LTV"). In the first column, the effect of an LTV shock on growth is not significant. In the second column, the effect of an LTV shock on wealth residual is *positive and long-lived*: in response to a percentage-point increase in LTV, the share of land value significantly increase (i.e. higher LTV $\rightarrow$ higher wealth residual) and stay above its steady-state level for over twenty years, with a peak response of 0.06 percentage point in the fifth year. In the third column, the effect of an LTV shock on inequality is *positive and long-lived*: in response to a percentage-point increase in LTV, the share of the top 1% income rises (i.e. higher LTV $\rightarrow$ higher inequality) and stays above its steady-state level for over twenty years, with a peak response of a 0.06 percentage point in the sixth year. Overall, these results imply that if the government maintains LTVs at low levels, wealth residual and inequality can be stayed at lower levels without damaging on the long-term growth.

Housing supply The last row in Figure 1.8e shows the marginal IRFs to a percentage-point shock to the homeownership rate (hereafter, "housing supply"). In the first column, the effect of a housing supply shock on growth is insignificant.[47] In the second column, the effect of a housing supply shock on wealth residual is *positive and short-lived*: in response to a percentage-point increase in the homeownerhship rate, the share of land value significantly increase (i.e. higher housing supply $\rightarrow$ higher wealth residual) and stay above its steady-state level for ten years, with a peak response of 0.1 percentage point in the fourth year. In the third column, the effect of a housing supply shock on inequality is *positive and long-lived*: in response to a percentage-point increase in the homeownerhship rate, the share of the top 1% income rises (i.e. higher housing supply $\rightarrow$ higher inequality) and stays above its steady-state level for over twenty years, with a peak response of a 0.20 percentage point in the third year. However, if we replace the share of the top 1% income with poverty ratio as a measure of inequality, we can get a different implication. The last row and the third column in Figure 1.8f shows that in response to the same shock, the poverty ratio decreases (i.e. higher housing supply $\rightarrow$ lower poverty) and stays below its steady-state level for twenty years, with a trough of -0.06 percentage point in the third year. Overall, these results imply that tenants who did not own assets were left behind from the housing boom over the period of 1978-2015; on the contrary, they have been adversely affected by rising scarcity rents, resulting in rising inequality without increases in the growth rates. However, the increase in homeownership (e.g. social housing) can contribute to reducing poverty ratio.

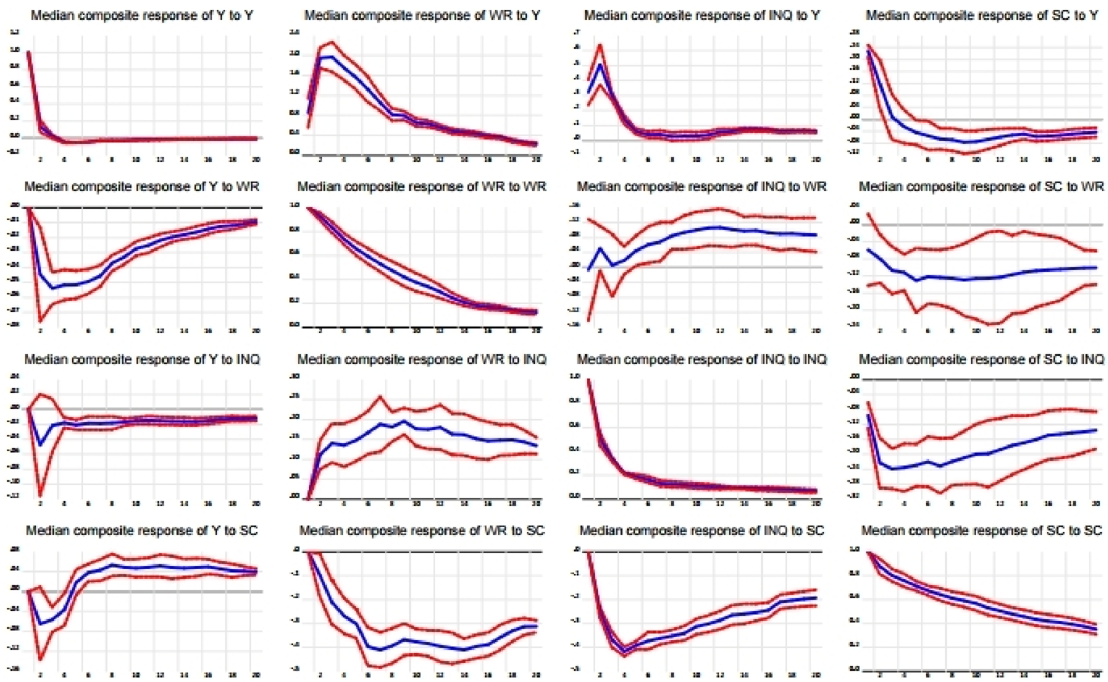
In summary, the US government must enact stricter regulation on rent-related activities by controlling the LTV ratio and managing credit supply at the level of curving the rapid expansion of rent-related finance. Moreover, a sharp increase in wealth tax or housing supply itself is not likely to solve the problem, while increasing affordable housing for the poor can reduce absolute poverty without hurting growth. Most importantly, the government can tax more on wealth residual instead of capital, and use the funds to finance either private or public investment, supporting the views of Stiglitz (2015a,b,c,d,e) and Atkinson (2015); such policies reduce inequality and lead to increased national savings. By doing so, industrial policies that can divert excess rent-related finance into productive investment must take precedence.

[47]Meanwhile, the first row in Figure 1.8e depicts that, in response to an increase in the growth rates, the homeownership rate substantially rises

Fig. 1.8 Marginal impulse responses of the quadvariate model



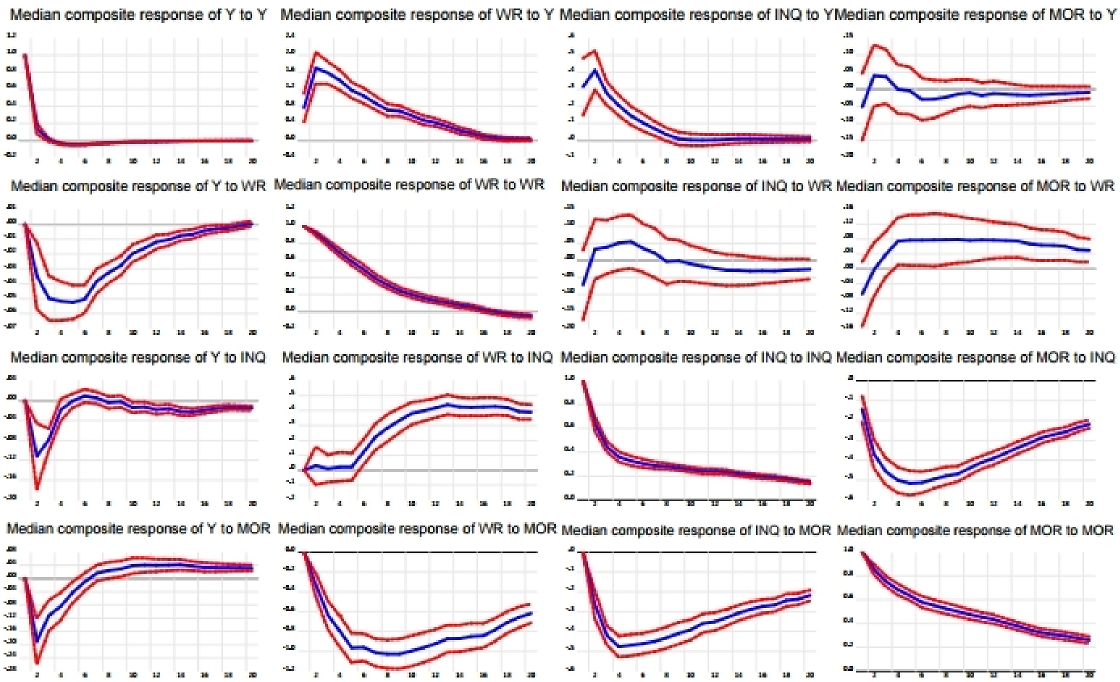
(a) Taxation



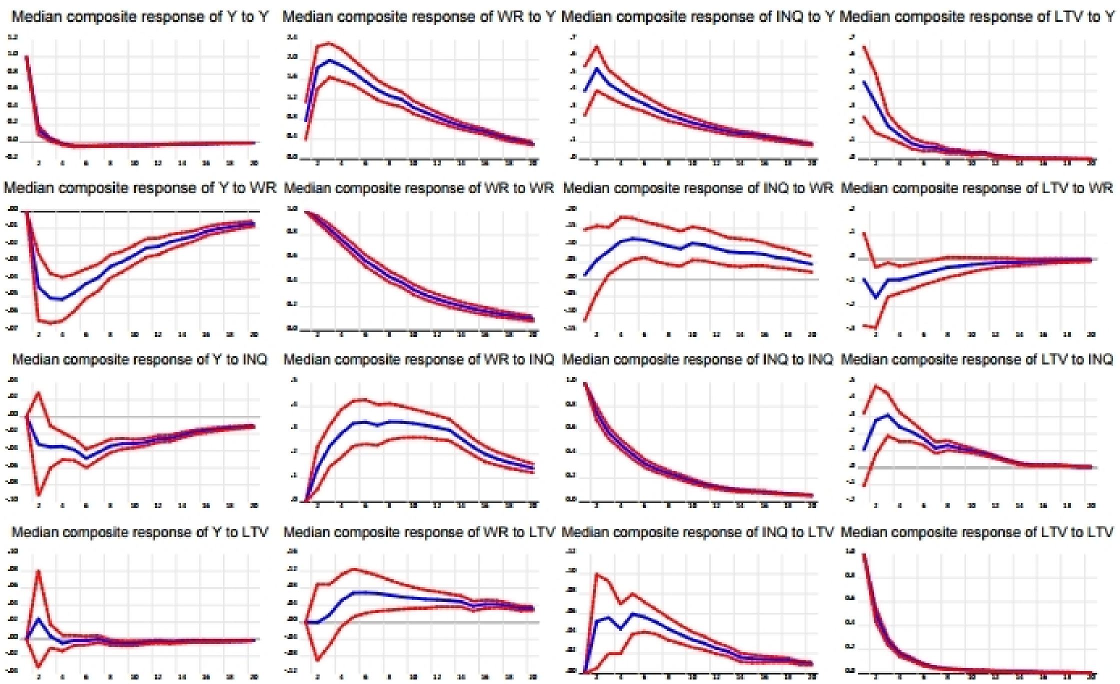
(b) Structural change

Notes: Each variable denotes the following data: Y, growth rate of real per capita personal income; INQ, share of the top 1% income; WR, share of land in the market value of real estate; TAX, maximum tax rate on wealth gains; and SC, employment share of manufacturing. Dotted lines denote 5% confidence intervals calculated from a resampling simulation with 500 repetitions.

Fig. 1.8 Marginal impulse responses of the quadvariate model (cont.)



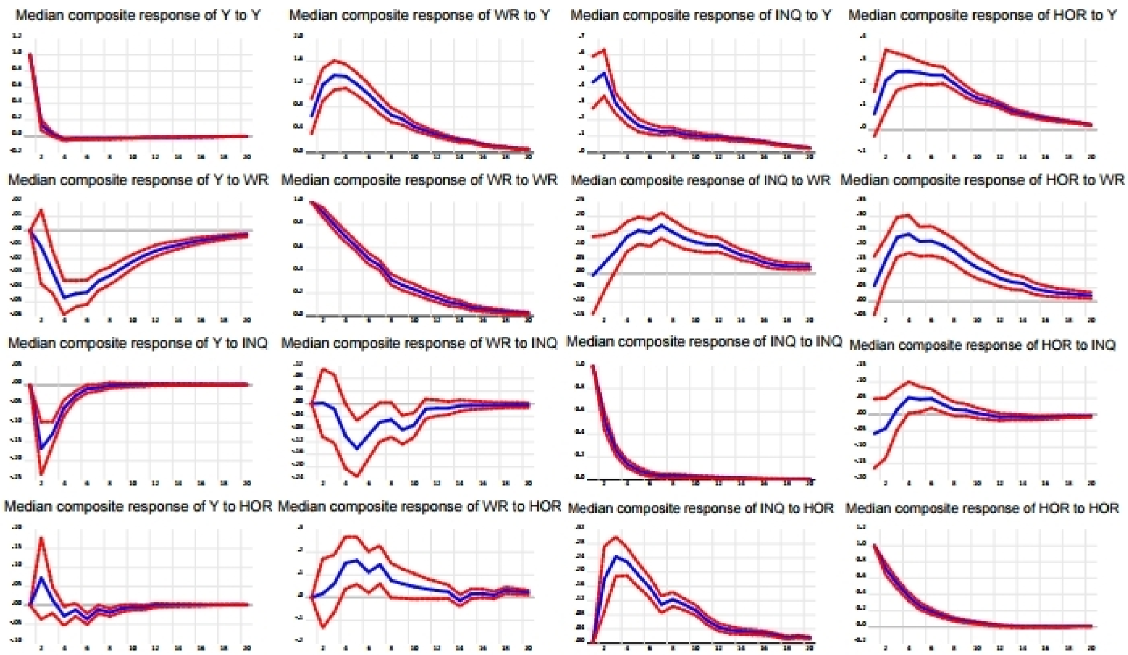
(c) Credit supply



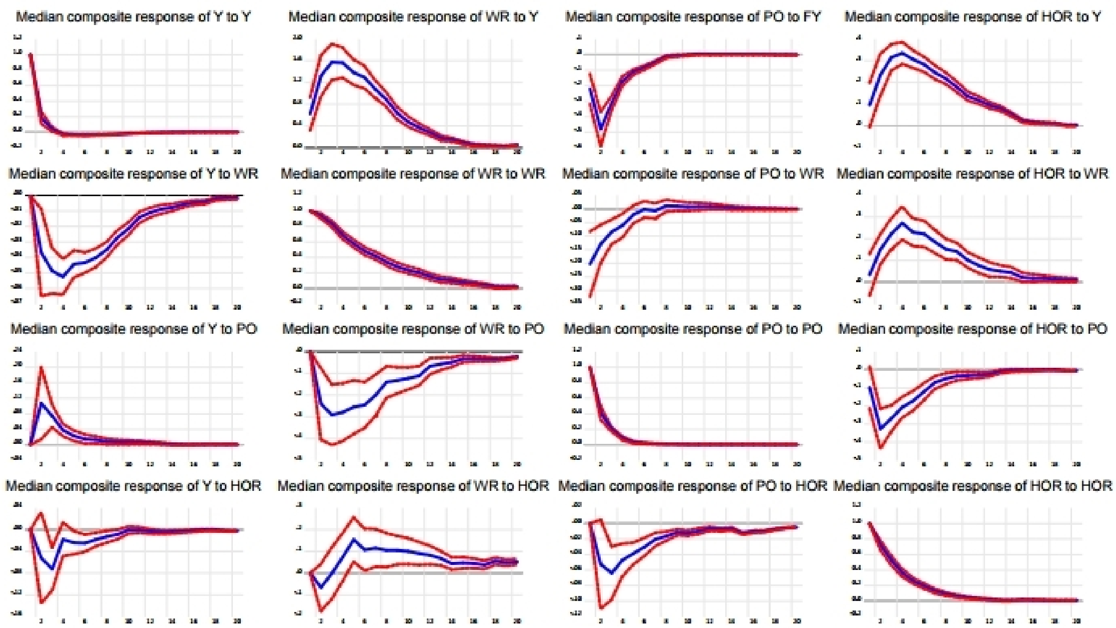
(d) Financial regulation

Notes: Each variable denotes the following data: Y, growth rate of real per capita personal income; INQ, share of the top 1% income; WR, share of land in the market value of real estate; MOR, average mortgage rate; and LTV, loan-to-value ratio. Dotted lines denote 5% confidence intervals calculated from a resampling simulation with 500 repetitions.

Fig. 1.8 Marginal impulse responses of the quadvariate model (cont.)



(e) Housing supply



(f) Housing supply and poverty ratio

Notes: Each variable denotes the following data: Y, growth rate of real per capita personal income; INQ, share of the top 1% income; WR, share of land in the market value of real estate; HOR, homeownership rate; and PO, poverty ratio. Dotted lines denote 5% confidence intervals calculated from a resampling simulation with 500 repetitions.

1.6.2 Political regime changes

Generally speaking, the type of the government is determined by elections. In case of the US government, the type has been changed frequently in binary options for every state– the party control of state legislatures or state government changes between the Democratic Party and the Republican Party with each election. But these economic consequences of the political regime changes are substantially different from each other (e.g. Mukand and Rodrik, 2020).

To investigate these, I carry out an additional exercise by assuming that the government has two distinct policy mixes: (a) a labour-friendly policy that uses more progressive taxation and rather stricter regulations (e.g. California) and (b) a rentier-friendly policy that conducts more regressive taxation and rather looser regulations (e.g. Texas). Comparing these two results makes it clear whether the government that is supported by those with unearned income remarkably expands wealth residual and aggravates inequality, without comparative advantage in growth rates for the whole economy. As in Table A.2, I construct the panel dataset of political regime changes between the two regimes (Democratic Party = '1') at state level for the 1978-2015 period using the State and Legislative Partisan Composition data from the National Conference of State Legislatures. When the same party holds both legislative chambers and the governorship, that party definitely has control over state policies. But when any of those three points of power is held by another party, I assume that state control is determined by the governorship.[48]

Table A.3 describes the statistics on rent-related policies and economic outcomes for the two regimes over the 1978-2015 period. It implies that the Democratic Party has conducted a rather labour-supportive policy that uses more progressive taxation, stricter financial regulations, and more housing supply than the Republican Party. As a result, labour-friendly governments have showed a limited growth of wealth residual and lower inequality than have rentier-friendly governments.

Thus, by examining the net effects of the increase in wealth-residual on growth and inequality when the political regime shifts from a rentier-friendly to a labour-friendly government, I quantitatively evaluate and predict the economic impacts of a policy mix of taxation and related regulations. To do this, based on the model specification in equation (1.23), I add a categorical variable, $D_{i,t}$, which takes a value of one if a state i at time t is controlled by the Democratic Party, and its cross-term with wealth residual, $D_{i,t} \cdot WR_{i,t}$. I also let the time lag between the dependent variable and our measure of a regime change and its cross-term with wealth residual vary from one to four years because there is a lag in the effect of regime

[48] One exception is Nebraska, where its state control party is determined by the governorship only since its legislators are elected on a non-partisan basis.

changes on the regional economy.[49] I estimate the following panel regression equation:

$$X_{i,t} = \alpha + \beta_1 D_{i,t-j} + \beta_2 WR_{i,t-j} + \beta_3 D_{i,t-j} \cdot WR_{i,t-j} + Z'_{i,t-1}\gamma + \lambda_i + \theta_t + \varepsilon_{i,t} \tag{1.30}$$

where $X_{i,t}$, a measure of either inequality or growth; $WR_{i,t}$, wealth residual; $Z_{i,t}$, a vector of control variables described in Table 1.4; λ_i, state-fixed effects; θ_t, time-fixed effects; $\varepsilon_{i,t}$, the error term; and j, the time lag ranging from 1 to 4.

Table 1.7 presents the result of estimating equation (1.30) by using the FE estimator. My main focus is on β_3 that captures the net wealth-residual effect on the dependent variable (inequality or growth) due to the regime shift towards a labour-friendly government. From columns 1-4, we find significantly negative interaction terms between wealth residual and regime changes. In contrast, from columns 5-8, we discover positive interaction terms between them but it is not significant. These results imply that the effect of wealth residual on top income inequality is dampened when labour-friendly governments take the state control. Furthermore, from the coefficients on lagged terms in columns 2-4, we can see that those magnitude decrease with the lag and remain significant for up to two years. The effect of regime changes eventually disappears as we increase the lag beyond three years – that is, lame duck session begins.

Conversely, a following question naturally arises: does the high degree of wealth residual incur a political regime change towards a more labour-friendly one? To answer this question, I consider binary outcome models, such as Probit RE, Logit RE, and Logit FE.[50] I estimate the following equation:

$$Pr(D_{i,t} = 1 | X_{i,t}, Z_{i,t}, \alpha_i) = X_{i,t}\beta + Z_{i,t}\gamma + \alpha_i + \mu_{i,t} \tag{1.31}$$

where $X_{i,t}$, a vector of key variables (wealth residual, growth, and inequality), $D_{i,t}$, a categorical variable that takes a value of one if a state i at time t is controlled by the Democratic Party; $Z_{i,t}$, a vector of control variables described in Table 1.4; and α_i is state-fixed effects.

Table 1.8 displays the result of estimating equation (1.31). Columns 1-3 show that increases in wealth residual and inequality are associated with the decline in the prevalence

[49] All US governors serve four-year terms except those in New Hampshire and Vermont, who serve two-year terms. Members of the smaller chamber (the Senate) represent more citizens and usually serve for generally four years, longer than two years of members of the larger chamber (House of Representatives).

[50] According to Cameron and Trivedi (2010), the advantage of Logit FE is to include the unobserved state-specific fixed effects, which are allowed to be correlated with other covariates, reducing the possibility of omitted variable bias. However, FE is not allowed for Probit estimators.

Table 1.7 The effects of political regime changes on inequality and growth

	(1) INQ	(2) INQ	(3) INQ	(4) INQ	(5) Y	(6) Y	(7) Y	(8) Y
WR(t-1)	0.040**				-0.053***			
	(0.016)				(0.011)			
WR(t-2)		0.028*				-0.062***		
		(0.015)				(0.009)		
WR(t-3)			0.016				-0.058***	
			(0.014)				(0.008)	
WR(t-4)				0.011				-0.056***
				(0.013)				(0.007)
D·WR(t-1)	-0.022**				0.007			
	(0.010)				(0.012)			
D·WR(t-2)		-0.016*				0.000		
		(0.009)				(0.010)		
D·WR(t-3)			-0.007				0.002	
			(0.009)				(0.010)	
D·WR(t-4)				-0.001				0.002
				(0.009)				(0.006)
N	1,887	1,836	1,785	1,734	1,887	1,836	1,785	1,734
R^2	0.35	0.34	0.31	0.28	0.57	0.54	0.55	0.50

Notes: Panel FE regressions with state and year fixed effects. Variable description is given in 1.4. Control variables are not reported but their values are similar to those in column 6 in Table 1.5 and 1.6. Robust standard errors are presented in parentheses. ***, **, and * indicate 0.01, 0.05, and 0.1 levels of significance, respectively.

of labour-friendly governments.[51] Columns 4-6 represent the marginal effect (assuming α_i=0) of an increase in key variables (dy/dx), saying that 10 percentage-points increases in wealth residual and inequality are associated with, respectively, 2-3 percentage-points decrease and 12-21 percentage-points decrease in the possibility of taking the state control by a labour-friendly government. In summary, these findings suggest that a labour-friendly government tends to win elections when it has maintained both wealth residual and inequality at low levels during its term in the office.

[51] From the Hausman test, I have confirmed strong rejection of the null hypothesis that RE provides consistent estimates.

Table 1.8 The effect of wealth residual on political regime changes

	(1) Probit RE	(2) Logit RE	(3) Logit FE	(4) Probit RE dy/dx	(5) Logit RE dy/dx	(6) Logit FE dy/dx
WR	-0.007*	-0.012*	-0.016**	-0.003*	-0.003*	-0.002**
	(0.004)	(0.007)	(0.007)	(0.002)	(0.006)	(0.001)
Y	0.016	0.024	0.022	0.002	0.006	0.003
	(0.013)	(0.022)	(0.022)	(0.005)	(0.005)	(0.003)
INQ	-0.053***	-0.088***	-0.086***	-0.020***	-0.021***	-0.012***
	(0.012)	(0.020)	(0.020)	(0.005)	(0.005)	(0.001)
N	1,938	1,938	1,482	1,938	1,938	1,482

Notes: Variable description is given in 1.4. Standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

1.7 Conclusion

"The ordinary progress of a society which increases in wealth, is at all times tending to augment the incomes of landlords; to give them both a greater amount and a greater proportion of the wealth of the community, independently of any trouble or outlay incurred by themselves. They grow richer, as it were in their sleep, without working, risking, or economizing." – Mill (1848), Principles of Political Economy (p.28)

The aim of this paper is to empirically examine the wealth residual hypothesis that a fast-growing unearned income from wealth residual is a key factor of the concurrence of surging inequality and sluggish growth in the post-1978 American economy. To demonstrate this hypothesis, this study not only applies a comprehensive panel dataset of the US at the state level but also conducts PVARs with heterogeneous dynamics.

The estimation results demonstrate that the upsurge of wealth residual, influenced by the government's regressive stance on rent-related activities, has contributed to a co-evolution of fast-growing inequality and falling growth since the late 1970s, chiefly due to intensified position competition. That is, an individual's attempts to acquire wealth residual can only benefit one agent at the expense of another, resulting in a zero-sum game without any increase in the productive capacity of the economy. If the chief beneficiary is the rich and the opponent is the poor, this economy will naturally experience rising inequality. In this respect, this study supports Stiglitz (2015a,b,c,d,e)'s assertion that appropriately defined aggregates for wealth may be moving in a direction opposite to what was supposed in the old-fashioned 'trickle-down' theory in which greater concentration of wealth at the top can increase growth, at least for the post-1978 period.

After the unwinding of the financial crisis in 2008, we have witnessed numerous examples of highly compensated individuals whose apparent contributions to social output proved illusory. Events like the housing and equity bubbles provide fertile ground for the pursuit of personal enrichment by the wealthy by extracting a slice of the existing economic pie rather than by increasing the size of that pie. In this vein, this study suggests that a new set of economic policies that can reduce the intensity of wealth residual (e.g. better regulated financial systems, more affordable housing, and more progressive tax-and-transfer policies), accompanied by industrial policies that can divert excess rent-related finance into productive investment, are required to achieve greater equality and faster growth for the US economy. These policies can encourage individuals to put their savings into more productive forms and restore the conventional growth-inequality relationship that Kuznets (1955) predicted.

Given that this study empirically shows that the wealth residual plays a crucial role in the post-1978 growth-inequality relationship, further research is needed to examine more closely a set of candidate explanations for the increase in US wealth residual over the last four decades. For instance, by using a macroeconomic general-equilibrium model with heterogeneous agents, we can decompose the rise in wealth residual into four parts– differences in time-varying observable characteristics (e.g. risk exposure), transitory variations (e.g. luck), persistent components (e.g. ability) or economic rents.

Chapter 2

Rentier Premium and Wealth Inequality

Why do some people stay wealthy while others remain poor? To what extent can government policies affect wealth inequality? Based on the empirical evidence of persistently heterogeneous returns on wealth across households, this study introduces the ‘rentier premium’ into the standard heterogeneous agents dynamic stochastic general equilibrium model. This study also examines the effects of two different types of government on wealth distribution and welfare gains. I conclude that the rise in rentier premium and the government’s regressive policies have acted as key drivers in the co-evolution of rising wealth inequality and declining share of labour income in the United States since the 1980s.

Keywords: wealth inequality; rentier premium; heterogeneous agents DSGE
JEL: E21, E13, E25

2.1 Introduction

The aggregate wealth of the US has rapidly expanded compared to national income since the early 1980s. According to the World Inequality Database (WID.world)[1], the ratio of net national wealth to net national income remained stable at around 3.6 between 1949 and 1978. However, after the housing and equity booms, the ratio increased to around 5.4 in 2007. Despite the 2008 financial crisis, the ratio stayed at around 4.8 in 2015. However, concurrently, fast-growing wealth inequality has led to a renewed interest in the study of wealth distributions. As argued by a large number of studies (e.g. Atkinson, 2015; Piketty, 2014; Saez and Zucman, 2016; Stiglitz, 2015a), there seems to be a new set of stylised facts regarding the distributions of income and wealth for US households. Many of them are remarkably different from those at the centre of attention six decades ago.[2]

First, there has been a significant rise in wealth inequality and the greatest upsurge of wealth in the highest group. For example, according to the WID.world, households at the top 10% of the wealth distribution have held roughly seventy percent of the net worth (total assets minus liabilities) while households at the bottom 50% have held almost zero over the last four decades. Second, the concentration of wealth is much higher than that of earnings. The WID.world further reported that the Gini indexes for pre-tax income and net worth in 2014 were 0.60 and 0.86, respectively. Finally, the share of labour income has declined; the Bureau of Labour Statistics (BLS) has revealed that the share of labour income has decreased from a peak of 66.4% in 1960 to 57.0% in 2014.

Based on the above empirical backgrounds, we can raise two questions: (i) 'despite the rapid expansion of aggregate wealth, why do some people stay wealthy while others remain poor?' and (ii) 'to what extent can government policies affect wealth inequality?' Answering these fundamental questions requires an understanding of why the wealthy have the high marginal propensity to save, unlike the predictions of the precautionary savings hypothesis.[3] To address this puzzle, I use a standard heterogeneous agents dynamic stochastic general equilibrium (DSGE) model that can help unravel the determinants of wealth inequality in reality and evaluate the consequences of the government's redistributional policies.

[1] The WID.world is an open source online database that presents data series on the distribution of income and wealth in a number of countries (see Atkinson et al. (2011) and Alvaredo et al. (2013) for the details).

[2] Among the old stylised facts (e.g. Kaldor, 1961), the key elements were the constancy of the capital-output ratio and the relative factor shares.

[3] The standard incomplete-markets models of household predict that saving rates are either independent of or decreasing with wealth (e.g. Aiyagari, 1994; Bewley, 1986; De Nardi and Fella, 2017; Huggett, 1993; Imrohoroglu, 1992) because of the $\beta R < 1$ condition, where β is the discount factor and R is the (gross) risk-free rate (see Benhabib and Bisin (2018) and Stachurski and Toda (2019) for a discussion). However, precise evidence on this is missing until now, leaving us with an open empirical question (Fagereng et al., 2019).

The literature on the heterogeneous agents DSGE model has largely emphasised the role of labour earnings and precautionary savings in explaining wealth inequality (notably in Aiyagari, 1994; Bewley, 1986; Huggett, 1993). However, models of labour earnings inequality and precautionary savings alone struggle to generate results that fit the above stylised facts because the nature of precautionary savings implies that the saving rate decreases with the rise in wealth (e.g. Benhabib et al., 2017; De Nardi, 2015). This outcome is contradicted by the observed US data that indicate the wealthy keep saving at high rates.[4] Therefore, recent strands of papers have shifted attention from heterogeneous returns on human capital to heterogeneous returns on wealth to better account for the recent wealth concentration, focusing on topics such as bequests (e.g. De Nardi, 2004), heterogeneity in patience (e.g. Hendricks, 2007), entrepreneurial risk (e.g. Cagetti and De Nardi, 2006; Quadrini, 2000) and capital income risk (e.g. Benhabib et al., 2015).

This study builds on these ongoing theoretical studies to find uncovered forces that are driving wealth concentration. In specific, based on empirical evidence of persistently heterogeneous returns on wealth (e.g. Bach et al., 2016; Cao and Luo, 2017; Fagereng et al., 2016, 2020, 2019), this study introduces the ‘rentier class’, defined as those who have a privileged access to wealth residual – the unexplained increase in wealth that is not accompanied by any increase in productive investment (Stiglitz, 2015a) – into the standard heterogeneous agents DSGE model (or say, ‘the Bewley-Hugget-Aiyagari model’). More specifically, I assume that there are two cohorts of representative agents distinguished by asset ownership: the rentier class and the working class. The former is able to accumulate assets and works at the same time, while the latter gets its income solely from labour supply. As a result, the working class may have to pay rents normally set above the rate of return on capital in a perfectly competitive market for access to property owned by rentiers.

This study also deals with the debate on tax policies and regulations on the rent-related activities that redistribute a significant share of national income across households. For instance, one current major US policy debate is centred around the issue of growing large government deficits. To eliminate these deficits, the government must raise tax revenues in the upcoming years. However, the question for policy makers is, ‘should policy makers raise taxes on lower-income earners, middle-income earners or top income earners?’ Or concurrently, ‘should higher taxes be imposed on wealth, or should stricter regulations be implemented on the excess returns?’ To address these questions, this study examines the effect of the government’s choice between labour-friendly and rentier-friendly policies on wealth distribution and welfare gains.

[4] According to Saez and Zucman (2016), the saving rate for households at the bottom 90% of the wealth distribution was 2% on average during 1980-2011, while that for the top 10% was 24%. If we limit ourselves to the top 0.1%, the saving rate reached 53%.

Two innovative components in my model are positional competition and rentier premium. First, I incorporate concerns that people have about their relative positions. According to Hirsch (1977), positional goods are goods valued only by how they are distributed among the population, not by how many goods exist in total. The total benefit from all instances of a positional good is zero; that is, attempts to acquire them can only benefit one agent at the expense of another agent, resulting in a zero-sum game.[5] Second, I build a tractable dynamic stochastic general equilibrium model, where households face heterogeneous returns on wealth. Most importantly, I have introduced a concept of rentier premium[6], defined as the gap between the return on asset ownership and the return on capital used for the production of real output in a perfectly competitive market.[7]

The main findings in this study are as follows. First, the model provides implications for our understanding of cross-sectional distribution, in line with the findings of the recent wealth inequality literature (e.g. Benhabib et al., 2019). More specifically, I have shown that stochastic processes of labour income and rentier premium provide a better way to understand distribution of wealth and income than the models that rely only on labour income. Each of them has a distinct role in the dynamics of wealth accumulation. A stochastic process of labour income allows agents to avoid poverty traps and to move upward near the borrowing constraints due to a precautionary savings motive, while a stochastic process of rentier premium has negative effects on households at low levels of wealth. Second, capital income and rentier premium help us explain the thick top tail of the wealth distribution as the wealthy get richer by accumulating their wealth at higher rates than the poor. Most strikingly, stochastic processes of labour income and rentier premium generate wealth fluctuation ("within-class" inequality), while rentier premium under positional competition generates a net transfer of wealth to the owner of assets from the renter of assets ("between-class" inequality). Finally, the government's policies favouring the rentier class have generated the increase in the top wealth shares and the decrease in welfare gains for the majority of households who live in poverty. This result implies that the government can alter the dynamics of wealth accumulation by way of tax policies or regulations on rent-related activities.

[5] For instance, let us imagine that rich individuals compete for houses in certain fashionable parts of New York. As the wealthy get wealthier, they compete more vigorously for these real estates, and the prices of those assets with fixed supply increase without any increase in real output. Accordingly, if some individuals consume positional goods, other individuals must be included in the consumption of related negative quantities.

[6] There are a number of papers studying the 'economic premium'. The most famous concept is *skill premium*, defined as the ratio of the wages of skilled to unskilled workers (e.g. Heathcote et al., 2010; Tinbergen, 1974). Also, Bloom et al. (2018) propose *large-firm wage premium*, defined as the gap between the average wages of employees in large versus small firms. Unlike most papers that focus on labour income differentials, *rentier premium* in this study is related to capital income differentials.

[7] Stiglitz (2015e) argued that the return on asset ownership falls into four parts: (i) the pure rate of interest (e.g. the T-bill rate); (ii) risk premium; (iii) the returns to the ability of the investment manager; and (iv) rents.

This study aims to make three contributions. First, it helps us better explain the new stylised facts by incorporating both precautionary savings (the so-called "thrift") and the inborn difference in asset ownership combined with rentier premium (the so-called "exploitation"). When an idiosyncratic labour earnings risk is the only source of heterogeneity, the model predicts far less cross-sectional dispersion and right-hand-side skewness than the actual US data (e.g. Benhabib and Bisin, 2018). Also, said model cannot fully explain why the share of labour income has declined since the 1980s (e.g. De Nardi, 2015). However, an extension that relies on a small amount of heterogeneity in rentier premium can better replicate the key features of the US wealth and income data; that is, the skewed distribution in the right-hand-side tail and the decline in labour income shares. Second, this study has discovered that wealth inequality is the result of preferences for redistribution as much as underlying economic forces. Even though market forces help shape the degree of inequality, government policies shape those market forces. Accordingly, the government can set and enforce rules to transfer wealth from the bottom to the top, or vice versa.

The rest of the paper is organised as follows. Section 2 reviews the related literature. Section 3 introduces the concept of rentier premium and provides empirical evidence of its rationality. Section 4 describes the model economy, while Section 5 presents the results of quantitative experiments and draws policy implications. The last section concludes.

2.2 Literature Review

This study builds on two strands of the wealth inequality literature: (i) the heterogeneous agents DSGE model and (ii) rent-related activities. While a rent-related activity is a conceptually crucial ingredient of this study, my main method rather closely follows the DSGE model with heterogeneous agents.

The literature on the heterogeneous agents DSGE model studying wealth inequality can be divided into two categories. The first category features heterogeneous agents who face partially insurable idiosyncratic earnings shocks (notably, Aiyagari, 1994; Bewley, 1986; Huggett, 1993).[8] It is closely related to 'within-class' inequality such that difference in labour productivity ("human capital") leads to difference in wealth. However, the Bewley-Huggett-Aiyagari model, relying solely on labour income differentials and precautionary savings,

[8]The Bewley-Huggett-Aiyagari model has been successful in the study of several macroeconomic phenomena of interest. Calibrated versions of this model have been used to study the welfare costs of inflation (Imrohoroglu, 1992), asset pricing (Huggett, 1993), unemployment benefits (Hansen and Imrohoroglu, 1992), fiscal policy (Heathcote, 2005) and labour productivity (Perri and Krueger, 2009).

generally finds it difficult to replicate the thick right-hand-side tail of the wealth distribution observed in the data (e.g. Benhabib and Bisin, 2018; De Nardi, 2015). Benhabib et al. (2017) also mention that the cross-country data does not display a statistically significant correlation between income inequality and wealth inequality, indicating a significant role for other factors in generating wealth inequality. The second category includes several cohorts of representative agents who face class-specific aggregate shocks (e.g. Guvenen, 2009; Krusell and Smith, 1998). It is closely related to 'between-class' inequality that different capabilities can manifest as a restricted access of a part of the population to certain institutions like housing, knowledge and financial markets. This study embraces 'within-class' inequality and 'between-class' inequality at the same time by combining the above two types of models to make our analysis become more realistically.

This study also draws on the literature on heterogeneous returns on wealth. Motivated by the empirical fact that wealth distribution generally tends to be much more skewed than earnings distribution, an important question in the literature has been whether wealth distribution becomes more skewed due to factors unrelated to skewed earnings distribution. Therefore, various works have attempted to uncover forces that keep the saving rates of the wealthy high and thus generate higher wealth concentration in the hands of a small fraction of households. These 'forces' include heterogeneity in patience (Hendricks, 2007; Krusell and Smith, 1998), bequests and transmission of human capital across generations (Cagetti and De Nardi, 2006), the higher earnings risk for the top earners (Castaneda et al., 2003; Kindermann and Krueger, 2014), capital income risk (Benhabib et al., 2015) and entrepreneurs with Schumpeterian creative destruction risk (Jones and Kim, 2018).[9]

There are growing number of studies in measuring the degree of heterogeneity in the returns on wealth. An earlier attempt is Flavin and Yamashita (2002), who compare the risk and return on housing to those of various asset categories and portfolios. More recently, by evaluating the portfolios of wealthy households in Sweden, Bach et al. (2016) find that the annualised returns on financial wealth are on average four percentage points higher for

[9]Contrary to the infinite-horizon setup of Krusell and Smith (1998), Hendricks (2007) studies the heterogeneity in the discount factor in a life-cycle framework with purely accidental bequests. He argues that the presence of an addition saving motive (e.g. retirement) generates a higher wealth-income ratio than the infinite-horizon model. Cagetti and De Nardi (2006) build a model of entrepreneurship with perfectly altruistic and finitely-lived agents who are endowed with two types of abilities, a worker and an entrepreneur. They endogenise the return to being an entrepreneur into the production function. Castaneda et al. (2003) were the first to highlight how a stochastic process featuring right-hand-side skewness may help generate a long right tail in wealth distribution by calibrating so that the highest productivity level is more than 100 times higher than the second highest. Benhabib et al. (2015) conduct a quantitative exploration of the extent to which idiosyncratic rates of return, skewed earnings risk, and luxury bequest motives account for both the US cross-sectional wealth distribution and its intergenerational wealth mobility. Finally, Jones and Kim (2018) study the link between innovation and top income inequality by exploring a model in which entrepreneurs exert effort to generate exponential growth in their incomes, which tends to raise inequality.

households at the top 1% of the wealth distribution, compared with the median household. They argue that these high returns are primarily compensations for bearing high level of systemic risk. More specifically, wealthier households allocate a much larger share of their funds to risky assets and within the risk portfolio wealthier households load more aggressively on market risk. In contrast, a work by Fagereng et al. (2016), using Norwegian tax data, documents that there is a spread of five percentage points in the rates of return between the top 10% and the bottom 10% of the returns distribution. They reveal that heterogeneity holds within asset classes, rather than just being the results of a different portfolio mix between safe and risky assets. That is, returns are heterogeneous even within asset classes and are positively correlated with wealth. Fagereng et al. (2020) further argue that heterogeneity in returns does not arise merely from differences in the allocation of wealth between safe and risky assets, and individual returns exhibit substantial persistence over time. For the US data, Cao and Luo (2017) documents that the returns on wealth across households are significantly and persistently heterogeneous after the early 1980s. Motivated by this observation, their DSGE model shows that financial deregulation and a reduction in corporate tax rates can explain the joint evolution of rising wealth inequality and declining share of labour income since the 1980s.

This study also draws on the literature that studies why the wealthy save so much. Carroll (1998) argues that the savings behaviour of the wealthy cannot be explained by models in which the only purpose of wealth accumulation is to finance future consumption. Indeed, actual data suggest that the wealthy behave in ways that are substantially different from the rest of the population. The author concludes that the model that explains the relevant facts is one in which either consumers regard wealth accumulation as an end in itself or wealth enters the utility function directly as a luxury good. In contrast, this study considers an alternative explanation that the savings behaviour of the wealthy is largely determined by what I call rent-related activities.[10] Similarly, Fagereng et al. (2019) argue that the relation between saving rates and wealth crucially depends on whether saving includes capital gains. That is, saving rates including capital gains increase markedly with wealth because the wealthy own assets that experience 'persistent capital gains' which they hold onto instead of selling them off to consume.

While many DSGE models have focused on the causes and the effects of wealth inequality, models of accommodating the phenomenon of rent-related activities are few. I present the first study of introducing the concept of 'positional competition' into the heterogeneous agents DSGE model. Hirsch (1977) defines positional goods as goods valued only by how

[10] Since rents are defined as 'that part of the payment to an owner of resources over and above that which those resources could command in any alternative use' (Buchanan, 1980), rent-related activities are thus denoted as all economic activities that can create rents.

they are distributed among the population, not by how many goods there are in total. In positional competition, the total benefit from all instances of a positional good is zero; that is, attempts to acquire them can only benefit one agent at the expense of another agent, resulting in a zero-sum game. The recent work of Stiglitz (2015a) argues that a significant increase in the wealth-income ratio in recent decades is due to an increase in the present value of land, which has a property of a positional good, combined with more generous provision of credit by the financial system. He uses the concept of a positional good to explain rent-related activities, where some earnings reflect the capture of transferring pre-existing output rather than the creation of new output.

Stiglitz (2015a) also introduces the concept of the 'wealth residual' to identify the causes of the divergence between wealth and capital. For instance, he points out that a large fraction of the increase in the recent US wealth is an increase in the value of land, not in the amount of capital. Considering the fact that national income data on savings can account for at most three-quarters of the wealth growth in the US, he refers to a large unexplained increase in wealth that is not related to production of real output as the 'wealth residual'. He argues that it is associated with rents from land, monopoly power and asymmetric information. This study shares this view, but extends this literature by incorporating it into the DSGE model.

2.3 Empirical Motivation for Rentier Premium

In this section I define rentier premium and estimate it through the national accounts of the US and the associated statistics. And then, based on empirical evidence, this study presents three assumptions that are key components in my model economy.

2.3.1 Definition

This study has defined rentier premium as the gap between the return on asset ownership and the return on capital used for the production of real output in a perfectly competitive market. To figure out this definition, we need to first understand two concepts: (i) rents and (ii) wealth residual.

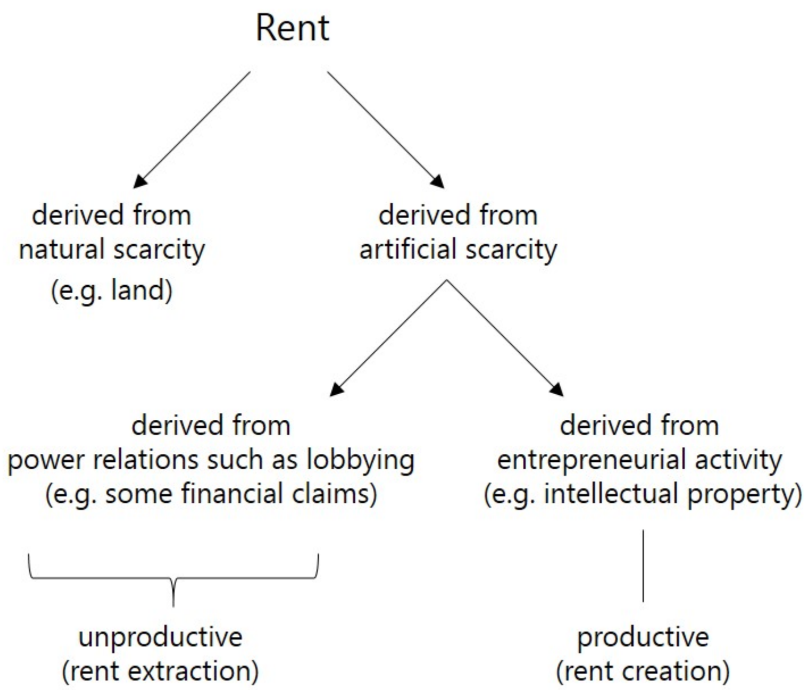
First, rents are defined as 'that part of the payment to an owner of resources over and above that which those resources could command in any alternative use' (Buchanan, 1980); that is, it is 'the receipt in excess of the owner's opportunity cost of the resources' (Chang, 1994). Though rents can take on many different forms, as in Figure 2.1, we can categorise them in the following ways. The first category comprises rents of *natural scarcity*, and the obvious example is land. Households derive their unearned income simply through the 'possession' of scarce assets. For instance, if the fashionable location of certain land is owned by a few households, they can rent it out to those who have no land for a high price. This income is rent, not profits from a productive activity, as the landlords do nothing to earn, drawing exclusively from their asset ownership.

However, the argument that rents arise from the scarcity of an asset is less convincing when these are reproducible. For example, specific talents and skills may be temporarily scarce in specific positions and for specific markets, but there is no intrinsic scarcity to justify the rentier income for the owners of these skills in the long run. Therefore, the second category comprises rents of *artificial scarcity*. It is imposed by the 'rules of the game' (i.e., institutional arrangement and power relations), which determine who gets an income from privileged access to specific assets and who will have to earn an income through traditional entrepreneurial activity or the provision of labour. Obvious examples are intellectual property, the access to which is artificially restricted by laws, and some financial claims (e.g., private equity funds), the access to which is artificially restricted by asymmetric information (Botta et al., 2019)[11].

Rents can be unproductive if they are derived from either the natural scarcity of assets or the power relations between economic interest groups and governments, unaccompanied by any increase in the flow of goods and service. In this case, the increase in rental price leads to *rent extraction*, the net wealth transfer from the one who has no assets (called "the poor") to the owner of those assets (called "the wealthy"). In addition, the very act of seeking rents imposes additional costs on society in the form of the efforts and resources spent by rent-seekers on gaining access to rents such as corporate lobbying for financial deregulations or tax cuts (Krueger, 1974), rather than productive activities. The exception to this is innovation rent, which is derived from entrepreneurial activities (e.g. intellectual property), since these are dynamically productive (Chang, 1994) and normally dissipate as others imitate the entrepreneurial innovation (Schumpeter, 1942).

[11]Botta et al. (2019) argue that the financial sector is the paradigmatic industry that can benefit from exploitation rents (i.e., rents accruing to economic agents as a consequence of monopoly power). For instance, financiers manage a huge amount of asymmetric information because they are in the middle of complex and often opaque network relationships, which certainly give rise to economic dynamics inconsistent with perfect competition.

Fig. 2.1 A categorisation of rents



Second, wealth and capital are distinctly different concepts. If wealth is equated with capital, the increase in wealth should be associated with a decline in the return to capital and an increase in real wages. However, this hypothesis is contradicted by the actual US data: the increase in the wealth-income ratio and the stagnation of median real wages during the past three decades (e.g., Eggertsson et al., 2018). Hence, there might be an unexplained causal factor of the increase in the wealth-income ratio, referred to as 'wealth residual' (Stiglitz, 2015a).

On top of Stiglitz (2015a)'s theoretical perspective, this paper is, I believe, the first to measure wealth residual by combining the Financial Accounts of the Federal Reserve (FED) and the Fixed Assets Accounts of the Bureau of Economic Analysis (BEA). First, household wealth is calculated by the market value of aggregate assets, net of their liabilities, from the Financial Accounts Table B.1 (FED code: FL152090005).[12] Second, capital is calculated by the current-cost net stock of fixed assets from the Fixed Assets Table 1.1 (BEA

[12] According to the FED, wealth consists of nonfinancial wealth and financial wealth. Nonfinancial wealth includes land, structures, machines and patent rights. Financial wealth includes foreign deposits, checkable deposits and currency, time and savings deposits, money market fund shares, debt securities, loans, corporate equities, mutual fund shares, trade receivables, life insurance reserves, pension entitlements, equity in noncorporate business and miscellaneous assets. To a large extent, households' financial position ultimately represents a claim on the nonfinancial assets of corporate business or government.

code: K1PTOTL1ES000).[13] Accordingly, wealth residual is naturally calculated by the gap between household wealth and capital. Next, the upsurge of wealth residual is being driven by either land[14] or rent-related finance (i.e., some financial assets that is not linked to capital or land) or both. I estimate a portion of land in wealth residual by using a concept of Davis and Heathcote (2007), who see that a house can be seen as a bundle comprising a reproducible tangible structure and a non-reproducible plot of land. Since the structure can be priced explicitly as the replacement-cost, the residuals are assumed to be the value of a home location[15] – referred to as 'scarcity rents' that are evidently capitalised in nonfinancial assets, as argued by Fagereng et al. (2019), Piketty (2014), Rognlie (2015) and Stiglitz (2015a,e).

In this context, I calculate wealth residual values for land by subtracting the replacement-cost value of structures from the market value of real estate from the Financial Accounts Table B.101 (FED code: LM155035005, LM155012605 and LM165013665). Rent-related finance is then naturally calculated as the gap between wealth residual and land – referred to as 'monopoly rents' that are unobtrusively capitalised in financial assets, as argued by Brun and Gonzalez (2017) and Stiglitz (2015a).

Table 2.1 shows the result for measuring wealth residual for the US during 1978-2018. Most notably, in 2018, roughly 40% of wealth in US households was not related to production of real output but was instead associated with rents from wealth residual. During the last four decades, wealth residual has increased by 59.3 times, while capital has increased by only 8.7 times. As a result, the ratio of wealth residual to total wealth in 2018 (39.7%) was much larger than 8.8% in 1978, implying that not all households' finances lead to productive investment. In addition, by far the largest contributor to the increase in the wealth-income ratio from 409.4% in 1978 to 607.6% in 2018 is rent-related finance, while land was a crucial source of wealth residual until the Subprime mortgage crisis of 2007. Most importantly, if we look at the co-evolution of top wealth inequality and growing wealth residual, the ownership of wealth residual may be a source of excessive returns (i.e. the returns over and above the return on capital) – that is, 'rentier premium'.

[13] According to the BEA, buildings, machines, software, and even the intellectual property product (e.g. creation of a song) can be fixed assets, if they are used for a year or more in the production of goods or services. Houses and apartments are also included, even if the owner lives there.

[14] Nonfinancial assets consist of real estate and consumer durables. Since consumer durables are solely calculated by the replacement-cost value, there must be no wealth residual for consumer durables.

[15] If housing were simply a manufactured good, the price of housing would be determined by construction costs, and housing prices would increase at roughly the same rate as the price of other goods. But since housing has a valuable location component that is in limited supply, an increase in the demand for housing can link directly to increases in the value of good locations.

Table 2.1 Wealth residual in the United States, 1978-2018

(Trillions of dollars, amounts outstanding end of period)

	1978(E)	1988	1998	2008	2018(F)	F/E
Wealth (A)	8.0	19.8	38.8	58.1	104.3	13.1
Capital (B)	7.3	15.2	24.5	46.0	62.9	8.7
Wealth residual (C=A-B)	0.7	4.6	14.3	12.1	41.4	59.3
Land (D)	0.5	2.6	3.4	5.0	8.9	16.5
Rent-related finance (C-D)	0.2	2.0	10.9	7.1	32.5	207.5
Wealth residual ratio (C/A, %)	8.8	23.1	36.8	20.8	39.7	4.5
Wealth-income ratio (%)	409.4	455.1	498.0	526.8	607.6	1.5

Note: Wealth-income ratio is the ratio of households' net worth as a percentage of personal income.
Source: The author's own estimation based on the Federal Reserve and the Bureau of Economic Analysis

2.3.2 Rentier premium

A question naturally arises from the above data: who gains (or losses) from the upsurge of wealth residual? I attempt to find a clue from the Distributional National Accounts (DNAs) compiled by Piketty et al. (2018).[16] By combining Table B2d, B3, E2, E2b and E3 of this statistic, I estimate the average income accruing to labour and capital, and the average wealth for each asset composition – equities, non-corporate business assets, real estate, fixed income claims and pensions – held by the bottom 50%, the middle 40%, the top 10% and the top 1% of the US households. All our data are expressed in constant 2018 dollars to adjust for inflation, using the national income deflator.

Table 2.2 describes the estimated portfolio of the US households in 2016. It clearly shows that households at the top 10% make a living through all income sources with about 2.3 millions of net worth while households at the bottom 50% significantly depend on labour income with little net worth. If we look at the top 1%, they earn a huge amount of capital income mostly from fixed income claims and equities unlike the middle 40% who significantly rely on pensions and real estate.

Furthermore, based on historical data of the DNAs, I have estimated the return on the portfolio of each wealth group over the period of 1962-2016 as $r_t^g = \sum_i A_{i,t}^g r_{i,t} / \sum_i A_{i,t}^g$ where $A_{i,t}^g$ refers to the amount of asset i held by wealth group g at t and $r_{i,t}$ refers to the return on asset i at t. Based on Chang (1994)'s definition on rents, I then define the gap between

[16] Data files and replication files are available at http://gabriel-zucmann.eu.usdina/.

Table 2.2 Households' portfolio (pre-tax, pre-transfer) in the Data, 2016

(Thousands of dollars, %)

	Top 1%		Top 10%		Middle 40%		Bottom 50%	
Labour income	568.3	(42.1)	184.7	(57.6)	58.8	(81.8)	17.3	(93.4)
Capital income	781.7	(57.9)	136.0	(42.4)	13.1	(18.2)	1.2	(6.6)
Total income	1,350.0	(100.0)	320.7	(100.0)	71.9	(100.0)	18.5	(100.0)
Equities	3,887.7	(32.2)	564.3	(24.0)	16.3	(7.0)	1.7	(70.4)
Business	1,454.3	(12.0)	267.3	(11.4)	24.3	(10.4)	1.4	(56.7)
Real estate	1,127.6	(9.3)	389.1	(16.5)	68.3	(29.3)	1.8	(76.8)
Fixed claims	4,042.5	(33.5)	556.6	(23.6)	-1.8	(-0.8)	-15.4	(-639.8)
Pensions	1,564.7	(13.0)	577.3	(24.5)	126.2	(54.1)	12.8	(535.9)
Total wealth	12,076.8	(100.0)	2,354.6	(100.0)	233.4	(100.0)	2.4	(100.0)

Notes: The share of each component is in parenthesis. Equities, corporate equities including S-corporation and money market shares; Business, non-corporate business assets including sole proprietorships, farms, partnerships, and intellectual property products; Real estate, owner- and tenant- occupied housing, net of mortgage debt; Fixed claims, currency, deposits, bonds, and other interest-paying assets, net of non-mortgage debts; and Pensions, pensions and insurances.
Source: The author's estimation based on the Distributional National Accounts (Piketty et al., 2018)

the return on the portfolio of the top 1% and the risk-free rate, measured by the 3-month treasury-bill rate from the Federal Reserve Bank at St. Louis, as 'rentier premium'. We can further decompose this rentier premium as top 1% premium, top 10% premium and middle 40% premium. For instance, top 1% premium refers to the gap between the return on the portfolio of the top 1% and that of the top 10%. This decomposition suggests that larger premium for the wealthier group is mainly due to a further monopoly of access to high-return equities (e.g. Lusardi et al., 2017; Stiglitz, 2015a). For instance, given that financial firms that offer these high-return products requires high minimum investments, as wealth increases, an individual's ability to absorb risk also increases, enabling the wealthy to take on more high-risk and high-return products.[17]

Table 2.3 shows that the estimated average return on the portfolio of the top 1% over the entire period is 9.48%, which is 4.73 percentage points higher than the risk-free rate (4.75%). In addition, rentier premium has increased from 1.81% in 1980-89 to 6.84% in 2010-16. The decomposition of rentier premium indicates that the premium becomes larger, as those who belong to the wealthier group. Furthermore, top 1% premium and top 10% premium have declined from 1.87% and 2.11% in 1962-69 to 0.18% and 0.08% in 2010-16, respectively, while middle 40% premium has increased from 0.12% in 1980-89 to 6.58% in 2010-16.

[17] For example, private equity finance (PEF) typically has horizons of 10-13 years with the high minimum required commitment (at least a median of $1 million for funds of $100 million) but at most 5% of US households can afford to invest in it (Shakhnov, 2014).

These results imply that those who had enough funds to invest high-return products during 1980-2016, their wealth have been increased geometrically; however, those who did not (i.e., the bottom 50%) have remained poor, in line with the findings of Fagereng et al. (2019) who argue that wealthier households hold assets like stocks and housing that experience persistent capital gains.

Table 2.3 Rentier premium in the United States, 1962-2016

(%, %p)

	62-69	70-79	80-89	90-99	00-09	10-16	62-16
Return on $P_{T1\%}$ (A)	11.38	11.51	10.35	8.92	7.27	6.99	9.48
Return on $P_{T10\%}$ (B)	9.51	10.15	9.80	8.46	6.88	6.81	8.68
Return on $P_{M40\%}$ (C)	7.40	8.46	8.66	7.78	6.32	6.73	7.62
Risk-free rate (D)	4.79	6.55	8.54	4.75	2.36	0.15	4.75
Top 1% premium (=A-B)	1.87	1.36	0.55	0.46	0.39	0.18	0.80
Top 10% premium (=B-C)	2.11	1.69	1.14	0.68	0.56	0.08	1.06
Middle 40% premium (=C-D)	2.61	1.91	0.12	3.03	3.96	6.58	2.87
Rentier premium (=A-D)	6.59	4.96	1.81	4.17	4.91	6.84	4.73

Notes: $P_{T1\%}$, $P_{T10\%}$, and $P_{M40\%}$ refer to the portfolio of the top 1%, the top 10%, and the middle 40%, respectively. The risk-free rate is measured by 3-month treasury bill rate.
Source: The author’s estimation based on the Distributional National Accounts (Piketty et al., 2018)

Moreover, many studies have witnessed that households face heterogeneous returns on wealth within wealth groups (e.g. Bach et al., 2016; Cao and Luo, 2017; Fagereng et al., 2016, 2020). I calculate the distribution of returns on wealth for the US households during 1984-2013 by using the Panel Study of Income Dynamics recompiled by Cao and Luo (2017).

Table 2.4 describes the distributions of annualised returns for pooling observations during the entire period, showing that the returns on wealth are heterogeneous across households. The cross-sectional standard deviation of annualised returns is 20.3 percentage points without capital gains and 37.2 percentage points with capital gains. The majority of households report little or no income from assets, and the distribution exhibits a long right tail. Households at the top 10% earn an annualised return above 10%, and households at the top 5% earn an annualised return above 25%, which is much higher than the risk-free rate (6.1%). When it comes to capital gains, the distribution of returns exhibits even more dispersion. While there are still masses of households who earn a low return close to zero, households at the top 10% earn an annualised return above 39%. Even further, there is a significant fraction of households who report capital losses, while households at the top 5% earn an annualised return above 70%.

Table 2.4 Heterogeneous returns on wealth across the US Households, 1984-2013

(%, %p)

	percentile of returns							
	10	25	50	75	90	95	mean	s.d.
Returns without capital gains	0.0	0.0	0.7	4.8	14.0	28.2	6.3	20.3
Returns with capital gains	-20.2	0.0	0.4	13.0	39.2	70.4	7.9	37.2

Source: The author utilised the Panel Study of Income Dynamics recompiled by Cao and Luo (2017)

Finally, we can ask ourselves why the rentier premium emerges. According to Fagereng et al. (2020, 2019) and Stiglitz (2015e), the return on asset ownership falls into four parts: (i) the pure rate of interest (e.g. the T-bill rate); (ii) risk premium; (iii) personal ability of the investment; and (iv) monopoly or scarcity rents. In this regard, the rentier premium, defined as the gap between the return on asset ownership and the return on safe assets (i.e. opportunity cost) based on Chang (1994)'s definition on economic rents, is a more encompassing term than the risk premium because it includes differences in time-varying observable characteristics (e.g. risk exposure), transitory variations (e.g. good or bad luck), persistent components (attributable to both observable factors, such as education, and unobservable ones, such as ability or risk tolerance), and economic rents (e.g. a privileged access to wealth residual). Accordingly, even though differences in risk exposure are important determinants of persistent return heterogeneity (e.g. Bach et al., 2016), the persistent rentier premium is closely related to a privileged access to wealth residual (e.g. Lusardi et al., 2017; Pagano, 2014; Stiglitz, 2015a,b,c,d,e). Benhabib et al. (2015) show that a positive correlation between returns to wealth and wealth is well documented in the actual US data, even though it would be due to personal ability or risk premium. In particular, Saez and Zucman (2016) show that well-off families have access to high-yielding investment strategies and opportunities than the rest of the population. According to Fagereng et al. (2020) and Piketty (2014), even after consideration of these confounding factors, returns are significantly increasing in wealth; for instance, the rich are able to bear more risk, and disproportionately own equities that give high returns.

The privileged access to wealth residual is closely related to asymmetric information (e.g. Lusardi et al., 2017; Stiglitz, 2015a,e). For instance, once we assume that not everyone has equal access to knowledge, knowing something slightly before others do can yield high returns due to insider trading. In particular, a few individuals would benefit from superiority in financial knowledge, as financial markets have become more complex.[18] Furthermore, given that financial firms that offer these complex products require high minimum investments, as wealth increases, an individual's ability to absorb risk also increases, enabling the wealthy to take on more high-risk and high-return products.[19]

More fundamentally, the government allowed through deregulation and lax standards banks to lend more, but much of that money did not go for creating new businesses or increasing the stock of capital goods (Stiglitz, 2015d). The effect of the expansion of credit has actually been an increase in the value of land and other fixed assets. Changes in financial regulations and monetary policy can lead to more wealth inequality. For instance, an increased flow of credit combined with a change in regulation that allows more lending against collateral, such as land; those who hold wealth become wealthier. Those who have little or no wealth do not benefit from that kind of credit expansion.

Finally, the growing influence of corporate lobbying to defend unproductive rents (e.g. Drutman, 2015; Esteban and Ray, 2006), such as lobbying for deregulations for activities with high entry barriers or for tax cuts for the wealthy, can create a *persistent* monopoly of access to wealth residual. For instance, lobbying spending of the US corporations during 1998-2017, obtained from the Center for Responsive Politics (CRP), confirms this. Until the global financial crisis in 2008, the financial sector was the top spender on lobbying and has invested almost $8 billion in political influence purchasing over the last two decade. Also, the 'lobbying intensity' that is defined as the ratio of the portion of lobbying spending to that of sectoral GDP in the financial sector (0.74) is much greater than that of the non-financial sector (0.58), including manufacturing, retail trade and wholesale trade. Therefore, the resulting deregulations on financial industries have provided fertile ground for pursuing persistent rent-related activities (e.g. Cao and Luo, 2017; Hubmer et al., 2016; Kuhn et al., 2020).

[18] Some people may argue that the higher returns earned by the wealthy seem to be mainly driven by exceptional investment skills rather than by privileged access to private information. However, we can refute this argument once we assume that there is no 'entry barrier' to these complex financial activities other than 'natural' differences in people's intelligence in perfect capital market where clever poor people can borrow money to be educated in these things. In reality, most global banks, such as Barclays, BNP Paribas, and Citibank, have a separate business unit with dedicated teams of client advisors and product specialists exclusively for the wealthy. They provide a wide range of investment opportunities, such as private equity finance (PEF).

[19] According to Shakhnov (2014), PEF typically has horizons of 10-13 years with the high minimum required commitment (at least a median of $1 million for funds of $100 million) but at most 5% of US households can afford to invest in it. Of course, this long-term horizon and high entry costs which erect high entry barrier are compensated for by substantially higher returns. For instance, the annualised return on PEF (9.2 percent) was much higher than those for stock (3.2 percent), T-bond (0.4 percent) and than the rate of inflation (2.4 percent) during 1997-2011.

2.3.3 Key assumptions

Based on the empirical evidence in the previous sub-section, I make the premise that a household's wealth consist of both capital and wealth residual.

Assumption 1 (Rentier premium) Wealth residual may create heterogeneous rentier premium across asset owners.[20]

Let us assume that there are heterogeneous returns on assets (e.g. Cao and Luo, 2017). It implies that an agent i's wealth accumulation process may be described as

$$c_{i,t} + a_{i,t+1} - a_{i,t} = \underbrace{w_t l_{i,t}}_{\text{labour income}} + \underbrace{(r_t^k - \delta)\lambda a_{i,t}}_{\text{capital income}} + \underbrace{r_{i,t}^{nk}(1-\lambda)a_{i,t}}_{\text{rentier premium}} \tag{2.1}$$

where c_i is consumption, a_i is total assets, w is the wage rate, l_i is labour supply, r^k is the rental price of capital (k), r_i^{nk} is the rate of return on wealth residual (nk), $1-\lambda$ is the portion of wealth residual in total assets, and δ is the depreciation rate. Rearranging equation (2.1) yields

$$c_{i,t} + a_{i,t+1} - a_{i,t} = w_t l_{i,t} + (r_t + \theta_{i,t})a_{i,t}, \qquad \theta_{i,t} = (1-\lambda)(r_{i,t}^{nk} - r_t) \in [-1, \infty) \tag{2.2}$$

where r (=$r^k - \delta$) is the market interest rate and θ_i is rentier premium. Therefore, in equilibrium, an agent i's rate of return on total assets, r_i, is described as

$$r_i = \bar{r} + \theta_i \tag{2.3}$$

where $\bar{r}$ is the risk-free rate.

Assumption 2 (Positional competition) For certain wealth residual with fixed supply (denote "land"), the aggregate return on land across households is zero.

[20] Mathematically, Assumption 1 is related to the work of Benhabib et al. (2017) that displays that the 'random growth process' in wealth accumulation generates an extremely skewed distribution under appropriate regularity conditions. More specifically, an agent i's wealth accumulation is described as

$$a_{i,t+1} = w_t l_{i,t} + (1 + r_{i,t} - \frac{c_{i,t}}{a_{i,t}})a_{i,t}$$

where r_i is an idiosyncratic stochastic growth component. The high saving rates for the wealthy that is generated from either highly persistent differences in r_i or c_i/a_i - that is, decreasing in wealth - provide an amplifying wealth concentration mechanism.

As in Stiglitz (2015d), land generates a rent that is fixed and lasts in perpetuity. Then, an asset owner i's total income, y_i, is given as

$$y_{i,t} = w_t l_{i,t} + r_t k_{i,t} + R_{i,t}\bar{T}_i \tag{2.4}$$

where R_i is the return on the fixed amount of land ($\bar{T}_i$). Then, let us assume that land is a positional good. Under positional competition, the aggregate return on land across households is zero, so that

$$\int R_{i,t}\bar{T}_i di = 0, \qquad \forall\, t. \tag{2.5}$$

Furthermore, an asset owner i's total assets is simply described as

$$a_i = k_i + p\bar{T}_i = k_i + \frac{R_i}{r}\bar{T}_i = k_i + \frac{R_i}{F_K}\bar{T}_i \tag{2.6}$$

where p is the (relative) price of land ($p = \sum_{t=0}^{\infty}(1+r)^t R_i$) and r equals the return on capital, F_K, for simplicity, so that

$$\frac{da_i}{dk_i} = 1 - \frac{R_i\bar{T}_i F_{KK}}{{F_K}^2} > 1 \qquad if \quad R_i > 0. \tag{2.7}$$

This implies that there can be an increase in the price of land without any increase in real output. Following Stiglitz (2015a), assume that the demand for a positional good is given by $M(a,p)$ with equilibrium given by $M(a,p) = p\bar{T}$. Thus, $K = a - p(a)\bar{T}$. Then,

$$\frac{dk}{da} = 1 - p'(a)\bar{T} = 1 - \frac{M_a\bar{T}}{1-M_p} < 1.$$

So, if the wealth elasticity of the demand for a positional good (M_a) is large enough and if the price elasticity of a positional good (M_p) is small enough, then an increase in a may even be associated with a decrease in k.

Assumption 3 (Heterogeneous classes) Each household belongs to either the rentier class or the working class in terms of their asset ownership.

Table 2.2 indicated that the whole population can be divided into two groups in terms of their asset ownership – the rentier class and the working class. In the same vein, as in Table 2.5, we can naturally assume that the rentier class is able to accumulate assets and work at the same time, while the working class earns their income entirely from labour supply. Most importantly, due to the monopoly of access to certain wealth residual, the rentier class can gain additional capital at the expense of the working class through rentier premium.

Table 2.5 Households' portfolio (pre-tax, pre-transfer) in the Model

	a rentier i	a worker j	whole economy
Labour income (A)	wl_i	wl_j	$w\{\int_{i\in R} l_i di + \int_{j\in W} l_j dj\}$
Capital income (B)	$r(k_i + nk_i)$	-	$r\int_{i\in R}(k_i + nk_i)di$
Rentier premium (C)	$\theta_i nk_i$	-$\theta_{i'} nk_{i'}$	$\int_{i\in R} \theta_i nk_i di$ - $\int_{i'\in R} \theta_{i'} nk_{i'} di'$
Total (=A+B+C)	wl_i + $r(k_i + nk_i)$	wl_j - $\theta_{i'} nk_{i'}$	$w\{\int_{i\in R} l_i di + \int_{j\in W} l_j dj\}$
	+ $\theta_i nk_i$		+ $r\int_{i\in R}(k_i + nk_i)di$ + $\hat{\theta}\int_{i\in R} nk_i di$
Net worth	k_i + nk_i	-	$\int_{i\in R}(k_i + nk_i)di$

Notes: w: wage rate; l: labour supply; r: return on capital; θ: rentier premium; k: capital; nk: wealth residual; R: rentier class; W: working class; $\hat{\theta}$: the weighted average rate of return on nk where $\hat{\theta}$=0 if all wealth residual are in positional competition.

2.4 The Model

This section formally describes my model economy. Two key forces that generate wealth inequality are (i) partially insurable income risk with different degrees of precautionary savings that all individuals face ("within-class inequality") and (ii) the innate difference in asset ownership between the rentier class and the working class combined with rentier premium ("between-class inequality"). The remaining setup follows the standard heterogeneous agent DSGE model with aggregate uncertainty (e.g. Benhabib et al., 2015).[21]

Time is discrete and indexed by $t \in [0, \infty)$, and there is a continuum of agents indexed by $i \in [0, 1]$. The total population is normalised to one.

Demographics: Agents are divided into two classes (e.g. Judd, 1985). The population share of the first class ("rentier class"), who is able to accumulate assets and work at the same time, is λ, while the share of the second class ("working class"), who earns income entirely from their labour supply, is $1 - \lambda$. Hereafter, the subscripts r and w represent a rentier and a worker, respectively.

[21]The description of optimal behaviour and equilibrium, and solution methods used in this model is presented in Appendix B.1 and B.2.

Rentier class: Each rentier i, born in time $t = k$, seeks to maximise

$$U_{r,i} = E_k \sum_{t=k}^{\infty} \beta_r^{\,t-k} \frac{c_{r,i,t}^{1-\gamma_r} - 1}{1-\gamma_r}, \qquad c_{r,i,t} \geq 0, \;\; \beta_r \in (0,1), \;\; \gamma_r > 0 \tag{2.8}$$

where β_r is the subjective discount factor, $c_{r,i}$ is consumption and γ_r is the risk aversion coefficient. Each period, a rentier i faces the budget constraint:[22]

$$c_{r,i,t} + a_{i,t+1} = (1-\tau_r)\{r_t^k(k_t, l_t, z_t) + (1-\lambda)\theta_{i,t}\}a_{i,t} + w_t(k_t, l_t, z_t)l_{r,i,t} + (1-\delta)a_{i,t} \tag{2.9}$$

where a_i is total assets, r^k is the rental price of capital, w is the wage rate, θ_i is rentier premium, λ is the population share of rentiers, $\delta \in (0,1)$ is the depreciation rate, and τ_r is the tax rate on capital income (including rentier premium).

A rentier i faces partially insurable labour income risk and rentier premium risk at the same time. First, each rentier is endowed with one unit of time (i.e. $\bar{l} = 1$). This endowment is transformed into labour input according to $l_{r,i,t} = e_{r,i,t}\bar{l}$ where $e_{r,i}$ represents an idiosyncratic labour productivity shock satisfying

$$e_{r,i,t+1} = (1-\rho^e)\mu^e + \rho^e e_{r,i,t} + \zeta_{r,i,t+1}^e, \qquad \zeta_{r,i}^e \sim \mathcal{N}(0, \sigma_r^{e2}) \tag{2.10}$$

where the steady state $\mu^e > 0$ and the adjustment coefficient $\rho^e \in (0,1)$.

Second, θ_i stands for a rentier premium shock which follows an autoregressive process with the steady state $\mu^{rp} \geq 0$, and the adjustment coefficient $\rho^{rp} \in (0,1)$ such that

$$\theta_{i,t+1} = (1-\rho^{rp})\mu^{rp} + \rho^{rp}\theta_{i,t} + \zeta_{i,t+1}^{rp}, \qquad \zeta_i^{rp} \sim \mathcal{N}(0, \sigma^{rp2}). \tag{2.11}$$

Working class: Workers consume their entire income since they have no assets, as adopted by class-based economic theories (e.g. Judd, 1985; Kaldor, 1955). Lifetime utility for a worker j, born in time $t = k$, is given by:

$$U_{w,j} = E_k\Big[\sum_{t=k}^{\infty} (\beta_w)^{t-k} \frac{c_{w,j,t}^{1-\gamma_w} - 1}{1-\gamma_w}\Big], \qquad \beta_w \in (0,1), \; c_{w,j,t} \geq 0, \; \gamma_w > \gamma_r \tag{2.12}$$

[22] Given the premise that an agent i's total assets (a_i) consist of capital (k_i) and wealth residual (nk_i), the rate of return on assets is calculated as

$$\underbrace{r_t^k k_{i,t}}_{\text{capital}} + \underbrace{(r_t^k + \theta_{i,t})nk_{i,t}}_{\text{wealth residual}} = r_t^k \lambda a_{i,t} + (r_t^k + \theta_{i,t})(1-\lambda)a_{i,t} = \underbrace{\{r_t^k + (1-\lambda)\theta_{i,t}\}a_{i,t}}_{\text{total assets}}$$

where λ is the portion of capital, which, for simplicity, is the same for all rentiers.

where β_w is the subjective discount factor, $c_{w,j}$ is consumption, and γ_w is the risk aversion coefficient. Each period, a worker j faces the budget constraints such that

$$c_{w,j,t} + R_{j,t} = w_t(k_t, l_t, z_t) l_{w,j,t} + Tr_{j,t} \tag{2.13}$$

where R_j is the lump-sum rent payment to rentiers (e.g. housing rent), and Tr_j is the lump-sum transfer from the government (e.g. unemployment benefit).

A worker j only faces partially insurable labour income risk. In addition, unlike that of a rentier, the labour productivity of a worker responds to an aggregate TFP shock that creates the cyclical behaviour of an agent's labour productivity. Since labour supply is the only source of income for a worker, a labour productivity shock can create much more variation in consumption for a worker than a rentier. Accordingly, the working class may be hit harder in recessions than the rentier class. Each worker j is endowed with one unit of time; this endowment is transformed into labour input according to $l_{w,j,t} = e_{w,j,t}\bar{l}$, where $e_{w,j}$ represents an idiosyncratic labour productivity shock satisfying

$$e_{w,j,t+1} = (1-\rho^e)\mu^e + \rho^e e_{w,j,t} + \rho^{ez}(z_t - \mu^z) + \zeta^e_{w,j,t+1}, \qquad \zeta^e_{w,i} \sim \mathcal{N}(0, \sigma^{e\,2}_w) \tag{2.14}$$

where the steady state $\mu^e > 0$, the adjustment coefficient $\rho^e \in (0,1)$, ρ^{ez} represents the sensitivity of labour productivity to an aggregate TFP, and z is an aggregate TFP shock with its steady state μ^z.

Firms: Firms operate in a perfectly competitive market. Aggregate output is produced according to a Cobb-Douglas production technology such that[23]

$$y_t = z_t k_t^{\alpha} l_t^{1-\alpha}, \qquad \alpha \in (0,1) \tag{2.15}$$

where k is the aggregate capital stock, l the aggregate labour input, α the share of capital income, and z an aggregate TFP shock that follows an autoregressive process with the steady

[23] Piketty (2014) claims that the elasticity of labour and capital lies between 1.3 and 1.6 on the basis of historical data. He thus concludes that the upsurge of wealth-to-income ratios is not accompanied by decreasing returns on capital, yielding the over-accumulation of capital and growing wealth inequality. In contrast, many authors have estimated that the elasticity of labour and capital lies beween 0.4 and 0.8 (e.g. Chirinko, 2008; Chirinko and Mallick, 2017; Klump et al., 2007; Raurich et al., 2012; Rowthorn, 1999, 2014). Meanwhile, using the US data, Balistreri et al. (2003) find a wide range of long-run elasticities of substitution between labour and capital at the industry level, with a median of around one. To allow for an increasing capital share over time, as argued in Piketty (2014), we need to conduct an experiment using a Constant Elasticity of Substitution (CES) production function with a somewhat higher than unitary elasticity between capital and labour, which I leave for future research.

state $\mu^z > 0$ and the adjustment coefficient $\rho^z \in (0,1)$ such that

$$z_{t+1} = (1-\rho^z)\mu^z + \rho^z z_t + \zeta^z_{t+1}, \qquad \zeta^z \sim \mathcal{N}(0, \sigma^{z2}). \tag{2.16}$$

Government: The government can implement two types of tax policies, favouring the rentier class or the working class. On one hand, a labour-friendly government can impose a distortionary tax on capital income and redistribute them among the working class through a lump-sum transfer. On the other hand, a rentier-friendly government can give a distortionary subsidy on capital income to increase capital investment and boost the economy by imposing a lump-sum tax on workers' labour income. The government may run balanced budgets under the condition:

$$\tau_r \int_{i\in R} (r^k_t + (1-\lambda)\theta_{i,t}) a_{i,t} di = \int_{j\in W} Tr_{j,t} dj, \qquad \begin{cases} \tau_r \geq 0, & \textit{labour-friendly} \\ \tau_r < 0, & \textit{rentier-friendly} \end{cases} \tag{2.17}$$

where R and W represent the rentier class and the working class, respectively, τ_r is the tax rate (or subsidy rate) on capital income, and Tr_j is a lump-sum transfer between the government and a worker.

Assets market: If an agent belongs to the rentier class, she may choose to invest her resources in two different investment technologies, as in Lusardi et al. (2017).[24] The first is a basic technique (e.g. a checking account) that yields a certain (but potentially low) risk-free return (r) that is flowed to capital, representing the expected return for everyone without any financial know-how. The second is a more sophisticated but risky technique that enables a rentier to receive a higher expected return ($r + \theta_i$) from wealth residual.[25] As a result, rentier premium risk (θ_i) may play a fundamental role in generating wealth concentration that far exceeds the earnings concentration because the rate of return on wealth accumulates multiplicatively over time, as similarly argued by Nirei and Souma (2007), Aoki and Nirei (2016), Aoki and Nirei (2017), Benhabib et al. (2017).[26]

[24] Lusardi et al. (2017) show that financial knowledge is a key determinant of wealth inequality in a stochastic life-cycle model with endogenous financial knowledge accumulation, where financial knowledge enables individuals to better allocate lifetime resources in a world of uncertainty and imperfect insurance.

[25] Admittedly, in reality, the market for wealth residual would be endogenously determined. However, for my purpose, this paper analyses the effect of an exogenous change in the price of wealth residual. Thus, internalising the market of wealth residual requires moving away from the standard heterogeneous agents DSGE model, which I leave for future research.

[26] Benhabib et al. (2017) demonstrate why models of earnings inequality and precautionary savings alone cannot reproduce the statistical properties of the wealth distribution that are observed in the data. Consider a

Especially, as in Assumption 2, if certain wealth residual lie in positional competition across households, a worker j should pay the rent to a rentier i and the aggregate returns are zero so that

$$\underbrace{(1-\tau_r)\int_{i\in R}\theta_{i,t}\omega(1-\lambda)a_{i,t}di}_{\text{total (after-tax) rents received by rentiers}} = \underbrace{\int_{j\in W}R_{j,t}dj}_{\text{total rents paid by workers}} \tag{2.18}$$

where ω is a portion of positional competition and R is the rent payment to rentiers.

In addition, asset markets are incomplete. To ensure satisfaction of intertemporal budget constraints, a rentier i's asset holdings are restricted by a borrowing limit b, ensuring the repayment of loans and the absence of Ponzi schemes. To impose this restriction, define the penalty function, as described in Preston and Roca (2007)

$$P(a_{i,t+1}) = \frac{\phi}{(a_{i,t+1}+b)^2}, \qquad b \geq 0 \tag{2.19}$$

where ϕ is barrier parameter and b is the borrowing limit. As an agent's asset holdings approach the borrowing constraint, the penalty function approaches infinity. That is, the inclusion of the penalty function restricts individual debt holding by punishing households in terms of utility for holding too few assets.

---

linear individual wealth accumulation equation

$$a_{t+1} = (1+r_t)a_t + y_t - c_t$$

where a, y, c and r are wealth, earnings, consumption, and the rate of return on wealth, respectively. We may assume $\{y_t, r_t\}$ are stationary stochastic processes. Consider a linear consumption function for simplicity

$$c_t = \Phi a_t + \chi_t.$$

where Φ is constant. We can then write the wealth accumulation equation as

$$a_{t+1} = (1+r_t-\Phi)a_t + y_t - \chi_t.$$

Since the distribution of wealth has a thicker tail than that of earnings, earnings differentials cannot by themselves explain the skewed wealth distribution. So, the burden for explaining the thick tails of the wealth distribution relies upon the rate of return on wealth r_t.

2.5 Results

While the actual data show that wealth is concentrated in the hands of a small number of the wealthy, chiefly due to their high saving rates (Saez and Zucman, 2016), many theoretical models fail to incorporate this fact (Benhabib et al., 2017; De Nardi and Fella, 2017). Even though some models that succeed in matching this fact give completely different policy implications according to their assumptions about the natures of saving motives (e.g. heterogeneity in patience, bequests and transmission of human capital, entrepreneurship and compensation for high earnings risk).[27]

This study suggests a new mechanism in generating wealth inequality based on the realistic saving motive for the wealthy in line with the recent US data, as we have seen in Section 3. More specifically, the purpose of our quantitative analysis is to numerically assess the impact of the centrifugal and centripetal forces in labour income, capital income and rentier premium on wealth distribution. These quantitative exercises require us to firstly calibrate the theoretical model. Once parameters are calibrated and the steady state is approximated, this study calculates through simulation the following outcomes of interest: the Lorenz curves, the associated Gini coefficients, and wealth distribution. In addition, this study examines the effect of political regime changes between labour-friendly and rentier-friendly on wealth distribution and welfare gains.

2.5.1 Calibration

The calibration follows the standard heterogeneous agents literature. The values of structure parameters are chosen to be consistent with aggregate features of the US economy after the early 1980s, while those of the stochastic process parameters are calibrated to fit the data using the Bayesian estimation. Table B.1 reports the values of all parameters and contains a comment on how they were selected in the model.

Each time period is taken to be one quarter. The subjective discount factor of the rentier class, β_r, is chosen to be 0.985, such that the average rate of real interest is 4.67% at an annual frequency, while that of the working class, β_w, is set to 0.960. The relative risk aversion of the rentier class, γ_r, is set to 2, while that of the working class, γ_w, is chosen to

[27] For instance, in a model with entrepreneurship, the adverse responses of savings to increased taxation are significant because higher taxation negatively affects the returns on running a business (Kitao, 2008; Lee and Lee, 2018). In contrast, in a model with high earnings risk for the top earners, the optimal tax rate is close to 90% (Kindermann and Krueger, 2014).

be 5, based on micro evidence in Mehra and Prescott (1985). The share of capital income, α, is taken to be 0.33, which matches the fact that the average share of labour income is roughly around two-thirds, according to the BLS. The depreciation rate of capital, δ, is set at 0.022 to reproduce the average share of capital investment (i.e. private fixed investment plus purchases of new durable goods): 22% at a quarterly frequency. The share of the working class who exhibits "rule-of-thumb behaviour" in this economy, $1-\lambda$, is set at 0.25 based on the DNAs.[28] Our analysis assumes that households are constrained to hold positive quantities of the capital stock, so that the borrowing limit, b, is set at zero. The barrier parameter, ϕ, is set to be 0.05, as in Preston and Roca (2007), to control the sensitivity to the borrowing constraint in the utility function, ensuring that no agent violates this constraint. The portion of positional competition, ω, is set to be 0.33 to match the wealth residual data (i.e. average ratio of land to wealth residual) in Table 2.1. Finally, the tax rate on capital income, τ, is set to zero in the baseline model.

The stochastic processes for aggregate total factor productivity (TFP), individual labour productivity, and rentier premium are estimated by Bayesian estimation, described in Table B.2 and Figure B.1.[29] First, the aggregate TFP shock is specified by μ^z=1, ρ^z=0.75 and σ^z=0.013 to correspond to the two-state Markov process, close to the one adopted by Krusell and Smith (1998) and Preston and Roca (2007). Second, an individual labour productivity is specified as both μ^e=1 and ρ^e=0.7, which are the same for two classes; however, the higher volatility of workers' labour productivity is given by a larger standard deviation (σ^e_w=0.2 compared to σ^e_r=0.1) and by the interaction term between labour productivity and business cycle (ρ^{ez}_w=0.45 compared to ρ^{ez}_r=0), as in Preston and Roca (2007). Lastly, rentier premium is specified by μ^{rp}=0, ρ^{rp}=0.9 and σ^{rp}=0.1, implying that the gross return on wealth residual lies in the interval [0.9874, 1.1437] for 90% of all agents in the initial steady state at a quarterly frequency with a high persistency, as in Hubmer et al. (2016).

2.5.2 Simulation

The computation of wealth distribution is based on simulating the model economy using policy functions and three shocks that affect an agent's earnings (i.e. labour income, capital income and rentier premium). For simulation purposes, I use 10,000 households and set

[28] Many authors have estimated the portion of income allocated to "rule-of-thumb" households (e.g. Blinder and Deaton, 1985; Hall and Mishkin, 1982; Weber, 2000) but the values vary quite a lot from 0.082 to 0.483.

[29] I use the matlab code of Petr Sedlacek, which was acquired in 2018 summer course of "Tools for Macroeconomist" at the London School of Economics.

their initial wealth to be completely uniform within their classes. More specifically, 7,500 households start with the steady-state assets, while the remaining households begin with zero assets. I then simulate a series of three shocks using a random number generator from a normal distribution and obtain 1,000 draws.[30]

After the model is simulated, I use the outcomes as a measurement tool for the following question: how far can we go in explaining wealth inequality when different histories of labour earnings risk are the only source of heterogeneity among households (denoted "Benchmark"), as in Aiyagari (1994)? The typical answer is that this Aiyagari-type model generates too much asset holdings at the bottom and too little at the top so that the wealth Gini is around 0.27, which is much smaller than the actual value (0.86) in 2014. Therefore, this model needs to be modified by either introducing an extra incentive for the wealthy to accumulate wealth or reducing an incentive for the poor to save for self-insurance purposes (or both). In addition, to generate an extremely skewed right tail of wealth distribution, it seems necessary to make the wealthy have higher propensities to save, give them higher returns on saving than the Aiyagari-type model, or even combine both solutions.

In this respect, this study presents a new mechanism in generating wealth inequality by simulating based on the following three models: (i) innate difference in asset ownership ("Model 1"), (ii) heterogeneity in rentier premium ("Model 2") to generate heterogeneous returns on wealth, and (iii) heterogeneity in rentier premium with a fat-tailed distribution of asset market returns ("Model 3") to reflect the non-normality of asset market returns in reality (as already seen in Table 2.4). Especially, the simulation with non-normality of asset market returns is constructed to create a realisation of very high earnings, which occurs with a very low probability in the so-called "awesome state", as in Castaneda et al. (2003) and Kindermann and Krueger (2014)[31], to be consistent with the wealthy class's extraordinarily high propensity to save.[32]

Figure 2.2 plots the evolution of the wealth Gini and the Lorenz curve for wealth. It reports that my model succeeds in generating empirically plausible degrees of wealth inequality and shows two implications. First, if we exclude heterogeneity in rentier premium, as in the Benchmark model and Model 1, the wealth Gini clearly depicts a very slow

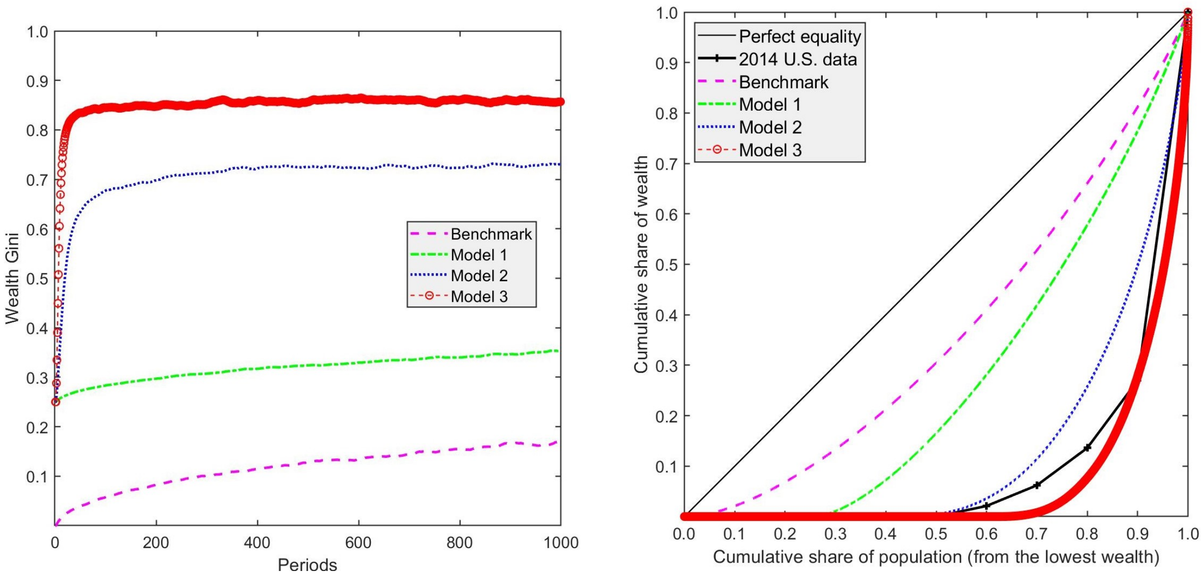
[30] I treat agents whose wealth reach zero because of many unfavourable shocks as "economically dead" so that they no longer participate in wealth accumulation. Mathematically, it corresponds to imposing an absorbing boundary condition on the density of agents who have zero wealth.

[31] Castaneda et al. (2003) calibrate the properties of an awesome state in an overlapping generation model such that the top 0.039% earners have roughly 1,000 times the average labour endowment of the bottom 61%. Similarly, Kindermann and Krueger (2014) estimate that 0.036% of the population's labour productivity lies in the awesome state.

[32] To generate the awesome state, I simulate a series of rentier premium shocks using a random number generator from a t location-scale distribution, fit to the actual fat-tailed distribution of asset market returns of the US during the 1984-2013 period.

convergence to the new steady state. However, once we allow the stochastic process of rentier premium, as in Model 2 and Model 3, it generates a very fast convergence to the new steady state of the wealth Gini. This is consistent with the work of Gabaix et al. (2016), who argues that persistent heterogeneity in returns (the so-called “type dependence”) is necessary for models to account not only for extreme wealth concentration but also for the speed of change of wealth concentration observed in the data. Second, after the convergence, the degree of wealth inequality becomes worse, such that it is close to the actual distribution, as in Model 2 and Model 3. Nevertheless, if we assume only the rentier premium risk, as in Model 2, it predicts somewhat too a low concentration in the extreme upper tail of the wealth distribution and too a high concentration in the middle, compared to the actual data. These shortcomings reflect the restrictions of the symmetric normal distribution that has been imposed upon the rentier premium shock in Model 2. In this respect, Model 3 that reflects the fat-tailed distribution of asset market returns best mimics the actual wealth distribution.

Fig. 2.2 Gini coefficients and Lorenz curves



(a) Evolution of wealth Gini (b) Lorenz curves for wealth

Notes: US data – the World Inequality Database; Benchmark – heterogeneity in labour productivity; Model 1 – plus innate difference in asset ownership; Model 2 – plus heterogeneity in rentier premium; Model 3 – plus non-normality of asset market returns.

The resulting stationary wealth distribution is described in Table 2.6. The first line refers to the 2014 US data from the WID.world. It shows that wealth is highly unevenly distributed: the wealth Gini is 0.86 and the wealthiest 10% of households hold 73 percent of the net worth, while the poorest 50% hold almost zero. The second line of data refers to the outcomes from Benchmark model that replicates Aiyagari (1994) with an aggregate TFP shock, while the third line of data refers to the outcomes from Model 1, which assumes the inborn difference in asset ownership under the Aiyagari-type economy. Comparing these lines makes it clear that the innate difference in asset ownership itself generates higher wealth inequality (0.45 in Model 1 compared to 0.27 in Benchmark), but it is still far below the actual value (0.86).

The fourth line of data refers to the outcomes from Model 2, which adds heterogeneity in rentier premium to Model 1, while the last line of data refers to the outcomes from Model 3, which introduces the non-normality of asset market returns into Model 2. Comparing these lines makes it obvious that Model 3 comes closest to matching both the concentration of wealth in the hands of the wealthy few and the main features of wealth distribution in the data, including the wealth Gini. For instance, in Model 3, the wealthiest 10% of households hold 73 percent of net worth, while the poorest 50% hold almost zero, and the wealth Gini is 0.86. My model also replicates the fact that almost a quarter of households hold non-positive wealth, close to 20-30 percent in the data, unlike the zero percent result found in the Aiyagari-type model.

Table 2.6 Wealth distribution: models and data

(%)

percentage of net worth held by			wealth	share of	fraction of
top 10%	middle 40%	bottom 50%	Gini	top 5%	wealth ≤ 0
2014 US data– WID.world					
73	27	0	0.86	60	20-30
Benchmark– heterogeneity in labour productivity ($\sigma^{rp} = 0$, $\sigma^e > 0$, $\lambda \simeq 1$)					
19	51	30	0.27	10	0
Model 1– plus innate difference in asset ownership ($\sigma^{rp} = 0$, $\sigma^e > 0$, $\lambda = 0.75$)					
24	59	17	0.45	13	25
Model 2– plus heterogeneity in rentier premium ($\sigma^{rp} > 0$, $\sigma^e > 0$, $\lambda = 0.75$)					
50	50	0	0.73	32	25
Model 3– plus non-normality of asset market returns ($\sigma^{rp} \gg 0$, $\sigma^e \gg 0$, $\lambda = 0.75$)					
73	27	0	0.86	51	25

Wealth concentration continues to increase, mainly due to the heterogeneous returns on wealth, as witnessed by the recent theoretical literature (e.g. Benhabib et al., 2017) and empirical findings (e.g. Cao and Luo, 2017; Saez and Zucman, 2016). As in Table B.3, the simulated data can mimic the actual returns on wealth, which are recompiled by Cao and Luo (2017).

The Aiyagari-type model assumes that all individuals are confronted with a common risk-free rate of return on savings, but it doesn't fit well with the observed data, as seen in the Benchmark model. Instead, my theoretical model and the observed data consistently support the conjecture that there is substantial heterogeneity in returns on wealth across households; in Model 3, the standard deviation of annualised returns is approximately 20%. While there are masses of households who earn low returns close to zero, a small fraction of households enjoy an annualised return higher than 30%. In addition, there is a significant fraction of households who report capital losses. Thus, we can infer that one of the main losers in rising wealth concentration at the top was the American lower-middle class - households between the bottom 25th and 50th of the wealth distribution who have lost their most assets as a result of many unfavourable shocks in asset trading, supporting the empirical findings of Kuhn et al. (2020).[33] In addition, since the households at the bottom 50% should pay the extra cost from a higher rentier premium, it directly reduces the incentives for the poor to save for self-insurance purposes while giving an extra incentive for the wealthy to accumulate assets.

Table 2.7 describes another important related observation that concentration of wealth is much higher than that of income. The first line refers to the 2014 US data from the WID.world. It shows that the Gini coefficient for market income (pre-tax, pre-transfer) is 0.60 and the top 10% of households earn 47 percent of gross national income, while the bottom 50% earn only 13 percent. The second line of data refers to the outcomes from Benchmark model, while the third line of data refers to the outcomes from Model 1. Comparing these lines makes it clear that introducing the inborn difference in asset ownership transfers income from the bottom 50% to the top 10% and the middle 40%. The income share of the bottom 50% decreases from 43 percent to 38 percent, while income share of both the top 10% and the middle 40% increase by 2 and 3 percentage points, respectively.

The fourth line of data refers to the outcomes from Model 2, while the last line of data refers to the outcomes from Model 3. Comparing these lines makes it obvious that Model 3 succeeds in replicating the concentration of income in the hands of the rich few and the high level of market income Gini. For instance, in Model 3, the top 10% of households earn 51 percent of gross national income, while the bottom 50% earn 14 percent, and the income Gini is 0.61 that is very close to the observed data value (0.60).

However, some authors (e.g. Atkinson, 2015) argue that the Gini index is over-sensitive to changes in the middle of the distribution and, as a consequence, insensitive to changes at the top and the bottom. Therefore, I explore a different measure of income inequality: the Palma ratio. The Palma is the ratio of the richest 10%'s share of gross national income

[33] Due to some limitations of this arguably simplistic framework, the model fit may not be perfect. This would require moving away from a simple heterogeneous agents economy which I leave for future research.

divided by the poorest 40%'s share.[34] Our estimated Palma ratio in Model 3 is 5.6, which implies that the top 10% earns 5.6 times more than the bottom 40%. This index reflects the fact that income of the middle class almost always represent about half of gross national income, while the other half is split between the richest 10% and the poorest 40%. In this regard, my assumption on rentier premium is closely related to the Palma ratio such that the income share of the rich grows at the expense of the poor through rentier premium. So, as argued by Palma (2011), distributional politics would be largely about the battle between the rich and the poor for the half of national income, and who the middle classes side with in that battle.

Table 2.7 Income distribution: models and data

(%)

percentage of income held by			market	share of	Palma
top 10%	middle 40%	bottom 50%	income Gini	top 5%	ratio
2014 US data					
47	40	13	0.60	36	6.1
Benchmark- heterogeneity in labour productivity ($\sigma_r = 0$, $\sigma^e > 0$, $\lambda \simeq 1$)					
13	44	43	0.10	7	0.4
Model 1- plus innate difference in asset ownership ($\sigma_r = 0$, $\sigma^e > 0$, $\lambda = 0.75$)					
15	47	38	0.16	8	0.5
Model 2- plus heterogeneity in rentier premium ($\sigma_r > 0$, $\sigma^e > 0$, $\lambda = 0.75$)					
41	39	20	0.49	28	2.7
Model 3- plus skewed distribution of earnings ($\sigma_r \gg 0$, $\sigma^e \gg 0$, $\lambda = 0.75$)					
51	35	14	0.61	36	5.6

Notes: Palma ratio, the ratio of the richest 10%'s share of gross national income divided by the poorest 40%'s share; US data- the Standardised World Income Inequality Database and World Inequality Database.

Finally, Table 2.8 describes the related observation that the share of labour income declines, as I introduce heterogeneous returns on wealth into the model. The share of labour income was thought to be relatively stable at around two-thirds in many theoretical and empirical studies. According to the BLS, however, the share of labour income has declined by roughly 9.4 percentage points, dropping from 66.4% in 1960 to 57.0% in 2014. In Model 1, an innate difference in asset ownership itself cannot generate the sharp declining share of labour income compared to the Aiyagari-type model; the estimated share is 67.0%. However, once we incorporate heterogeneous returns on wealth, the share of labour income significantly

[34]For instance, if the richest 10% earn half of the national income and the poorest 40% earn one-tenth of the national income, the Palma ratio is 5 (i.e. 0.5 divided by 0.1).

fell to 58.5%. Furthermore, introducing the non-normality of asset market returns brings about an additional drop in the share of labour income by 1.9 percentage points. Therefore, we can infer that the sharp decline in the share of labour income after the 1980s may be closely related to fast-growing capital income from wealth residual relative to labour income.

Table 2.8 Share of labour income: models and data

Benchmark	Model 1	Model 2	Model 3	US data (1960 $\to$ 2014)
67.0%	67.0%	58.5%	56.6%	66.4% $\longrightarrow$ 57.0%

Notes: Benchmark – heterogeneity in labour productivity; Model 1 – plus innate difference in asset ownership; Model 2 – plus heterogeneity in rentier premium; Model 3 – plus non-normality of asset market returns; US data – the BLS.

2.5.3 Policy implications

Put simply, there are two types of governments: labour-friendly and rentier-friendly. For instance, the labour-friendly government collects distortionary capital income taxes and redistributes them among workers in the form of lump-sum transfers. In contrast, the rentier-friendly government collects lump-sum labour income taxes and redistributes them among rentiers in the form of distortionary capital income subsidies. Normally, since the type of the government is determined by elections, we can regard changes in political regime as an exogenous process. But the economic consequences of the political regime changes are substantially different (see Mukand and Rodrik (2020) for further discussion). Thus, by investigating the effects of political regime changes between labour-friendly and rentier-friendly governments on wealth distribution and welfare gains, I quantitatively evaluate and predict the distributional impacts of tax policies and regulations. In this respect, I conduct various experiments by changing institutional parameters related to the government's intervention and comparing the results with the calibrated baseline model.

Taxation. The progressivity of the US tax system has substantially decreased since the 1980s. For instance, the maximum tax rate on ordinary income (including short-term capital gains) significantly reduced from 90% in 1963 to 28% in 1988. To investigate how the changes in the tax rate on capital income affect wealth distribution and welfare gains, I perform several counterfactual exercises. Results are reported in Table 2.9, which quantitatively suggests that the welfare implications are large and the effects are asymmetric across the wealth distribution.

In experiment 1, I reduce the tax rate on capital income by 20 percentage points, which matches the Reagan tax cuts during the early 1980s. The welfare gain of all households is roughly 37.3 percent of the permanent consumption equivalent of the baseline economy. However, when it comes to the welfare implications per wealth decile, the welfare gain of households at the bottom 50% is 7.9 percent, while the welfare gain of households at the top 10% and the middle 40% are 53.2 percent and 32.2 percent, respectively. This is because subsidies on capital income naturally increase the tax burden of workers and the after-tax returns on assets of rentiers at the same time while the tax cut boosts investment. The changes in wealth Gini and consumption Gini, compared to the baseline economy, rise by 0.01 and 0.05 points, respectively, implying that households at the bottom 50% would be relatively worse off in the aftermath of tax cuts for the wealthy.

In experiments 2-5, I then increase the tax rate from -10% to 40%. The outcomes give the following two implications. First, the government is confronted with a 'welfare-equality trade-off' problem. More specifically, as the tax rate increases, inequality falls while the welfare gain falls; the changes in wealth Gini and consumption Gini, compared to the baseline economy, fall from 0.00 and 0.03 points to -0.03 and -0.26 points, respectively, whereas the welfare gain decreases from 17.3% to -44.5%. Second, the government also faces a 'class conflict' problem. Concretely, as the tax rate goes up, the welfare losses of households at the top 10% and the middle 40% become more significant, while the welfare gain of households at the bottom 50% become larger. Since a distortionary tax on capital income changes the optimal amount of assets that rentiers wish to hold, it would decrease capital accumulation in the whole economy, which naturally results in a declining gross national product.

In this respect, I conduct another experiment, assuming that the government only imposes a 40% tax rate on capital income from wealth residual (the so-called "Henry George Tax") instead of total assets, as in experiment 6. There remains a 37.9% aggregate welfare loss, but it is lower than the 44.5% welfare loss found in experiment 5. Given the better equality in wealth and consumption, we can infer that a well-designed progressive taxation would be less distortionary and allows us to divert resources from non-productive activities to productive ones that can recover a part of welfare losses.

The main implication of these experiments is that my model can differ from the representative agent model with complete markets on some important policy issues. For instance, with complete markets, the capital income tax should be zero in the long run (e.g. Chamley, 1986; Judd, 1985). However, as Aiyagari (1994) mentions, with idiosyncratic shocks and incomplete markets, the optimal capital income tax is strictly positive even in the long run.[35] Therefore, the large welfare gains of reducing the capital income tax to zero calculated by Lucas (1990) in the complete markets model may well generate welfare losses in the incomplete markets model.[36] My analysis further suggests that if the economy has a property of positional externality in rentier premium, the government needs to impose a positive (and high) tax rate on rentier premium to attain the distributional outcomes of the Aiyagari-type model. That is, since a disproportionate share of the wealth at the top is associated with rents from wealth residual, the Henry George Tax would reduce inequality efficiently. Then, in this economy, wealth inequality might only be created from labour productivity differentials, which most people may convince.

Table 2.9 Policy experiments: tax or subsidy on capital income

	average welfare gain (%) held by top				changes in Gini	
	0-10%	10-50%	50-100%	all	wealth	consumption
Baseline (τ=0%)	-	-	-	-	-	-
Experiment 1 (τ=-20%)	53.2	32.2	7.9	37.3	0.01	0.05
Experiment 2 (τ=-10%)	24.5	15.3	2.8	17.3	0.00	0.03
Experiment 3 (τ=10%)	-20.7	-13.8	10.5	-13.3	-0.00	-0.05
Experiment 4 (τ=20%)	-28.1	-26.3	23.0	-24.5	-0.01	-0.13
Experiment 5 (τ=40%)	-64.6	-47.6	27.4	-44.5	-0.03	-0.26
Experiment 6 (τ^p=40%)	-57.3	-40.4	29.8	-37.9	-0.03	-0.22

Note: τ^p stands for the tax rate on rentier premium only.

Regulations. Now I change three key regulatory parameters that influence an agent's wealth accumulation: (i) the population share of asset owners λ, (ii) the intensity of positional competition ω, and (iii) the volatility of rentier premium σ_r. Results are reported in Table 2.10, which quantitatively shows that the most powerful distributional policy is expanding access to asset ownership.

In the first counterfactual exercise, I adjust the population share of asset owners. For instance, the US government promoted homeownership using government sponsored entities (e.g. Freddie Mac and Fannie Mae) which fund borrowing by home buyers or guarantee

[35]The recent work of Straub and Werning (2020) also questions the Chamley-Judd result that capital should not be taxed in the long run, proving that the long-run tax on capital is positive and significant, whenever the intertemporal elasticity of substitution is below one.

[36]The recent work of Acikgoz et al. (2018), using incomplete markets models of the Bewley-Huggett-Aiyagari type, quantitatively shows that in the long run the government debt-to-GDP ratio is high, capital is taxed at a low rate and labour income at a high rate when compared to current US values.

mortgage-backed securities. Notably, the Community Reinvestment Act of 1977 expanded homeownership for low- and middle-income earners by encouraging depository institutions to help meet the credit needs of the communities in which they operate. As a result, the US homeownership rate increased from 62.1% in 1960 to 69.0% in 2004. This rate fell after the 2004 peak down to 63.7% in 2015.

Based on this background, in experiment 1, I increase the population share of asset owners λ, from 75% to almost 100%. The welfare loss of all households is roughly 26.5 percent of the consumption equivalent of the baseline economy. However, when it comes to the welfare implications per wealth decile, welfare falls among the top 10% and the middle 40%; on the contrary, a huge welfare gain is given to the lowest tail of wealth distribution. This is explained by two factors. First, boosting asset ownership enables the bottom 50% to accumulate capital and thus to gain capital income from their asset ownership. As a result, wealth Gini and consumption Gini fall quite a lot, by 0.46 and 0.43 points, respectively. Second, the rentier class cannot enjoy rentier premium any more; thus, their permanent income naturally decreases. In contrast, when I reduce the population share of asset owners by up to 60%, as in experiment 2, the results are totally reversed. A further monopoly of access to wealth residual results in a huge welfare gain to the top 10% at the expense of increasing inequality in wealth and consumption for the whole economy.

In the second counterfactual exercise, I control the intensity of positional competition (e.g. land use regulation) such that parameter ω changes from 0 to 1, as in experiment 3 and 4. The welfare gain of all households decreases from 11.4 percent to -0.3 percent. Interestingly, the welfare loss is concentrated in the lowest tail of wealth distribution. This is because as positional competition intensifies, asset owners would like to gain high rentier premium from positional goods (e.g. land), resulting in the total exploitation of the working class without any increases in total output like a zero-sum game.

In the last counterfactual exercise, I restrict the range of the returns on assets (e.g. rate-of-return regulations) such that a standard deviation of rentier premium $|\sigma_r|$, reduces from infinity to 5 percent, as in experiment 5. Both wealth Gini and consumption Gini decrease by 0.05 and 0.21 points, whereas the welfare loss of all households is 19.7 percent of the consumption equivalent of the baseline economy. Especially, the welfare loss of households at the top 10% is 44.9 percent, while the bottom 50% shows 63.3 percent of the welfare gain. This implies that with the direct control of returns, those who do not own assets can rely on the government to ensure that asset owners must earn similar capital income that would ideally prevail in a perfectly-competitive market, equal to a market-determined rate of return on capital. However, in experiment 6, as I relax the range of σ_r, the impact of rate-of-return regulation on inequality in wealth and consumption diminishes.

Table 2.10 Policy experiments: regulations

	average welfare gain (%) held by top				changes in Gini	
	0-10%	10-50%	50-100%	all	wealth	consumption
Baseline (λ=0.75, ω=0.33, $\lvert\sigma_r\rvert \leq$ inf)	-	-	-	-	-	-
Adjust the population share of asset owners (λ)						
Experiment 1 ($\lambda \simeq$1)	-73.6	-30.4	129.4	-26.5	-0.46	-0.43
Experiment 2 (λ=0.6)	76.6	22.2	-78.4	30.4	0.06	0.16
Control positional competition (ω)						
Experiment 3 (ω=0)	0.0	0.0	87.0	11.4	-	-0.12
Experiment 4 (ω=1)	0.0	0.0	-2.7	-0.3	-	0.00
Restrict the range of returns on assets ($\lvert\sigma_r\rvert$)						
Experiment 5 ($\lvert\sigma_r\rvert \leq$5%)	-44.9	-21.6	63.3	-19.7	-0.05	-0.21
Experiment 6 ($\lvert\sigma_r\rvert \leq$15%)	-18.6	-8.4	14.3	-9.5	-0.01	-0.05

Political regime changes To investigate the total effect of political regime changes, I carry out an additional exercise by assuming that the government has two distinct policy mixes: (a) a labour-friendly policy ($\tau = 0.1$, $\lambda = 0.85$, $\omega = 0$, $|\sigma_r| \leq 0.10$) that uses more progressive taxation and rather stricter regulations and (b) a rentier-friendly policy ($\tau = -0.1$, $\lambda = 0.65$, $\omega = 1$, $|\sigma_r| \leq$ inf) that conducts more regressive taxation and rather looser regulations than the baseline economy. As reported in Table 2.11, comparing these two results makes it clear that the government, which may be supported by those with rentier income (i.e. the upper half of population), remarkably aggravates inequality and lowers the share of labour income but expands the welfare gain for the whole economy. When it comes to the outcomes per wealth decile, the labour-friendly government definitely gives huge gains in wealth and welfare to households at the bottom 50%, while the rentier-friendly government yields substantial gains in wealth and welfare to households at the top 10% and middle 40%. Therefore, it is not easy for the government to solve 'welfare-equality trade-off' and 'between-class conflict' problems at the same time.

Figure 2.3 displays three-dimensional graphs of the population density. The density is on the z-axis, while the x-axis and y-axis contain disposable income (in logs) and asset holdings (in logs), respectively. The upper graph depicts two 'witch' hats, whereas the lower graph shows three 'folded' witch hats with a 'new towering' witch hat. Comparing these two figures makes it obvious that, quantitatively, political regime changes like those after the 1980s have a detrimental effect on a low income and low wealth position, while only a few

Table 2.11 Impact of political regime changes on the model economy

(a) The aggregate outcomes

	Gini		gain (or loss)		share of	
	wealth	consumption	wealth	welfare	top 10% wealth	labour income
LF gov't	-0.10	-0.27	-40%	-28%	-11%p	+9%p
RF gov't	+0.04	+0.13	+46%	+40%	+6%p	-8%p

(b) The outcomes per weatlh decile

	wealth gain (or loss)			welfare gain (or loss)		
	top 10%	middle 40%	bottom 50%	top 10%	middle 40%	bottom 50%
LF gov't	-53%	-33%	+592%	-56%	-30%	+68%
RF gov't	+63%	+31%	-100%	+82%	+32%	-60%

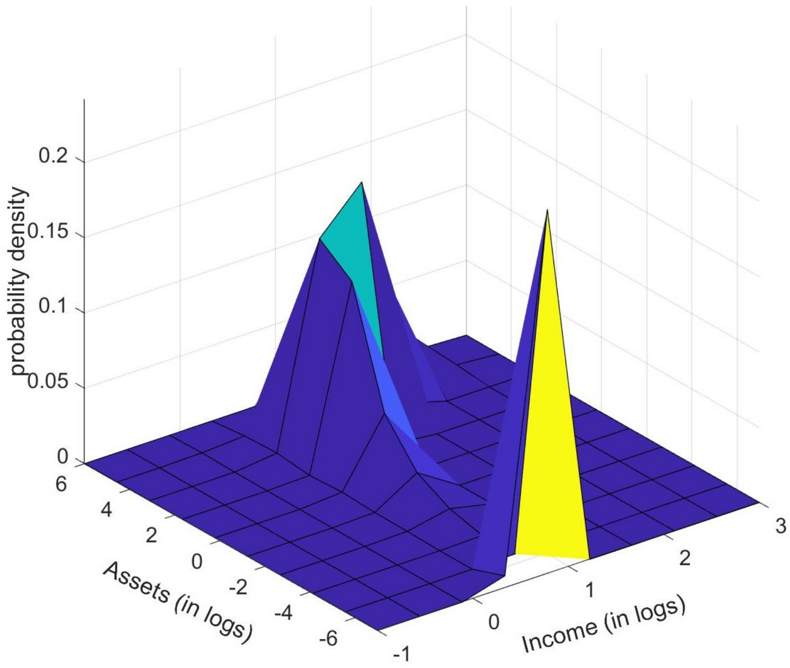
Notes: LF gov't implies the labour-friendly government whose policy mix is $\{\tau = 0.1,\ \lambda = 0.85,\ \omega = 0,\ |\sigma_r| \leq 0.10\}$, and RF gov't implies the rentier-friendly government whose policy mix is $\{\tau = -0.1,\ \lambda = 0.65,\ \omega = 1,\ |\sigma_r| \leq \text{inf}\}$.

wealthy individuals benefit greatly from loose regulations and tax-cuts on capital income. We can also reconfirm that another loser in rising wealth inequality is the lower-middle class who stepped down to the poor as a result of many unfavourable earnings shocks.

2.6 Conclusion

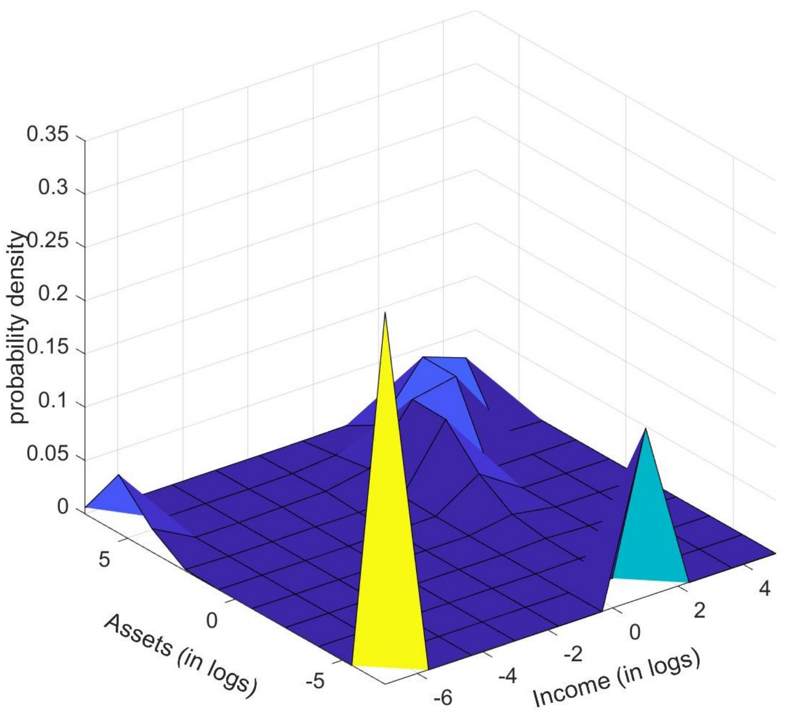
This study tried to answer the two fundamental questions: (i) 'despite the rapid expansion of aggregate wealth, why do some people stay wealthy while others remain poor?' and (ii) 'to what extent can government policies affect wealth inequality?' To address these questions, I focused on finding a new mechanism that complimented the precautionary savings hypothesis to show the wealthy's high marginal propensity to save. Based on the empirical evidence of persistently heterogeneous returns on wealth across households, this study introduced the rentier class, defined as those who have a monopoly of access to wealth residual, into the canonical heterogeneous agents DSGE model. More specifically, those who belong to this class receive rentier premium from those who do not. This study also compared the effects of two different types of government on wealth distribution and welfare gains. I conclude

Fig. 2.3 The impact of political regime changes on the population density



(a) Labour-friendly government

$\{\tau = 0.1,\, \lambda = 0.85,\, \omega = 0,\, |\sigma_r| \leq 0.10\}$



(b) Rentier-friendly government

$\{\tau = -0.1,\, \lambda = 0.65,\, \omega = 1,\, |\sigma_r| \leq \inf\}$

Note: Income stands for post-tax, -transfer, and -rent payment income.

that the rise in rentier premium and the government's policies favouring rentiers have acted as key drivers in generating the co-evolution of rising wealth inequality and declining share of labour income in the US since the 1980s.

In short, by paraphrasing a famous quote from Malthus (1798), we can predict a grim future, as capital gains would increase geometrically, but labour income would only grow arithmetically, which would result in extreme wealth concentration, unless the government enforces 'rebalancing rules' to transfer wealth from the top decile to the bottom decile.

Given that my model shows that the reniter premium plays an important role in exacerbating wealth inequality, an important next step of inquiry is to explain why the gap between the level of inequality in the US and those in other rich countries has become larger. More specifically, we need to empirically assess which institutional factors (e.g. financial system, lobbying intensity, and regressive policies) have aggravated the recent US wealth distribution (Alvaredo et al., 2018). Further research is also needed to show more explicitly how the rentier premium relates to excess returns arising in standard portfolio choice models. For instance, we can decompse the rentier premium into four parts– differences in time-varying observable characteristics (e.g. risk exposure), transitory variations (e.g. luck), persistent components (e.g. ability) or (monopoly and scaricty) rents.

Chapter 3

How Much Higher Can the Minimum Wage Safely Rise?

The unprecedented impeachment of South Korea's former president in mid-2017 and the subsequent large rise in the minimum wage allows us to investigate the non-linear employment effects of increases in the minimum wage on low-paid workers. To demonstrate these effects, this study uses two-step generalized method of moments, difference-in-differences, and regression discontinuity designs based on individual-level panel data over the period of 2009–2018. The estimation results show that a 14.9% rise in the real minimum wage in 2018 seems to hit the tipping point, which is 5.5% in my estimation. This unexpected double-digit growth in the minimum wage has led to a reduction in the number of hours worked by low-wage workers, compared with modest increases over a decade under business-friendly governments. Paradoxically, the large increase in the minimum wage has had an unintended positive consequence of promoting better work-life balance in South Korea – a place that is notorious for over-working. We ask: Can increases in the minimum wage actually help low-income workers? It depends.

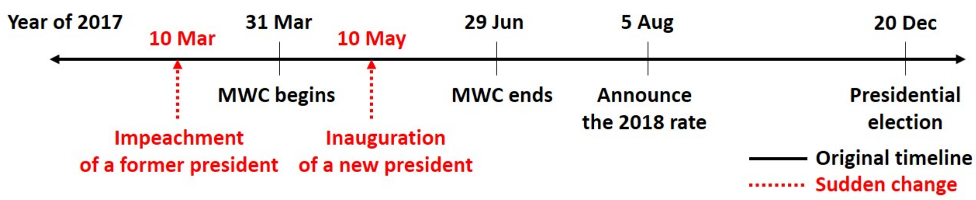
Keywords: Minimum wage; non-linear employment effect; generalized method of moments; difference-in-differences; regression discontinuity designs
JEL classification: J21, J38, J42

3.1 Introduction

The goal of this paper is to investigate how much higher the minimum wage can safely rise without leading to employment losses. Many studies have discussed the relationship between the minimum wage and employment (e.g. Card and Krueger, 1995; Neumark and Wascher, 2008) but there is still considerable disagreement over the sign and strength of the minimum-wage employment effects (e.g. Harasztosi and Lindner, 2019). In particular, not many studies have recognised that the relationship might be non-linear Manning (2016). This is partly because minimum wages have been raised at modest rates in most developed countries.[1]

However, the unexpected impeachment of a former right-wing president in South Korea allows us to study the non-linear impact of minimum wages on low-wage workers. In particular, the leaning of the Minimum Wage Commission (MWC) towards the employers' side was originally supposed to determine the 2018 minimum wage between 31 March and 29 June 2017 (see Figure 3.1). However, the sudden regime change to a labour-friendly government overturned the decade-long business-friendly stance on minimum wages. The new government raised the real minimum wage by 14.9% over the previous year, much higher than the 5.3% that was expected had the impeachment not happened (see Figure 3.2).

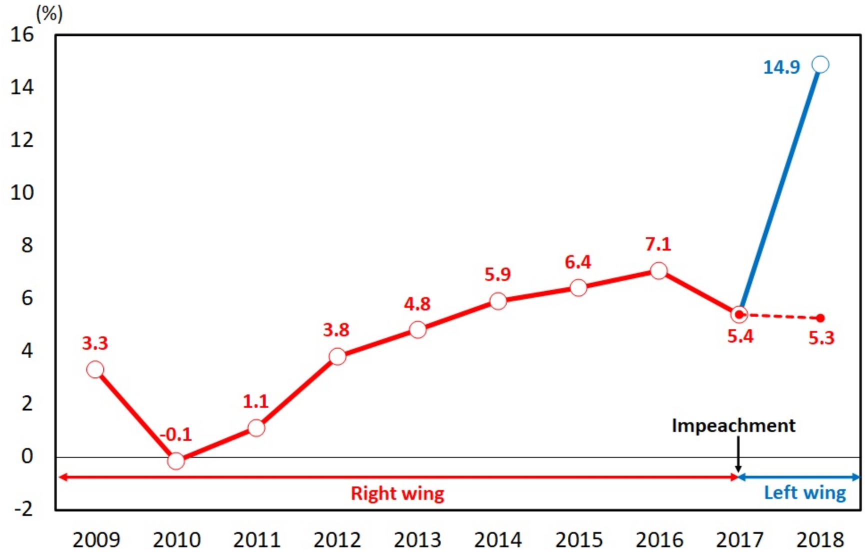
Fig. 3.1 Timeline of the regime change in South Korea



Because of the upsurge in the 2018 minimum wage, it was expected that employers would experience non-trivial increases in labour costs. When labour costs increase, an employer's immediate options are: (i) to pass on (at least part of) the increased cost to the consumers (Aaronson and French, 2007; Lemos, 2008); (ii) to absorb the increased cost by accepting lower profits (Draca et al., 2011; MaCurdy, 2015); (iii) to reduce the number

[1]From 2009 to 2018, on average, the real minimum wage grew by 2.1% across 28 OECD countries that had a statutory minimum wage, but it grew 1.5% if Eastern Europe is excluded (OECD.stat, 2020). Given the relatively small cost to employers of modest increases in the minimum wage, various adjustment mechanisms (e.g. reductions in non-wage benefits or profits) appear to be more than sufficient to avoid employment losses, even for employers with a large share of low-wage workers (Hirsch et al., 2015; Schmitt, 2013).

Fig. 3.2 Growth rates of the real minimum wage in South Korea



Note: The dotted line implies the estimate under the assumption of no impeachment

of people employed (i.e. reducing the cost of labour at the 'extensive margins'); and (iv) to reduce the number of hours worked per worker (i.e. reducing the cost of labour at the 'intensive margins', Brown, 1999; Hamermesh, 1995; Stewart and Swaffield, 2008; Zavodny, 2000).[2]

In the case of South Korea, there was little evidence to support the first three possibilities. The increase in the minimum wage was hardly passed on to consumers through higher prices – the rate of inflation actually fell afterwards (see Figure C.2a). Nor was there evidence that it was paid by employers through lower profits; the rate of profit stayed more or less the same after the hike in the minimum wage (see Figure C.2b). Furthermore, there was negligible change in the employment rate and the unemployment rate for all age groups (see Figure C.2c and C.2d). However, as this paper shows, there is evidence that the immediate adjustment in labour demand against the 2018 minimum wage hike in South Korea took place at the intensive margins.[3]

Another factor to consider in understanding the effects on employment of increases in

[2]In the longer run, an employer can keep her current level of profit at the current level of employment by making the operation more efficient through altering capital investments in machinery, buildings and technology (e.g. Kaas and Madden, 2008).

[3]In response to an increase in the minimum wage, a firm owner can cut back on labour services either by laying off workers or by reducing the length of the scheduled workweek (Michl (2000) called it the 'rescheduling hypothesis'). Since total payroll hours are the product of the number of workers and their average workweek, it is algebraically possible for the number of workers to remain constant while weekly hours per worker (and thus total payroll hours) decline when the minimum wage rises (see Appendix C.1 for the theoretical background).

the minimum wage is the mechanism of minimum-wage setting (Boeri, 2012). In much of the standard literature (e.g. Neumark and Wascher, 2008), the effects of raising the minimum wage are analysed under extreme assumptions – either of a fully competitive labour market or a fully monopsonistic labour market. On the one hand, if the minimum wage is increased in a competitive labour market, some firms will be unwilling to pay the higher wages, so they lay off employees (e.g. Neumark and Wascher, 2008). On the other hand, in a monopsonistic labour market in which firms may have already been paying wages lower than the laissez-faire equilibrium wage, an increase in the minimum wage can raise the incomes of low-wage workers at the expense of the firms' profit without a corresponding decrease in employment (e.g. Manning, 2003). However, in contrast to these simplistic models, heavy bargaining and power struggles are involved in determining the minimum wage (Brown, 2009). Thus, I have used the 'bargained minimum wage' model, a theory based on the right-to-manage model suggested by Nickell and Andrews (1983), as guidance in conducting my empirical analysis, as that model better reflects the actual labour market environment in South Korea.[4]

Using the bargained minimum wage model, this paper examines two hypotheses. The first one assumes a non-linear relationship between increases in the minimum wage and hours worked, and the second one assumes that the effects on employment from modest increases of the minimum wage under business-friendly governments are different from the effects from a sharp increase in the minimum wage under a labour-friendly government. To validate these hypotheses, this study applies two-step generalized method of moments (GMMs), difference-in-differences, and regression discontinuity (hereafter, 'RD') designs based on a longitudinal survey of labour markets and income activities in South Korea in the period from 2009 to 2018. Moreover, these estimators are constructed using different definitions of the group directly affected to examine whether there is a heterogeneous impact under the single minimum wage system in South Korea.

In particular, I implemented the RD design, which has emerged as one of the most credible non-experimental strategies to analyse causal effects (Cattaneo et al., 2019). This framework uses the fact that the observed and unobserved characteristics of workers earning on either side of near the cut-off wage are very similar, and therefore the main difference between them is the wage-push pressure on employers. In this regard, by comparing workers who are earning just above the cut-off wage with those just below this wage, we can evaluate the employment effects of the large increase in the minimum wage in 2018.

[4]The right-to-manage model assumes that the minimum wage is set during a collective bargaining procedure between the labour unions and the employers' association at the national level to maximise the surplus of their members. In this respect, the right-to-manage structure is consistent with the actual labour market in South Korea where minimum wages are set at the national level by agents who cannot bargain over employment because they do not know the technological constraints of individual firms. Institutional backgrounds for the minimum wage system in South Korea are illustrated in Section 3.

The estimation results show that reducing the hours worked appears to be the tactic adopted by employers against the large increase in the minimum wage in South Korea, with three key dimensions. First, the results from GMM estimations show that a 14.9% rise in the real minimum wage in 2018 seems to hit the tipping point, which is 5.5% in my estimation. That is, the minimum wage and employment outcomes in terms of work hours can show both a positive relationship at a relatively low wage floor and a negative relationship at higher levels. In addition, different curvatures of non-linearity are observed among different groups by employment status, firm size and the age of the employee for the single minimum wage system in South Korea. Second, the unexpected double-digit growth in the 2018 minimum wage has led to a reduction in the hours worked by workers whose wages were lower than the 2018 minimum wage ('affected workers'), compared with the modest increases of minimum wages under business-friendly governments during the period 2009–2017, and compared with the workers whose wages were modestly higher than the 2018 minimum wage ('unaffected workers'). In addition, I have confirmed that certain sub-groups are more susceptible to the adverse effects of a large increase in the minimum wage in 2018. For example, regular workers, workers whose firm size is relatively large and teenage and young adult workers (aged 16-29 years) are particularly susceptible to cuts in hours, partly because of the strong employment protection for temporary workers and small businesses under the left-wing government. In this respect, if there is no discernible effect of the minimum wage on overall employment, it might mask the non-linearity in the minimum wage shock, otherwise there is some heterogeneity in responses across types of workers.

Thus, a question naturally arises from these estimation results: Is the reduction in the hours worked actually harmful for low-wage workers in South Korea? Evidence that increases in the minimum wage result in adverse employment outcomes in terms of work hours does not imply that raising the minimum wage necessarily hurts low-wage workers. South Korea is notorious for its workaholic culture, which has contributed to its rapid industrialisation over the past half-century and transformed the war-torn country into the world's eleventh largest economy. For example, South Koreans worked an average of 2,228 hours in 2008, the longest hours among the countries in the Organisation for Economic Cooperation and Development (OECD) (see Figure C.3), although this had decreased to 1,993 hours by 2018 (ranked the second longest after Mexico).

However, the long working hours of low-paid workers have not changed as much as their productivity has since the 1997–98 Asian financial crisis (see Figure 3.3). This gap between labour productivity and real wages is referred to as 'exploitation rent' (Stiglitz, 2015a,e).[5]

[5] Stiglitz (2015a,e) defines 'exploitation rents' as rents arising from monopoly power and political influence (i.e. taking advantage of imperfections in corporate governance laws and of asymmetries in bargaining power between workers and firms).

In this regard, the reduced hours worked can be helpful for low-paid workers by removing the exploitation rents that have been in place since the International Monetary Fund (IMF) restructuring program after the 1997–1998 Asian financial crisis. Under the bargained minimum wage, because employers try to maximise their profit, the sharply increased minimum wage and the reduction in the hours worked offset each other, implying that the earnings of low-paid workers will roughly remain the same.

We ask whether higher minimum wages can actually help low-paid workers. The answer depends on what a given worker values more. For those who hope for greater earnings at the expense of more working hours, it cannot help. However, for those who desire a better work-life balance at the cost of greater earnings, it definitely can help.

This study offers three contributions to the research on the effects of the minimum wage on employment. First, the bargained minimum wage framework allows us to model the determination of minimum wages more realistically than models that make the extreme assumptions of perfect competition or monopsonistic competition in the labour market. Second, this study examines labour demand adjustments at the intensive margins, which are rarely studied, despite adjustments at the extensive margins being a less plausible option for employers as an immediate response to increases in the minimum wage, as argued by Brown (1999), Michl (2000), Stewart and Swaffield (2008), and Hirsch et al. (2015).[6] Last, this paper is believed to be the first to estimate a non-linear relationship between increases in the minimum wage and hours worked. If the relationship is non-linear, models that do not allow for non-linearity will lead to a downward bias in the relationship between the two variables, and the weak employment effects found in the literature may be driven by the fact that the models are mis-specified.

The rest of the paper proceeds as follows. Section 2 reviews the relevant literature. Section 3 describes the institutional context of the minimum wage in South Korea and the theoretical considerations that motivate my analysis. Section 4 reviews the data, and Section 5 presents the empirical results and those implications. Section 6 concludes.

3.2 Literature Review

This study relates closely to the scantiness of the literature that has examined the non-linear effects of the minimum wage. Despite several decades of microeconometric evidence, the

[6] In response to an increase in the minimum wage, managers see employment cuts as a relatively costly and perhaps counterproductive option, regarding them as a last resort (Hirsch et al., 2015).

impact of the minimum wage on low-wage employment remains an empirically controversial question. Opponents of increasing the minimum wage argue that it makes low-paid workers worse off because many of them lose their jobs (e.g. Meer and West, 2016; Neumark and Wascher, 2008). In contrast, proponents argue that increasing the minimum wage has no discernible effect on employment and that sometimes it might even have a positive effect (e.g. Azar et al., 2019; Card and Krueger, 1995; Cengiz et al., 2019; Doucouliagos and Stanley, 2009; Dube et al., 2010; Schmitt, 2013). Even though there have been some studies on the non-monotonic impact of the minimum wage on inequality (e.g. Autor et al., 2016; Lee, 1999), non-linearity of the effects on employment has not been investigated much in the literature, perhaps because it is thought hard enough to detect a linear effect (Manning, 2016). Nevertheless, using a standard search model with minimum wages, Flinn (2006) argues that a higher minimum wage has non-linear effects on unemployment because a higher minimum wage affects a greater proportion of workers as it cuts deeper into the underlying productivity distribution among the workers, which is not uniform.

This study is also similar to research that introduces monopsony or oligopsony into an analysis of the effects on employment of increases in the minimum wage (notably, Manning, 2011). When one departs from the assumption of a perfectly competitive labour market, theoretically, a rise in the minimum wage can increase employment (Bhaskar et al., 2002). Notably, Azar et al. (2019) provide the first direct test for the mediating role of labour market concentration, a key source of monopsonistic and oligopsonistic power, on the employment effect of the minimum wage in the United States' general merchandise sector. They present a robust and significant increase in the employment elasticity with respect to the minimum wage in concentrated occupational labour markets. Similarly, Manning (2016) addresses the question of how high the minimum wage could be raised without having a significant employment effect, dubbed the 'elusive employment effect' of minimum wages. I contribute to this debate by suggesting the non-linear employment effects predicted by the bargained minimum wage model, which is positioned between perfect competition and monopolistic competition.

This study is also related to research that explores the impact of a large increase in the minimum wage on employment, which is rarely studied because minimum wages have been raised at modest rates in most advanced economies. Remarkably, Harasztosi and Lindner (2019) assess the margins of adjustment used by firms in response to an unexpected double-digit increase in the minimum wage in Hungary. They showed that most firms responded to the increase in the minimum wage by raising wages instead of destroying jobs, and these large increases in labour costs were absorbed by higher output prices (approximately 75% of the increase in the minimum wage) and lower profits (approximately 25%). While Harasztosi and Lindner (2019) deal with firms' responses and adjustments to employment at the extensive margins by exploiting administrative data on firms filing balance sheet statements to the tax authority, I focus on workers' responses and adjustments to employment at the intensive margins by using a longitudinal survey of labour markets and income activities of individuals.

This study focuses on potential adjustments of hours by an employer in response to an increase in the minimum wage. Although research in this area is limited, Brown (1999) suggests that when the minimum wage rises, the effect on the hours worked per employee is more pronounced than the effect on the number of workers. Consistent with this, Stewart and Swaffield (2008) find a reduction in the hours worked of low-paid workers because of the introduction and uprating of the minimum wage using difference-in-difference estimators based on the employer-based New Earnings Surveys and the Labour Force Survey of the United Kingdom. Interestingly, Michl (2000) suggests that one explanation for the differences between the results of Card and Krueger (1994) and those of Neumark and Wascher (2000) in their analyses of the impact of the increase in the 1992 minimum wage on the fast-food industry in New Jersey may lie in their different treatment of the hours worked per employee, besides the differences in the data they use. In particular, Neumark and Wascher (2000) use total payroll hours as their dependent variable, so they would have captured any adjustment in hours, while Card and Krueger (1994) use the number of workers as their dependent variable, so they would not have.

In addition, Neumark and Wascher (2008) suggest that, when employers react to changes in the minimum wage, they may adjust the level of labour inputs by reducing the total number of hours worked across all minimum-wage employees rather than making specific workers redundant. Metcalf (2008), using data from the British Retail Consortium and the Union of Shop, Distributive, and Allied Workers, also argues that, following an uprating of the minimum wage, managers look closely at potential adjustments of working hours to offset rising labour costs. Couch and Wittenburg (2001), using state-level data on teenage working hours in the United States, argue that studies based on aggregate employment understate by roughly 10–30% the impact of minimum wages on labour demand compared to the studies based on working hours. In contrast, Zavodny (2000), who investigates the same question for teen workers using individual-level panel data in the United States, finds no evidence of a negative association between hours of work and the minimum wage.

Finally, this study is related to studies of 'localised' changes to employment around the change in the minimum wage. Dickens et al. (2015), using the RD design, consider the effect of age-related increases in the minimum wage on the employment of low-skilled young workers in the United Kingdom. They find that low-skilled young workers are significantly more likely to be employed as they turn 22. Furthermore, Fidrmuc and Tena (2018), also using the RD design, investigate discontinuities in terms of both the level and the slope of employment probabilities at different ages around the threshold. Their results show that turning 22 did not significantly change the employment probability, but the results did show a significant change in the slope of the probability of being employed approximately one year before. I also use the RD design to measure the impact of a large increase in the minimum wage on the hours worked of low-paid employees in South Korea.

3.3 Background

3.3.1 Institutional context

Currently, 28 out of 36 OECD countries have a statutory minimum wage.[7] As shown in Table 3.1, the minimum wage system varies considerably across these 28 OECD countries in terms of determination process, legal grounds, uniformity and enforcement. In South Korea, the minimum wage system has four characteristics: (i) a bargaining process combined with the government's political preference when the minimum wage is being set, (ii) a single minimum wage that applies to all workers, (iii) strong enforcement and (iv) high effectiveness as a poverty-alleviation policy. These charateristics are considered in turn.

Bargaining process. South Korea's minimum wage is determined by the relative bargaining power of employees and employers combined with the government's political preference. The minimum wage is determined by the Minimum Wage Commission (MWC) under the control of the Ministry of Employment and Labour. The MWC has 27 members: 9 representatives of employees, 9 representatives of employers, and 9 representatives of the public interest. The government has the *de facto* decision-making power because all public interest committee members are appointed by the government, and the chairperson and vice chairperson of the MWC must be elected from those public interest committee members. The Minster of Employment and Labour asks the MWC to review the minimum wage by March 31 every year. The MWC must then submit the minimum wage bill within 90 days of the day after the request was received. If the Minister considers that it would be difficult

[7] In the 8 OECD countries without a legal minimum wage, such as Nordic countries, Austria, Italy and Switzerland, a large part of the workforce is covered by sector-level collective agreements.

Table 3.1 Minimum wage system in 28 OECD countries

Country	Determination process	Legal ground	Uniformity (Yes/No)	Enforcement	% of median wages (2009 → 2018)
Australia	Consultation	Act	No	Strong	54.4 → 54.1 (Δ0.3)
Belgium	Bargaining	Constitution	No	Weak	49.4 → 46.3 (Δ3.2)
Canada	Consultation	Act	No	Weak	42.2 → 51.4 (+9.2)
Chile	Consultation	Act	Yes	Weak	70.7 → 69.3 (Δ1.4)
Czech Rep.	Consultation	Act	Yes	Weak	38.3 → 41.8 (+3.5)
Estonia	Bargaining	Act	Yes	Weak	39.8 → 43.1 (+3.3)
France	Consultation	Act	Yes	Weak	62.8 → 61.6 (Δ1.2)
Germany	Bargaining	Act	Yes	Strong	– → 45.6 (–)
Greece	Bargaining	Act	No	Strong	48.2 → 47.5 (Δ0.7)
Hungary	Consultation	Act	No	Weak	46.7 → 51.8 (+5.1)
Israel	Bargaining	Act	Yes	Weak	56.3 → 59.4 (+3.1)
Ireland	Consultation	Act	No	Strong	46.8 → 47.5 (+0.7)
Japan	Consultation	Constitution	No	Weak	36.2 → 42.0 (+5.8)
South Korea	Bargaining	Constitution	Yes	Strong	45.2 → 58.6 (+13.4)
Latvia	Consultation	Constitution	Yes	Weak	47.3 → 50.4 (+3.1)
Lithuania	Bargaining	Constitution	Yes	Weak	44.2 → 51.2 (+7.0)
Luxembourg	Government	Act	Yes	Weak	54.6 → 53.8 (Δ0.8)
Mexico	Bargaining	Constitution	No	Strong	36.8 → 42.2 (+5.4)
Netherlands	Government	Act	Yes	Weak	50.7 → 47.0 (Δ3.7)
New Zealand	Government	Act	Yes	Strong	59.1 → 61.4 (+2.3)
Poland	Bargaining	Constitution	Yes	Weak	45.9 → 53.1 (+7.2)
Portugal	Consultation	Constitution	Yes	Strong	49.9 → 61.4 (+11.5)
Slovak Rep.	Bargaining	Constitution	Yes	Weak	45.4 → 49.3 (+3.9)
Slovenia	Bargaining	Act	Yes	Weak	50.8 → 58.7 (+7.9)
Spain	Consultation	Constitution	Yes	Weak	39.1 → 41.2 (+2.1)
Turkey	Consultation	Constitution	Yes	Weak	71.3 → 70.9 (Δ0.4)
UK	Consultation	Act	No	Weak	46.1 → 54.5 (+8.4)
US	Government	Act	No	Strong	37.1 → 32.7 (Δ4.4)

Notes: Eight OECD countries that do not have a national minimum wage are Austria, Denmark, Finland, Iceland, Italy, Norway, Sweden and Switzerland. Germany introduced its national minimum wage in 2015. Boeri (2012) classified the determination process into (i) bargaining – the minimum wage is set by social partners and then simply ratified by the government, or determined by a tripartite body where representatives of the government, unions and employers; organizations are represented on an equal stance (the government typically acts as a go-between); (ii) consultation – the minimum wage is set after formal consultations between the government and representatives of employers and workers; and (iii) the government sets the minimum wage unilaterally with no formal consultation with social partners. Neumark and Corella (2019) divided countries based on the enforcement of the minimum wage as: (i) countries with 'no enforcement' do not penalise violations of the minimum wage law, (ii) countries with 'weak enforcement' have low-cost fees for a violation; and (iii) countries with 'strong enforcement' have severe penalties for not abiding by the law, such as imprisonment or shutting down the firm.
Source: The author utilised OECD.stat (2020) and ILO (2014)

to set the minimum wage based on the proposal, the MWC is requested to deliberate on it again. Otherwise, the new minimum wage is announced by August 5. The new minimum wage takes effect on January 1 of the following year.

Single minimum wage. The minimum wage applies to all businesses or workplaces with one employee or more, regardless of their employment status and nationality; that is, it applies to temporary and part-time employees (as well as regular ones) and to employees who are foreign nationals.[8] There is also no variation in the minimum wage across age groups, regions or industries. Therefore, the minimum wage is likely to lead to different employment effects for different types of workers (by employment status, firm size and age) due to the distributions of productivity and bargaining power (e.g. Flinn, 2010).

Strong enforcement. According to the classification of Neumark and Corella (2019),[9] South Korea's degree of enforcement of the minimum wage law is strong. An employer paying less than the minimum wage is sentenced to up to three years of imprisonment or fines of up to 20 million Korean won (approximately USD16,300). This implies that once the minimum wage is set, there may be a negligible amount of non-compliance. Moreover, as described in Figure C.4, the strength of South Korea's employment protection for temporary workers lies in the top third of OECD countries, at a level similar to Italy or Greece (OECD.stat, 2020). However, protection for permanent workers is in the bottom third. So, it is likely that restrictions on firing low-paid workers substantially prevent employers from responding strongly (i.e. laying off workers) to the increased minimum wage.

High effectiveness. For those who might benefit from higher minimum wages, higher taxes and reduced benefits connected to these wages can offset a large part of the increase in the minimum wage. Figure C.5 shows the share of an increase in the minimum wage that adds to net income across the OECD countries. For example, minimum-wage workers in Ireland and Japan would receive less than a tenth of an increase in the minimum wage in their pay packets. In contrast, the net gains seen by minimum-wage workers in South Korea are among the top three for workers in OECD countries. This implies that the minimum wage

[8]With the government's permission, 10% reduction can be applied for very exceptional cases, such as workers who have been on probation for less than three months and those engaged in monitoring activities.

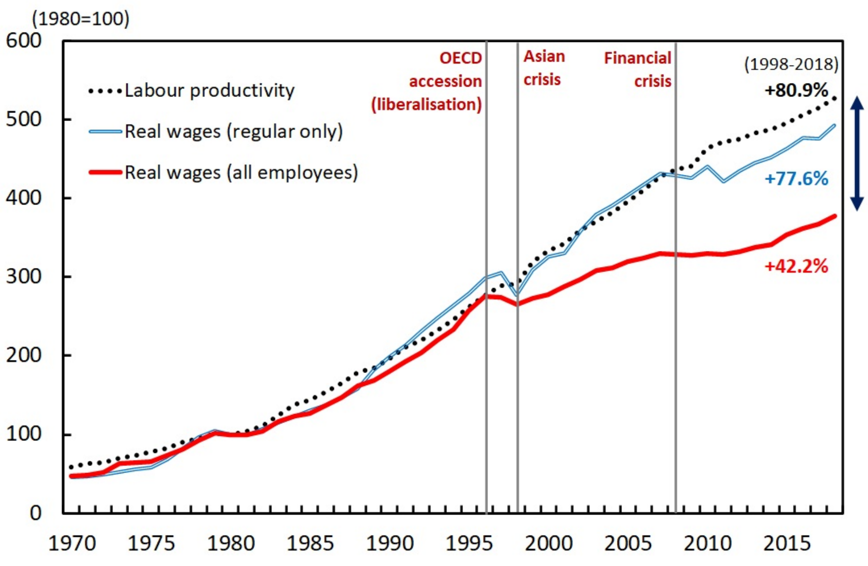
[9]Using the International Labour Organisation (ILO) Database of National Labour, Social Security and Related Human Rights Legislation (NALEX), Neumark and Corella (2019) presented three categories: (1) countries with 'no enforcement' are those that do not penalise violations of the minimum wage law, (2) countries with 'weak enforcement' are those that have low-cost fees for a violation, and (3) countries with 'strong enforcement' are those that have severe penalties for not abiding by the law, such as imprisonment or shutting down the firm.

policy in South Korea would be highly effective as a poverty-alleviation policy, because an increase in the minimum wage may lead to significant income gains for the targeted workers because of low tax burdens on the incomes of low-wage workers and low transfers to them.

3.3.2 The post-2018 evolution of the minimum wage

Figure 3.3 plots labour productivity in South Korea, measured by real GDP per employee (1980 = 100), and real wages, measured by labour compensation per employee adjusted for inflation (1980 = 100), over the period 1970–2018. While labour productivity has continued to grow since 1970, real wages have stagnated since the 1997–98 Asian financial crisis. Real wages increased by 42.2% from 1998 to 2018, a little more than half of labour productivity, which grew by 80.9% during the same period.

Fig. 3.3 Divergence of labour productivity and real wages in South Korea



Notes: Labour productivity stands for GDP per employee adjusted for inflation (1980 = 100), and real wages represent labour compensation per employee adjusted for inflation (1980 = 100). Regular only implies regular workers working at firms with 10 or more employees.
Source: The author's calculation using labour productivity and inflation data from the Bank of Korea and wage data from the Ministry of Employment and Labour

To understand how significant this divergence in labour productivity and real wages is, consider that both increased at the same pace between 1970 and 1996, approximately tripling in that period including for regular and other workers. If we consider only regular workers working at firms with 10 or more employees, the divergence was not significant. These results imply that, although employees in South Korea were increasing their productivity at an almost constant pace regardless of the period, the fruits of their labours primarily accrued to corporate profits after 1997 and the restructuring reforms imposed by the IMF (e.g. Kim, 2010; Shin, 2013). The divergence is called 'exploitation rent' (Stiglitz, 2015a,e). Regular workers received wages that were in line with the increases in their labour productivity, but non-regular workers (generally, low-paid workers) did not. We might say they have been exploited in the sense that the growth in their real wages was lower than the growth of their productivity.

Against this backdrop, after the unprecedented impeachment of a former right-wing president in mid-2017 in South Korea, raising the minimum wage became the chief component of the new left-leaning government's economic plan. It relied on what it called an 'income-led growth policy' instead of using exports to prop up the economy. As a result, the minimum wage leapt by 16.4% to 7,530 Korean won (USD6.7) an hour in 2018 in nominal terms. Moreover, it was announced that the minimum wage would be raised to 10,000 won (USD8.9) an hour by 2020. The 2018 increase brought the minimum wage in South Korea from just over 45% of the median wage in 2009 to just under 60% of it. If the government plan to raise it to 10,000 won by 2020 came true, it would mean that the South Korean minimum wage would be roughly 70% of the median wage, far higher than the level in other advanced economies (see Table 3.1).

This change in policy broke the pattern that had been established under decades-long rule by earlier business-friendly governments. This is highlighted in Table 3.2, which describes the determination of minimum wages through a bargaining process during the period 2009–2018. For example, at the beginning of the 2017 minimum wage negotiation, the representatives of employees and employers proposed increases in the minimum wage of 65.8% and 0%, respectively. However, the right-wing government at that time took the employers' side, awarding only the 7.3% increase that the employers offered in the final round of negotiation. In contrast, for the 2018 minimum wage, the left-leaning government fully accepted the 16.4% increase proposed by the representatives of employees in the final round of the negotiation.[10] It is clear that the large increase in the 2018 minimum wage represents a permanent departure from the previous pattern, in which the government largely sided with employers.

Accordingly, by exploiting the post-2018 evolution of the minimum wage in South

[10]This rate was unexpectedly high because most public interest committee members appointed by the former right-wing government remained unchanged.

Table 3.2 Bargained rates of increases in the minimum wage in South Korea

		2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
First	Employees	26.3	28.7	26.0	25.2	26.2	21.6	28.6	79.2	65.8	54.6
round	Employers	0.0	-5.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.4
Final	Employees	8.9	3.9	8.8	10.6	9.1	19.1	15.0	45.2	–	16.4
round	Employers	4.1	1.1	2.8	6.0	3.4	1.0	2.1	2.4	7.3	12.8
Bargained rate		6.1	2.8	5.1	6.0	6.1	7.2	7.1	8.1	7.3	16.4

Notes: A left-wing government took office in mid-2017 after decade-long right-wing governments. In the final round of the 2017 minimum wage negotiation, all the representatives of employees left the MWC without voting. Years represent the following year of the MWC.
Source: The Minimum Wage Committee

Korea, we can ask whether the sudden regime change and the resulting hike in the minimum wage had affected the employment of low-paid workers. To evaluate this, we must first set up a theoretical framework that reflects the actual Korean labour market.

3.3.3 Motivating theory

The potential effects on employment of increases in the minimum wage are closely related to the structure of setting the minimum wage (Boeri, 2012). On the one hand, if the minimum wage is increased in a competitive labour market, some firms will be unwilling to pay the higher wages, and so they lay off employees (e.g. Neumark and Wascher, 2008). On the other hand, in a monopsonistic labour market in which firms may have already been paying wages lower than the laissez-faire equilibrium wage, a rise in the minimum wage can raise the incomes of low-wage workers at the expense of the firms' profits without a corresponding decrease in employment (e.g. Manning, 2003).[11]

However, in contrast to these simplistic models, heavy bargaining and power struggles are involved in determining the minimum wage (Brown, 2009). Thus, I have used the bargained minimum wage model, a theory based on the right-to-manage model suggested by Nickell and Andrews (1983), as guidance in conducting my empirical analysis. More specifically, this model considers the regime in which the minimum wage is set within a collective bargaining procedure between the labour unions and the employers' association

[11] Similar results can arise in search and matching models of the labour market (e.g. Flinn, 2010).

at the national level, in which both sides try to maximise the surplus of their members. In this regard, the right-to-manage structure is consistent with the actual labour market in South Korea: The minimum wage is set at the national level by agents who cannot bargain over employment because they do not know the technological constraints of individual firms.

Let us now consider two extreme cases of labour markets in determining the bargained minimum wage.

First, assume that in a competitive labour market, the marginal value of a job (v) is a decreasing function of employment (l): $v = Al^{-\eta}$, where A represents labour productivity and $\eta \in [0, \infty)$ stands for the demand elasticity of labour. Replacing the marginal value of a job (v) with the market wage (w), and solving for employment (l) yields the labour demand schedule $L^D = (A/w)^{1/\eta}$. The supply side of the labour market is described as the cumulative distribution function $G(\cdot)$ of the reservation wages. For simplicity, we assume that the labour supply schedule is given by $L^S = G(w) = w^{1/\varepsilon}$, where $\varepsilon \in [0, \infty)$ is the supply elasticity of labour.

The laissez-faire equilibrium wage (w^*) will then be at the intersection of the two curves: $w^* = A^{\eta/\eta+\varepsilon}$. The total surplus of workers can be obtained by integrating the density function of the distribution of reservation wages over the relevant range (for simplicity, let us neglect constants of integration):

$$wl - \int_{\zeta}^{l} x^{\varepsilon} dx = wl - \frac{l^{1+\varepsilon}}{1+\varepsilon}. \tag{3.1}$$

Next, consider the monopsony case.[12] Labour demand originates from a sole employer facing the upward labour supply curve $L^S = w^{1/\varepsilon}$.[13] This employer chooses the employment level (l) that maximises her profits:

$$\frac{Al^{1-\eta}}{1-\eta} - wl. \tag{3.2}$$

The monopsony wage (w^m) will be derived from the first-order-condition for employment: $w^m = (A/1+\varepsilon)^{\varepsilon/\varepsilon+\eta}$, which is clearly lower than the laissez-faire equilibrium wage (w^*).

Finally, consider the cases in which the minimum wage is determined by collective bargaining over wages, allowing employers to choose the profit-maximising employment

[12] Extensions to n-firms with some monopsony power do not significantly alter the results (Manning, 2003).

[13] There are different sources of monopsony power that generate an upward labour supply curve: search frictions, non-wage working conditions (e.g. flexible working time arrangements, commuting time and training opportunities), firm concentration and limited mobility of workers across labour market regions (Bachmann and Frings, 2017).

level. Under this condition, the bargained minimum wage (w^b) will maximise the product of the surplus of employees (from equation (3.1)) and that of employers (from equation (3.2)) under the Nash-bargaining rule:

$$\underbrace{[\omega^b l - \frac{l^{1+\varepsilon}}{1+\varepsilon}]^{\beta}}_{\text{employee}} \underbrace{[\frac{Al^{1-\eta}}{1-\eta} - \omega^b l]^{1-\beta}}_{\text{employer}} \tag{3.3}$$

where the two surpluses are weighted by the bargaining power of employees, $\beta \in [0,1]$. Maximising equation (3.3) with respect to the bargained minimum wage (ω^b) under the binding constraint on the labour demand schedule $L^D = (A/\omega)^{1/\eta}$ yields:

$$\omega^b = (\mu^b)^{\frac{\varepsilon}{\varepsilon+\eta}} A^{\frac{\varepsilon}{\eta+\varepsilon}} \tag{3.4}$$

where $\mu^b \equiv \{1-(1-\beta)(\eta+\varepsilon)/(1+\varepsilon)\}/(1-\eta)$ is a shifter of wages over the opportunity cost of working (Bertola and Boeri, 2002). Equation (3.4) shows that a higher bargaining power of employees can increase the minimum wage (i.e. $\partial\omega^b/\partial\beta > 0$). Moreover, considering that labour's bargaining power is associated with the government's political preference, the effect of an increase in the minimum wage on employment will differ according to the government's position on the political spectrum.

Figure C.6 shows that the effect on employment of a change in the minimum wage would be different based on how the minimum wage is set. In a perfectly competitive labour market, there is an unambiguously negative relationship between the minimum wage and employment when the minimum wage is set above the laissez-faire equilibrium wage (w^*). In contrast, in a monopsonistic labour market, we can see an inverted U-type relationship between the minimum wage and employment. Any minimum wage set in the range between the monopsony wage (ω^m) and the laissez-faire equilibrium wage (ω^*) increases employment (i.e., $l \in [l^m, l^*]$), but it reduces employment when it is set above the laissez-faire equilibrium wage (w^*). The reason the effects on employment are not negative in certain intervals in a monopsonistic labour market is that the minimum wage plays a role in partially reducing the excessive rents appropriated by a firm's monopsony power.

Most importantly, the bargained minimum wage (w^c) in the bargained minimum wage model can limit the range of the minimum wage between the monopsony wage (ω^m) and the laissez-faire equilibrium wage (ω^*). That is, the bargained minimum wage could be set within the range between the monopsonistic equilibrium wage and the laissez-faire equilibrium wage (i.e. $\omega^m = \omega^C_{\beta=0} \leq \omega^C \leq \omega^C_{\beta=1} < \omega^*$). However, if the government intervenes and sets minimum wages higher than the laissez-faire equilibrium wage ($\omega^C > \omega^*$), the employment effects may turn out to be negative.

This simple theory predicts that an increase in the minimum wage and employment can have a positive relationship at a relatively low wage floor but a negative one at higher levels. This comes from the influence of the bargaining power of employees combined with the government's political preference.

In this respect, two hypotheses can be tested. First, there may be 'non-linear' effects on employment of increases in the minimum wage over the medium and the long term, as suggested by Manning (2016). Second, if that is so, then in the short run, increases in the minimum wage may have 'different' effects on employment when the minimum wage is set between a very high rate (e.g. that of the year 2018 in South Korea) and a moderate rate (e.g. the period 2009–2017 in South Korea), as argued by Harasztosi and Lindner (2019).

3.4 Data

I investigate the issue at hand using the Korean Labour and Income Panel Study (KLIPS),[14] a longitudinal survey of labour markets and income activities of individuals who live in urban areas conducted once a year. The cohorts at the start of the survey in 1998 were 13,321 individuals from 5,000 households. As argued by Currie and Fallick (1996), by using individual-level data, we can reduce the chances that estimation results are contaminated by correlations between aggregate working hours and general economic conditions. The KLIPS collects information on the sample households and their individual members aged 15 or older. The target sample of this study is wage and salary workers. I restrict this analysis to the period from 2009 to 2018 because the macroeconomic trends of the Korean economy, such as growth rates, inflation rates, and employment rates, have been much more stable since the 2008 global financial crisis than before (see Figure C.2). I also intend to compare the effects on employment of increases in the minimum wage between the two distinct periods: the period of right-wing governments from 2009 to 2017, and that of the left-wing government from 2017 to 2018.

The dependent variable is the hours worked per month. Recently, there appears to have grown a consensus that when the minimum wage is set at a moderate level, the employment

[14]The KLIPS was originally designed to follow the basic structure of the Panel Study of Income Dynamics (PSID) of the United States, and it was supplemented with questionnaires related to labour market activity to follow the National Longitudinal Survey (NLS) of the United States. See Appendix C.1 for the detailed description of the KLIPS.

effect is not significant (e.g. Doucouliagos and Stanley, 2009; Hirsch et al., 2015; Schmitt, 2013). One plausible explanation for this modest effect on employment is that, at a moderate level, minimum wages account for only a small share of an employer's total costs, so firms can absorb the increase in the minimum wage in a variety of ways other than by cutting payroll. The options include reducing non-wage costs, raising prices, boosting productivity and accepting lower profits (see Harasztosi and Lindner (2019) for the extensive surveys). However, even when the minimum wage rises at a large rate, immediate adjustments to employment can take place at the intensive margin (i.e. a reduction in the hours worked per employee), and not just at the extensive margin (i.e. a reduction in the number of workers). This is because it takes time to alter capital investments in machinery, buildings and technology, which might allow a more efficient operation. However, employers usually adjust working hours because – particularly in the short run – it is easier to alter hours than to change the number of employees (Brown, 1999; Hamermesh, 1995, 2014; Stewart and Swaffield, 2008; Zavodny, 2000).[15] For example, to put it bluntly, most employers would rather choose that all workers work for 10% fewer hours than see 10% of them lose their jobs, while the remaining 90% keep their jobs with no change in weekly working hours.

In the Korean case, there was little evidence that the increase in the minimum wage was passed on to consumers through higher prices (see Figure C.2a), a possibility suggested by Aaronson and French (2007) and Lemos (2008). Nor was it paid by employers through lower profits (see Figure C.2b), a possibility suggested by Draca et al. (2011) and MaCurdy (2015). Furthermore, there were negligible changes in the employment rate and the unemployment rate for all age groups (see Figure C.2c and C.2d) partly because of strong enforcement of the minimum wage and restrictions on firing temporary employees in South Korea, as seen in Section 3.1.

Nevertheless, to consider the possibility of an adjustment at the extensive margins to the large increase in the minimum wage, I treat workers in involuntary unemployment as having worked zero hours, rather than as missing data points. In addition, reflecting the limitation of a self-administered survey, approximately 9% of the surveyed workers reported working more than the legal maximum hours (68 hours per week). I treat them as measurement errors and adjust them by using interpolation and top coding.

To check whether the increased minimum wage has been passed on to employee earnings as labour costs for employers, I use total real wages, the surveyed nominal wages (including such things as overtime pay and meal allowance) divided by consumer price levels, instead of the hours worked as the dependent variable.

[15] Admittedly, in the long run, a firm's choice of the mix of worker-hours depends on the extent of the fixed costs of employment, technology and the productivity-hours schedule, the labour supply, and the presence and effectiveness of a union (Stewart and Swaffield, 2008).

Finally, I examine the evolution of the frequency distribution of weekly hour worked and monthly earnings over time. Figure C.1 shows the distribution of hours worked per week and real wages per month in 2017 (the last year before the minimum wage hike by the business-friendly government) and in 2018 (one year after the minimum wage hike by the labour-friendly government). First, by comparing Figure C.1a and C.1b, we can see that the substantial increase in the minimum wage clearly altered the distribution of hours worked: the portion of workers who were working more than forty hours was decreased; instead there was a much larger spike appeared in the 2018 distribution. Second, by comparing Figure C.1c and C.1d, we can see that the upsurge of the minimum wage shifted the distribution of earnings to the right.

The main explanatory variable in my empirical study is the monthly real minimum wage. Because the MWC announces the weekly minimum wage in nominal terms, I convert it into the monthly minimum wage in real terms using the Consumer Price Index (CPI) and each individual's weekly hours worked. An important procedure that is sometimes overlooked in the literature is taking weekly holiday allowances into account. More specifically, according to the Labour Standards Act of South Korea, if the hours worked per week are greater than 15 hours,[16] the employer must pay a holiday allowance, and the maximum working week for which the employer must pay a holiday allowance is limited to 40 hours. Accordingly, for those workers who have worked at around 15 and 40 hours per week, there is some variation in the reference period (and thus in the labour costs per worker) between employees.[17] These variations may induce cost-sensitive employers to shorten workweeks or to reduce some workers' working hours to less than 15 hours (so they do not have to pay holiday allowances) or to less than 40 hours (so they do not have to pay costly overtime remuneration).

Control variables include gender, marital status, age, education attainment, employment status, firm size and industry classification, as usual. To guard against the possibility that the macroeconomic environment might contaminate our results, I investigate the macroeconomic trends of South Korea during the period 2009–2018. However, the evidence in Figure C.2 shows that there were no dramatic macroeconomic shocks around the year 2018 that could alter our conclusions significantly. For example, real GDP growth was approximately 3% before and after the exceptionally large increase in the minimum wage in 2018. In line with

[16] Employers in South Korea are not required to grant weekly holiday allowance, annual paid leave, severance pay or social insurance to employees who work less than 15 hours a week. So, by splitting working hours down to less than 15 hours a week, employers can reduce the burden from the higher minimum wage.

[17] For example, those who are working eight hours a day and five days a week, the reference period is calculated as 8 hours $\times$ (5+1) days $\times$ 4.35 weeks = 209 hours, and the monthly minimum wage is 1,573,770 Korean won, given that the minimum wage in 2018 is 7,530 Korean won. In contrast, those who are working four hours a day and three days a week, the reference period is calculated as 3 hours $\times$ 3 days $\times$ 4.35 weeks = 117 hours, and the monthly minimum wage is 393,066 Korean won.

the modest growth rate, aggregate labour market conditions were stable. The employment rate and the unemployment rate were increasing gradually every year between 2009 and 2018 for all age groups, while the non-regular employment ratio was falling gently. The inflation rate was relatively stable at approximately 1.5% from 2012 onward. Nevertheless, I apply time-fixed effects to control for unexpected variation or special events that may affect the outcome variable.

As suggested in Draca et al. (2011), to allow differential impacts on employment of workers, I include a continuous measure of the treatment intensity (TI), defined as

$$TI_{i,t} = -(\frac{w_{i,t-1}}{MW_{i,t}} - 1) \times 100 \tag{3.5}$$

where $w_{i,t-1}$ is the current wage and $MW_{i,t}$ is the new minimum wage. This implies that because employers must pay at least the minimum, it is likely that raising the minimum wage would become more burdensome for employers with a larger difference between the current wage and the new minimum wage. This is defined as 'wage-push pressure' for employers. Table 3.3 describes the summary statistics for the key variables for the whole period.

Table 3.3 Summary statistics for the key variables

Variable	Unit	Obs.	Mean	Std.	Min.	Max.
Real minimum wage (monthly)	10^3 KRW	55,324	1,209.7	334.1	20.6	2,380.8
Real wage (monthly)	10^3 KRW	55,324	2,335.1	1,544.3	19.1	56,527.8
Hours worked (monthly)	hours	55,324	190.2	49.3	4.3	292.4
Wage-push pressure		55,324	-81.9	135.7	-5,506.2	100.0

Notes: KRW, Korean won; Obs., the number of observations; and Std., standard deviation.

As argued in Zavodny (2000), Stewart and Swaffield (2008), and Harasztosi and Lindner (2019), when the effects on employment of an increase in the minimum wage are evaluated, the group that would be most affected by the increase is those employees whose current wage is below the new minimum wage. For example, workers earning below 7,530 Korean won per hour in 2017 would have had their hourly wage rate increased after the 2018 minimum wage legislation. As such, this paper defines the 'affected workers' (or 'treatment group') as those whose current wage is below the new minimum wage.

In contrast, the comparison group is defined based on the wage-push pressure (TI)

and the living wage (about 1.22 times the minimum wage, set by the government).[18] I partition the comparison group into the 'unaffected workers' (or 'control group'), defined as those whose current wage is between the new minimum wage and the living wage (i.e. $-22 \leq TI < 0$), and the 'non-target workers', defined as those whose current wage is larger than the living wage (i.e. $TI < -22$). This classification implies that in response to increases in the minimum wage, the unaffected workers are not directly affected, but they still have characteristics similar to those of the affected workers. The non-target workers are not directly affected by increases in the minimum wage, and they have no characteristics similar to those of the affected workers. Table C.1 shows that the discriptive statistics for the socio-economic characteristics of the affected workers is similar to those of the unaffected workers, unlike those of the non-target workers.

3.5 Empirical Analysis

The longitudinal data over a relatively long period, combined with the very large shock in the minimum wage, allows us to examine the two hypotheses presented in Section 3. The first one is the non-linear relationship between the increase in the minimum wage and the hours worked, and the second one is the different employment effects on between modest increases in the minimum wage over a decade under business-friendly governments and a sharp increase under the recent labour-friendly government.

3.5.1 Non-linear effects on employment

To capture the non-linear effects of increases in the minimum wage on hours worked, I estimate the following equation:

$$e_{i,t} = \alpha + \beta_1 e_{i,t-1} + \beta_2 MW_{i,t} + \beta_3 MW_{i,t}^2 + \beta_4 TI_{i,t} + X_{i,t}' \gamma + \lambda_i + \delta_t + \varepsilon_{i,t} \tag{3.6}$$

[18]Neumark and Wascher (2008) also show that the increase in the minimum wage has ripple effects on wages up to approximately 1.2–1.3 times the minimum wage. Admittedly, there is a trade-off to this choice of the upper limit of the unaffected workers – widening the range of the unaffected workers improves sample sizes and hence the precision of estimation, but it may lessen the similarity between the affected workers and the unaffected workers.

where $e_{i,t}$ is the hours worked per month, as used in Zavodny (2000) and Stewart and Swaffield (2008); $MW_{i,t}$ is the growth rate of the real minimum wage; $TI_{i,t}$ is the wage-push pressure on an employer; $X_{i,t}$ is a vector of control variables including gender, marital status, age, education attainment, employment status, firm size and industry classification; λ_i is individual-fixed effects, δ_t is time-fixed effects, and $\varepsilon_{i,t}$ are the error terms.

To control for the unobserved heterogeneity, after using the generalised least square (GLS) as a reference (column 1), I use a two-step GMM estimator as the main regression model (columns 2–5).[19] The variables of interest are β_2, β_3 and β_4. If $\beta_2 \geq 0$ and $\beta_3 < 0$, then they imply non-linear (i.e. inverted-U) employment effects. Furthermore, if $\beta_4 > 0$, a larger difference between the current wage and the upcoming minimum wage is associated with longer hours worked.

Table 3.4 shows the estimation results. Columns 1 and 2 suggest that for the affected workers, whose current wage was lower than the new minimum wage, a higher minimum wage is associated with longer hours worked, unlike what would be predicted by models assuming a competitive labour market. For instance, in column 2, with the unobserved heterogeneity controlled, the increased minimum wage is shown to have a positive and statistically significant effect on the hours worked; a 10 percentage point increase in the minimum wage is associated with an increase of approximately 1.3-hour in monthly hours worked. However, this result can be misleading if we disregard the potential non-linear relationship between the two variables.

In this respect, column 3 explores the hypothesis of the 'excessive increase in the minimum wage'. It proposes that there is a threshold above which the increase in the minimum wage no longer has a positive effect on the hours worked. It does this by adding a quadratic term of the increase in the minimum wage (MW^2). If the coefficient associated with the linear term (MW) is positive and the quadratic term in this variable (MW^2) is negative, there is an inverted-U relationship between the increase in the minimum wage and the hours worked for low-paid workers. Our estimates have confirmed this: the increase in the minimum wage begins to have a negative effect on the hours worked when the rate of the increases in the minimum wage reaches approximately 5.5% with an interval of 3.8% to 8.0% at the 90% confidence level. This outcome shows that the 1–7% increases in real minimum wages in South Korea during the period 2009–2017 (except for -0.2% in 2010) would have had a positive effect on the hours worked by low-paid workers, but the 14.9% increase in 2018 would have had a significantly negative effect on their hours worked.

[19] I excluded the fixed-effects estimator to avoid the possibility of Nichell bias Nickell (1981) in the dynamic panel model. More specifically, according to Baltagi and Baltagi (2001), with the presence of lagged dependent variables in the right-hand-side equation, the fixed-effects estimator would be biased even with a large sample, and the bias approaches zero as the time dimension of the panel gets larger.

Table 3.4 Effects of increases in the minimum wage on hours worked

	Affected workers			Unaffected workers	Non-target
	(1) GLS	(2) GMM	(3) GMM	(4) GMM	(5) GMM
Hours worked (t-1)	0.607***	0.301***	0.920***	0.679***	0.885***
	(0.019)	(0.043)	(0.070)	(0.071)	(0.023)
MW	0.160***	0.131***	1.489***	1.271***	2.045***
	(0.027)	(0.029)	(0.120)	(0.206)	(0.063)
MW-squared			-0.136***	-0.130***	-0.293***
			(0.026)	(0.040)	(0.031)
TI	0.607***	0.558***	0.156***	0.068	-0.011**
	(0.038)	(0.073)	(0.047)	(0.050)	(0.005)
Fixed effects	No	Yes	Yes	Yes	Yes
Time effects	Yes	Yes	Yes	Yes	Yes
Observations	7,219	7,219	7,219	4,254	25,972
dY/dMW = 0			5.5	4.9	3.5
CI (90%)			3.8–8.0	4.1–5.7	3.1–3.8

Notes: Gender, marital status, age, education attainment, employment status, firm size and industry classification were controlled. Affected workers are those whose current wage is less than the new minimum wage. Unaffected workers are those whose current wage is between the new minimum wage and the living wage. Non-target workers are those whose current wage is greater than the living wage. CI implies a confidence interval. Robust standard errors in parentheses. ***, $p < 0.01$; **, $p < 0.05$, and *, $p < 0.1$.

Meanwhile, columns 4 and 5 focus on the unaffected workers, whose current wage was between the upcoming minimum wage and the living wage (i.e. 1.22 times the new minimum wage) and the non-target workers whose current wage was higher than the living wage. Again, the linear and the quadratic terms indicate that the marginal effect of the increases in the minimum wage becomes negative at 4.9% for the unaffected workers and at 3.5% for the non-target workers. The unaffected workers and the non-target workers also showed a non-linear relationship between minimum wages and the hours worked, but their curvatures were much steeper than one for the affected workers. This implies that, when the minimum wage rises sharply, employers would reduce the hours worked for the costliest employees, perhaps because their labour costs per unit are relatively high. Typically, employers would reduce the amounts paid out for holidays and for overtime for those workers by reducing their hours worked.[20] In addition, the estimate of the wage-push pressure on an employer for the affected workers show that a larger difference between the current wage and the new

[20]Especially in the Korean context, employers may first attempt to reduce the hours worked on overtime because the extra cost for overtime is 50% higher than the standard rate.

minimum wage is associated with longer hours worked, and the estimate of the wage-push pressure on an employer for the unaffected workers is about half of that for the affected workers.

Overall, employers increased the hours worked for all employees against the modest increases in minimum wages (approximately 3.5%). For growth rates of the minimum wage between 3.5% and 5.5%, the hours worked by the affected workers increased at the expense of the hours worked by the unaffected workers and the non-target workers. However, above the tipping point, 5.5% in my estimation, employers would reduce the hours worked for all employees.

Next, to check whether the increased minimum wage has been passed on to employee earnings as labour costs for employers, I use total real wages per month (including such things as overtime pay and meal allowance) instead of the hours worked as the dependent variable in equation (3.6).

Table 3.5 shows the estimation results. Column 1 indicates that a higher minimum wage is associated with higher earnings for a low-paid worker, but column 2, with the unobserved heterogeneity controlled, shows that it is not statistically significant. In column 3, which adds a quadratic term of the minimum wage, the coefficients associated with the linear term and the quadratic term are not statistically significant. This provides little evidence of a non-linear relationship between increases in the minimum wage and real wages for the affected workers. Moreover, there is little support that threshold-type non-linearities are also present for the unaffected workers and the non-target workers, as shown in columns 4 and 5.

The estimation outcomes have the following implications. First, an increase in the real minimum wage has a positive effect on the hours worked of low-paid workers up to the tipping point, which is 5.5% in my estimation. Second, even though the growth rate of the minimum wage hits the tipping point, an employer will attempt to absorb the minimum wage shock by reducing the hours worked to an extent that the labour cost per employee (i.e. hours worked times hourly wage rate) remains unchanged. This is the reason that, even though there is a large shock in the minimum wage, there is no discernible impact on the earnings of low-paid workers. Third, the higher the wage-push pressure is, the more benefit low-paid workers would have in terms of hours worked as a result of the same increase in the minimum wage.

However, those responses to the increased minimum wage are averages. There would be heterogeneous employment effects among low-paid employees because some workers have more experience, better skills, more education or stricter employment protection. The extent to which employers' labour demand for workers may change as a result of a rise in the minimum wage differs across all these characteristics. In this regard, Table 3.6 reports the

Table 3.5 Effects of increases in the minimum wage on real wages

	Affected workers			Unaffected workers	Non-target
	(1) GLS	(2) GMM	(3) GMM	(4) GMM	(5) GMM
Real wages (t-1)	0.765***	-0.682***	-0.472***	-0.164	0.246***
	(0.029)	(0.124)	(0.168)	(0.131)	(0.089)
MW	0.124	0.119	3.057	1.826	-0.847
	(0.098)	(0.115)	(2.455)	(1.530)	(5.512)
MW-squared			-0.299	-0.210	-0.318
			(0.234)	(0.408)	(1.936)
TI	7.127***	0.539	1.147	-0.166	-1.202
	(0.423)	(0.925)	(1.888)	(0.770)	(0.774)
Fixed effects	No	Yes	Yes	Yes	Yes
Time effects	Yes	Yes	Yes	Yes	Yes
Observations	7,140	7,140	7,140	4,254	25,972

Notes: Gender, marital status, age, education attainment, employment status, firm size and industry classification were controlled. Affected workers are those whose current wage is less than the new minimum wage. Unaffected workers are those whose current wage is between the new minimum wage and the living wage. Non-target workers are those whose current wage is greater than the living wage. Robust standard errors in parentheses. ***, $p < 0.01$; **, $p < 0.05$, and *, $p < 0.1$.

estimation results for the affected workers disaggregated by employment status (columns 1 and 2), firm size (columns 3 and 4) and age (columns 5 and 6).

As noted by the relevant literature (e.g. Dickens et al., 2015; Giuliano, 2013), certain vulnerable groups may be susceptible to the adverse effects of increases in the minimum wage. In particular, non-regular workers may be susceptible to cuts in the hours worked, given their more precarious employment status. Columns 1 and 2 indicate that this prediction is true. The linear and the quadratic terms for non-regular workers and regular workers indicate that the marginal effect of increases in the minimum wage becomes negative at 5.2% and 7.6%, respectively. That is, even though both groups show a non-linear relationship between increases in the minimum wage and hours worked, the non-regular workers' curvature was much steeper than that of the regular workers.

In contrast, columns 3 and 4 show that workers whose firm size is less than or equal to nine are less susceptible to cuts in hours worked because of the increase in the minimum wage, compared with workers whose firm size is greater than nine. The linear and the quadratic terms for the former and the latter indicate that the marginal effect of increases in the minimum wage becomes negative at 7.7% and 5.8%, respectively. This indicates that workers whose firm size is relatively large seems to be more substitutable by spreading working hours across colleagues or reducing overtime work.

Last, columns 5 and 6 show that teenage and young adult workers (aged 16–29 years) are more susceptible than mature workers (aged 30 years and over) to cuts in hours worked because of the shock of a large increase in the minimum wage. This is because of their more precarious employment status, in line with the literature (e.g. Clemens, 2015). The linear and the quadratic terms for teenage and young adult workers and mature workers indicate that the marginal effect of increases in the minimum wage becomes negative at 5.8% and 10.2%, respectively. That is, even though both groups show a non-linear relationship between the minimum wage and the hours worked, the curvature for the teenage and young adult workers was steeper than that of the mature workers.

Table 3.6 Heterogeneous effects of increases in the minimum wage on hours worked

	Employment status		Firm size		Age	
	(1) Temp.	(2) Regular	(3) 1-9	(4) 10+	(5) 16-29	(6) 30+
Hours worked (t-1)	0.743***	0.606***	0.721***	0.866***	0.744***	0.726***
	(0.082)	(0.118)	(0.123)	(0.078)	(0.104)	(0.101)
MW	1.417***	0.550**	1.000***	1.416***	0.758***	1.111***
	(0.235)	(0.263)	(0.306)	(0.235)	(0.254)	(0.213)
MW-squared	-0.136***	-0.036	-0.065*	-0.123***	-0.095***	-0.037
	(0.030)	(0.029)	(0.037)	(0.025)	(0.026)	(0.024)
TI	-0.074	0.258*	-0.007	-0.198	0.037	-0.428**
	(0.112)	(0.150)	(0.171)	(0.128)	(0.134)	(0.209)
Observations	3,368	3,851	3,725	3,494	806	6,413
dY/dMW=0	5.2	7.6	7.7	5.8	5.8	10.2
CI (90%)	4.5–6.1	5.4–10.5	5.6–10.4	4.6–7.3	3.9–7.9	8.8–11.5

Notes: The regressions were conducted on the two-step GMM estimator for the affected workers, those whose current wage is less than the new minimum wage. Gender, marital status, age, education attainment, employment status, firm size and industry classification were controlled. CI implies a confidence interval. Robust standard errors in parentheses. ***, $p < 0.01$; **, $p < 0.05$, and *, $p < 0.1$.

3.5.2 Employment effects of a large minimum wage increase

Next, if there is a non-linear relationship between increases in the minimum wage and hours worked, by exploiting the large shock in the minimum wage in 2018, we can find the different employment effects generated (i) between the modest increases between 2009 and 2017 under business-friendly governments and the sharp increase during 2018 under the labour-friendly government, and (ii) between the affected workers and the unaffected workers in 2018. These situations are considered in turn.

Modest shock vs. large shock. Following Stewart and Swaffield (2008), I estimate the effect of increases in the minimum wage on hours worked using a difference-in-difference estimator.[21] This involves calculating the change in hours worked by the treatment group (i.e. those whose current wage is below the new minimum wage) in the period following the large shock in the minimum wage, compared with the period preceding the shock, and subtracting it from the change in hours worked by the comparison group (i.e. those whose current wage is above the new minimum wage).

For our analysis of the effect of the increase in the minimum wage on the hours worked, I pool yearly cross-sections of data, with each year containing both the treatment group and the comparison group. The treatment effect can then be estimated using standard linear regression techniques. With two time periods, namely 2009–17 and 2018, the difference-in-difference estimator can be implemented such that:

$$e_{i,t} = \beta_0 + \beta_1 e_{i,t-1} + \beta_2 T_i + \beta_3 Post_t + \beta_4 T_i \cdot Post_t + \beta_5 TI_i + X'_{i,t}\gamma + \varepsilon_{i,t} \tag{3.7}$$

where $e_{i,t}$ is the hours worked per month; $Post_t$ is a dummy variable which equals one for observations in 2018 and zero for observations between 2009 and 2017; T_i is a treatment dummy variable, which equals one if worker i is in the treatment group, and zero if she is in

[21]To formalise, the difference-in-difference estimator is given by:

$$\theta = (\bar{y}^{(1)}_{t=2} - \bar{y}^{(1)}_{t=1}) - (\bar{y}^{(2)}_{t=2} - \bar{y}^{(2)}_{t=1})$$

where $\bar{y}^{(g)}_t$ denotes the average of y_{it} over i in group g. The validity of this estimator requires a crucial assumption. Suppose that in the absence of an increase in the minimum wage, the change in hours worked can be decomposed into two components – the first fixed over time and the second common across groups. This assumes that without the increase in the minimum wage, the difference in the average change in the hours of work between groups would be the same in each time period. That is to say, the time-paths of changes in hours worked would be the same for each group. Accordingly, if we observe a reduction in the hours worked for the affected workers and no reduction (or even an increase) in the hours worked for the unaffected workers, then I attribute the decline in the hours worked by low-paid workers to the large shock in the minimum wage in 2018.

the comparison group; and $TI_{i,t}$ is treatment intensity.

The variable of interest is the coefficient on the interaction term $(T_i \cdot Post_t)$. If $\beta_4 < 0$, the sharp increase in the minimum wage in 2018 is estimated to show an adverse employment effect because that rate of the increase in the minimum wage is far above the tipping point (5.5% in my estimation). I also include a vector for additional covariates, $X_{i,t}$, such as gender, marital status, age, education attainment, employment status, firm size and industry classification. Including additional covariates in a difference-in-difference model controls for any potential compositional changes in the treatment and comparison groups over time.[22]

Table 3.7 presents the estimation results. Column 1 indicates that the large increase in the minimum wage in 2018 had a negative and statistically significant effect on the hours worked – a 2.2–hour reduction in the hours worked per month for the treatment group. Columns 2 and 3 show that the results for regular workers were more pronounced than for non-regular workers. The 2018 increase in the minimum wage led to a reduction of approximately 2.6 hours worked per month for the affected workers on regular contracts. Moreover, there was an increase of approximately 3 hours per month for the treatment group on temporary contracts though this is not statistically significant. This result can partly explain the puzzle that – in response to the large increase in the minimum wage in South Korea – the employment rates for all age groups were increasing. That is, employers tried to manage total payroll hours by reducing costly overtime remuneration of regular workers and instead increasing working hours of non-regular workers under South Korea's strong employment protection for temporary workers (as already seen in Figure C.4). Columns 4-7 reveal that working in bigger firms and being younger is associated with shorter hours worked at the expense of the other workers.

Overall, on average, the large shock in the minimum wage shortened the hours worked for regular workers, workers whose firm size is relatively large, and teenage and young adult workers in the treatment group – the main target of the minimum wage policy. This is partly because of the strong employment protection for temporary workers and small businesses under the left-wing government.

Affected vs. Unaffected. To compare the employment effects on the affected workers (say, those who have a current wage below the new minimum wage) to the effects on the unaffected workers (say, those who have a current wage between the new minimum wage and the living wage) after the large increase in the minimum wage in 2018, I implemented RD designs using the 2018 sample only.

[22] Even if the treatment is independent of these covariates, it is common practice to include them in order to improve the precision of the difference-in-difference estimate.

Table 3.7 Employment effects from modest increases versus the sharp increase

	All	Employment status		Firm size		Age	
	(1)	(2) Temp.	(3) Reg.	(4) 1-9	(5) 10+	(6) 15-29	(7) 30+
Hours (t-1)	0.550***	0.618***	0.498***	0.582***	0.521***	0.497***	0.556***
	(0.005)	(0.009)	(0.006)	(0.009)	(0.007)	(0.018)	(0.006)
T	20.172***	25.588***	16.528***	18.011***	18.573***	10.979***	21.068***
	(0.742)	(2.105)	(0.828)	(1.845)	(0.944)	(2.365)	(0.797)
Post	-10.230***	-15.165***	-9.340***	-11.437***	-9.985***	-11.213***	-10.057***
	(0.569)	(2.008)	(0.560)	(1.280)	(0.626)	(1.738)	(0.603)
T · Post	-2.178*	3.034	-2.601*	0.291	-3.621**	-3.373	-1.931
	(1.292)	(2.847)	(1.463)	(2.136)	(1.667)	(3.351)	(1.406)
TI	0.064***	0.095***	0.055***	0.118***	0.057***	0.139***	0.062***
	(0.005)	(0.020)	(0.005)	(0.019)	(0.005)	(0.026)	(0.005)
Constant	43.006***	13.949	85.604***	43.573***	72.018***	88.909***	65.700***
	(12.518)	(16.181)	(9.279)	(16.432)	(13.448)	(19.059)	(7.059)
Observations	37,445	7,980	29,465	11,479	25,966	3,910	32,678
R-squared	0.49	0.56	0.42	0.55	0.46	0.45	0.50

Notes: Dependent variable is monthly hours worked. Gender, marital status, age, education attainment, employment status, firm size and industry classification were controlled. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

The RD design has emerged in recent decades as one of the most credible non-experimental research strategies to evaluate causal effects (Cattaneo et al., 2020b). The distinctive feature of this framework is that all units receive a score, and a treatment is assigned to those whose score exceeds a *known* cut-off, and no treatment is assigned to those whose score is below the cut-off. Under the assumption that the units' observed and unobserved characteristics do not change abruptly at the cut-off, the change in treatment status induced by the discontinuous treatment assignment rule can be used to study different treatment effects on our outcomes of interest (see Cattaneo et al. (2019) for the details).[23] Accordingly, in the RD design, we can get the average treatment effects by comparing treated units that are *slightly above* the

[23] To formalise, let us assume that there are n units, each unit i has a score X_i, and c is a known cut-off. Units with $X_i \geq c$ are assigned to the treatment condition, while $X_i < c$ are assigned to the control condition. This treatment assignment (T_i) is defined as $T_i = \mathbb{1}(X_i \geq c)$, where $\mathbb{1}(\cdot)$ is the indicator function. This implies that the probability of treatment assignment as a function of the score changes discontinuously at the cut-off. To see the average treatment effect at a given value of the score, we imagine having units with a score exactly equal to c, and units with score barely below c (i.e., with score $c - \varepsilon$ for a small and positive ε). The former units would receive treatment, and the latter would be control. The average treatment effect (ρ) is then formally defined as:

$$\rho \equiv \mathbb{E}[Y_i(1) - Y_i(0) \mid X_i = c] = lim_{x \downarrow c}\mathbb{E}[Y_i \mid X_i = x] - lim_{x \uparrow c}\mathbb{E}[Y_i \mid X_i = x].$$

cut-off to control units that are *slightly below* it because treated and control units in a small neighbourhood around the cut-off are comparable in the sense of having similar observed and unobserved characteristics. Thus, observing the outcomes of units just below the cut-off provides a valid measure of the average outcome that treated units just above the cut-off would have had if they had not received the treatment.[24]

I have used this method to assess the impact of the large increase in the minimum wage on the hours worked of low-paid workers because the increase in the minimum wage itself introduces randomness by assigning samples into either the control or the treatment group as if it had come from a purely randomised experiment. According to Cattaneo et al. (2020a), it requires two assumptions. First one is excludability assumption that nothing else changes at the cut-off wage that may affect the outcome. Second one is the exogeneity assumption that the cut-off wage was not chosen to purposefully include units near the cut-off. In this regard, the setting in this study is valid for the study of the minimum wage because the cut-off is determined solely by the minimum wage committee at the time when the rate of the minimum wage is set. The logic is that treatment and control groups are likely to be similar to each other at near the cut-off wage except for the fact that one received the treatment and the other did not.[25]

In this sense, the wage-push pressure on an employer will be used as the score, and I then compare the hours worked by workers with wages on either side of the cut-off wages that are likely to cause substantially different cost-burdens for the employers. Arguably, the characteristics of workers on either side of the cut-off wages are very similar, as already seen in Table C.1. Therefore, the main difference between them is the wage-push pressure on an employer. In this sense, by comparing workers who are earning just above the cut-off wage with those just below this wage, we can estimate the employment effect of the 2018 minimum wage shock.

First, I estimate equation (3.8) using ordinary least squares (OLS) with global polynomials (the so-called 'global RD designs') such that:

$$\Delta e_i = \alpha_0 + \alpha_1 T_i + \Sigma_{p=1}^{k} \alpha_{p+1} TI_i^p + \Sigma_{q=1}^{k} \alpha_{k+q+1} T_i \cdot TI_i^q + \varepsilon_i \tag{3.8}$$

[24] The RD parameter is often referred to as a 'local' average treatment effect, because it is informative of the effect of the treatment only for units whose value of the score is in a local neighbourhood of the cut-off. Since this limits the external validity of the RD parameter, a growing amount of literature is analysing how to extrapolate treatment effects in RD designs (e.g. Angrist and Rokkanen, 2015; Bertanha and Imbens, 2020).

[25] Unlike an instrumental variables approach, the RD Design does not require the minimum wage to be highly correlated with the outcome variable. Unlike a fixed-effects approach, the RD Design also does not require unmeasured individual characteristics that were assumed fixed to not change over time.

where Δe_i is the change in the monthly hours worked from pre- to post-increase period of worker i; T_i is the assignment of the worker (i.e. $T_i = 1$ if i is exposed to treatment); TI_i^k is the wage-push pressure on the employer of worker i on the kth polynomial order; and ε_i are the error terms.

In Figure 3.4, the outcome of the affected workers (i.e. treatment group) and the unaffected workers (i.e. control group) is graphed around the cut-off – the left side represents the unaffected workers and the right side represents the affected workers. The vertical distance at the cut-off implies the average treatment effect in response to the increase in the 2018 minimum wage, assuming that these groups would be identical except for their treatment status at the cut-off. Each panel represents the outcome using linear, quadratic, cubic and reciprocal polynomial, respectively. Visual inspection of the graphs enables us to see if there is a discontinuity in the outcomes between the two groups. We can see that there is a discontinuity at the cut-off, regardless of the different polynomials used in the fitting. These results imply a negative impact of the large increase in the minimum wage in 2018 on the hours worked. This is in line with our previous results in Section 5.1.

To check the significance of the discontinuity and to estimate the effect of the treatment, the OLS regressions are run. Table 3.8 presents the results of the global RD designs, confirming that the 14.9% rise in the real minimum wage had a negative effect on the hours worked of the affected workers by, on average, approximately -4.8 to -10.9 hours per month. That is, we can see that the growth rate of the minimum wage in 2018 seems to hit the tipping point that curbs the negative effects on employment.

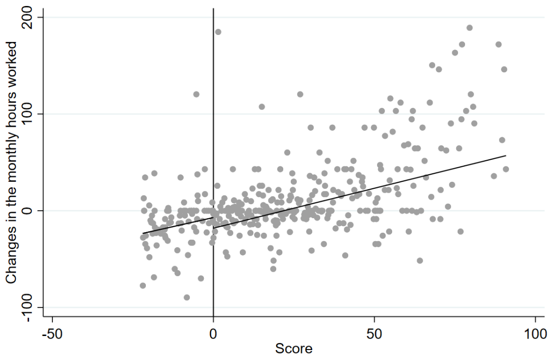
Table 3.8 Results of global RD designs

	Linear	Quadratic	Cubic	Reciprocal
RD effect	-10.934***	-4.757	-5.000	-8.651
	(3.486)	(4.786)	(6.067)	(7.027)
Observations	1,611	1,611	1,611	1,611
F-statistics	60.29	39.11	34.04	34.01
R-squared	0.17	0.19	0.19	0.19

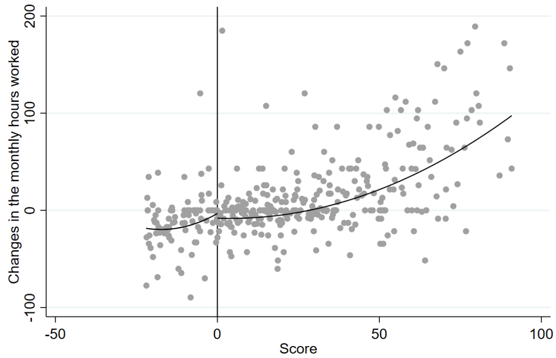
Notes: Robust standard errors in parentheses. ***, $p < 0.01$; **, $p < 0.05$, and *, $p < 0.1$

However, the RD design using global high-order polynomial approximations can cause misspecification bias, such as noisy estimates, and sensitivity to the degree of the polynomial (Gelman and Imbens, 2019). Therefore, I alternatively used the RD design relying on local low-order polynomial approximations (i.e. local linear or local quadratic) with selected bandwidths on both sides of the cut-off (the so-called ‘local RD designs’), as discussed

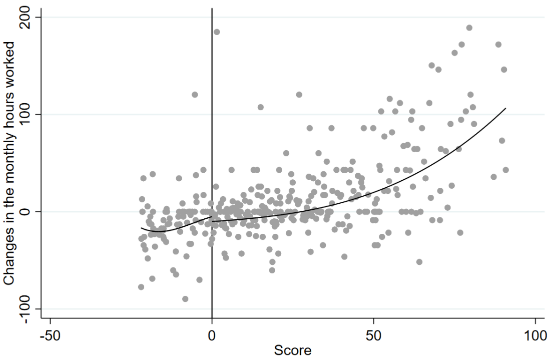
Fig. 3.4 Plot for global RD designs



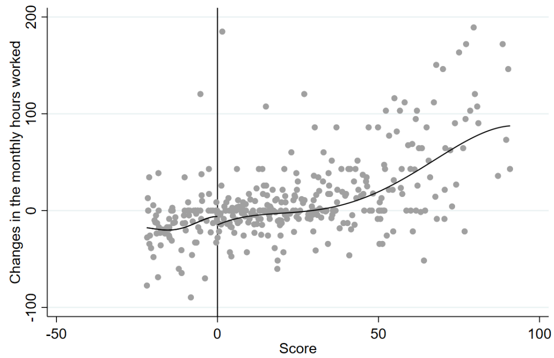
(a) Linear



(b) Quadratic



(c) Cubic



(d) Reciprocal

Note: The solid lines are estimated using ordinary least squares with global polynomials

by Hahn et al. (2001) and Calonico et al. (2014b).[26] Selecting the bandwidth around the cut-off in which to estimate the effect is a crucial step in the RD analysis, as the results and conclusions are typically sensitive to this choice (Cattaneo et al., 2020a). So, first, I choose a bandwidth based on intuition (or prior knowledge) about the particular context (the so-called '*ad-hoc* bandwidth'). I adjusted the bandwidth down to the growth rate of the real minimum wage in 2018 (i.e. bandwidth is set to 15). However, since it could lack objectivity (or a rigorous justification), second, I use the mean squared error criterion (the so-called 'MSE-optimal bandwidth'), a systemic data-driven criterion. (See Calonico et al. (2014a) and Cattaneo and Vazquez-Bare (2016) for an overview of bandwidth selection methods

[26]Despite the implementation and algebraic similarities between the OLS methods and local polynomial methods, there is a crucial difference: OLS methods assume that the polynomial used for estimation is the true form of the function, while local polynomial methods see it as just an approximation to an *unknown* regression function (Cattaneo et al., 2020b).

in the RD designs.) Furthermore, since local polynomial approximations usually include a weighting scheme that places more weight on observations that are closer to the cut-off (the so-called 'kernel function'), I use uniform weights as a kernel function as usual. Once the bandwidth and the kernel function have been chosen, the implementation of local RD designs reduces to simply fitting linear or quadratic regressions using weighted least-squares (Cattaneo et al., 2019).

Figure 3.5 shows a discontinuity in the outcomes between the two groups, regardless of the polynomial order and the bandwidth. A stark jump is seen at the cut-off, where the hours worked abruptly decreases as the score crosses the cut-off. This indicates that those whose current wage is barely below the new minimum wage have less hours worked compared to those whose current wage is barely above the new minimum wage.

Next, I analyse this effect formally. I estimate RD effects using local polynomial methods with a linear polynomial and an ad-hoc bandwidth (column 1), a quadratic polynomial and an ad-hoc bandwidth (column 2), a linear polynomial and a MSE-optimal bandwidth (column 3), and a quadratic polynomial and a MSE-optimal bandwidth (column 4). As we can see in Table 3.9, the MSE-optimal bandwidths are estimated to be around 7.3 percentage points for a linear approximation and around 13.2 percentage points for a quadratic approximation, and the RD local-polynomial point estimates are -4.0 – -10.3. This shows that the large increase in the minimum wage in 2018 was harmful for low-wage workers, with a reduction of approximately 4.0–10.3 hours in the hours worked per month. These results are similar to the results of the global RD designs.

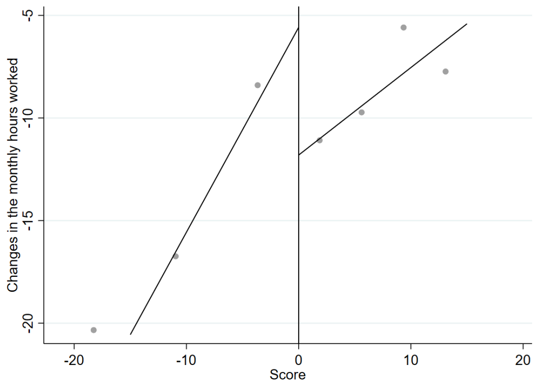
Table 3.9 Results of local RD designs

	Ad-hoc bandwidth		MSE-optimal bandwidth	
	(1) Linear	(2) Quadratic	(3) Linear	(4) Quadratic
RD effect	-10.319^{**}	-4.015	-9.836	-7.027
	(4.590)	(7.775)	(7.471)	(8.911)
Bandwidth	15.0	15.0	7.3	13.2
Observations	680	680	378	603

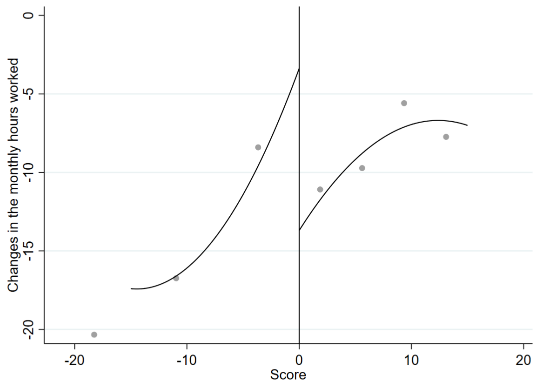
Notes: Robust standard errors in parentheses. ***, $p < 0.01$; **, $p < 0.05$, and *, $p < 0.1$

However, because these results are averages, we still cannot rule out the possibility that the effects on employment of the increase in the minimum wage are heterogeneous among employees. In particular, if the growth rate of the minimum wage is unexpectedly high, employment-related decisions will differ depending on the different situation for each employer. Thus, the impacts of the increase in the minimum wage on the employment

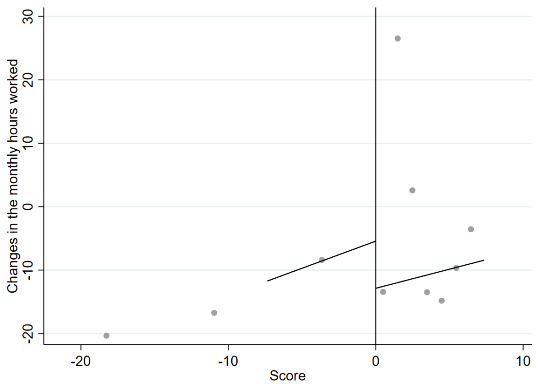
Fig. 3.5 Plot for local RD designs



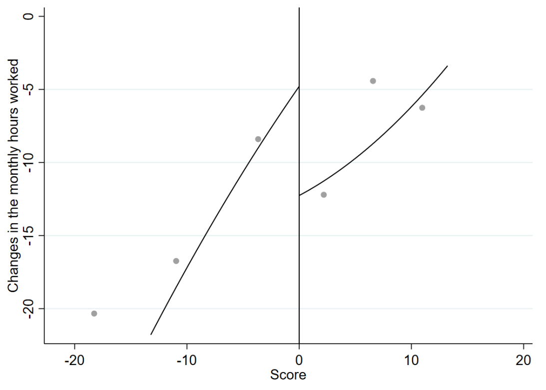
(a) Linear with ad-hoc bandwidth



(b) Quadratic with ad-hoc bandwidth



(c) Linear with MSE-optimal bandwidth



(d) Quadratic with MSE-optimal bandwidth

Notes: The solid lines are estimated using local polynomial methods with uniform weights. The number of bins is selected using the IMSE-optimal even-spaced method.

of specific workers may differ markedly. To check the effect of heterogeneity, this study compares the effects on employment of the 2018 minimum wage shock by employment status, firm size and age.

Table 3.10 shows the estimation results using local RD designs with a linear approximation. In particular, the impact on hours worked of a large shock in the minimum wage is significant for regular workers, workers whose firm size is relatively large and teenage and young adult workers (aged 16-29 years). These results imply that under the single minimum wage system, like the one in South Korea, the impacts of a sharp increase in the minimum wage on the employment of low-paid workers can vary based on the types of workers.

Table 3.10 Heterogeneous RD effects of the large increase in the minimum wage

	Employment status		Firm size		Age	
	(1) Temp.	(2) Regular	(3) 1-9	(4) 10+	(5) 16-29	(6) 30+
RD effect	-5.519	-10.644**	4.097	-26.966***	-12.116**	-2.318
	(10.110)	(4.811)	(5.978)	(6.812)	(5.069)	(8.845)
Observations	455	225	258	422	106	574

Notes: Bandwidth is set to 15 with a linear approximation and uniform weights. Robust standard errors in parentheses. ***, $p < 0.01$; **, $p < 0.05$, and *, $p < 0.1$

3.5.3 Robustness check

The key component in the RD designs lies in accurately capturing the treatment effect (i.e. a 16.4% rise in the minimum wage in my case) if the treatment status changes near the cut-off (Imbens and Lemieux, 2008). In this respect, I check the robustness of the RD designs that I have used in the previous section as follows.

First, I assumed that compliance with the assignment of treatment (commonly known as the 'sharp RD design') is perfect. However, since compliance with treatment coud be imperfect, I use a fuzzy RD design to check robustness (see Cattaneo et al. (2020a) for the details). In the fuzzy RD design, the probability of receiving the treatment does not have to change abruptly from zero to one. It allows for smaller jumps in the probability of assignment to the treatment at the cut-off such that:

$$lim_{x \downarrow c} Pr[Y_i = 1 \mid X_i = c] \neq lim_{x \uparrow c} Pr[Y_i = 1 \mid X_i = c]. \tag{3.9}$$

In this sense, the fuzzy RD design involves a concept of estimating instrumental variables.[27] In Table C.2, we can confirm that even using the fuzzy RD design, there is a negative impact of the large increase in the minimum wage on the hours worked of low-wage workers.

Second, instead of relying on continuity of regression functions and their approximations,

[27]To formalise, the ratio of the jump in the regression of the outcome on the covariate to the jump in the regression of the treatment indicator on the covariate is interpreted as the average treatment effect such that:

$$\rho \equiv \frac{lim_{x \downarrow c} \mathbb{E}[Y_i \mid X_i = x] - lim_{x \uparrow c} \mathbb{E}[Y_i \mid X_i = x]}{lim_{x \downarrow c} \mathbb{E}[T_i \mid X_i = x] - lim_{x \uparrow c} \mathbb{E}[T_i \mid X_i = x]}.$$

I conduct a local randomisation framework that assumes the treatment of interest is 'as-if' randomly assigned in a very small region around the cut-off (see Cattaneo et al. (2020b) for the details). The intuition is that, if units either have no knowledge of the cut-off or have no ability to precisely manipulate their own score, units whose scores are close enough to the cut-off will have as much possibility of being barely above the cut-off as barely below like it. If this is true, close enough to the cut-off, the RD design may create experimental-like variation in the assignment of treatment. For this, we must first choose the window around the cut-off where the assumption of local randomisation appears plausible if such a window exists. I set our window to one which has 19 and 16 observations on each side of the cut-off, and the number of repetitions is set to a thousand.

In Table C.2, I present my inference results, reporting Fisherian inference. The RD effect is -17.932, with a (Fisherian) p-value of approximately .078. This means that the null hypothesis of no average effect is rejected at 10% level. The fact that the point estimate continues to be negative and that the p-value is 8% and below suggests that the estimated RD effects in the previous section are broadly robust to a local-randomisation assumption, as both approaches lead to similar conclusions. That is, even using a local randomisation method, we can confirm a negative effect on employment of low-paid workers from the 2018 increase in the minimum wage.

3.5.4 Are the reduced hours worked harmful for low-paid workers?

A question naturally arises from the empirical results in the previous sections: Are the reduced worked hours very harmful for South Korea's low-paid workers? My answer is that it depends on whether the worker places more value on leisure or earnings.

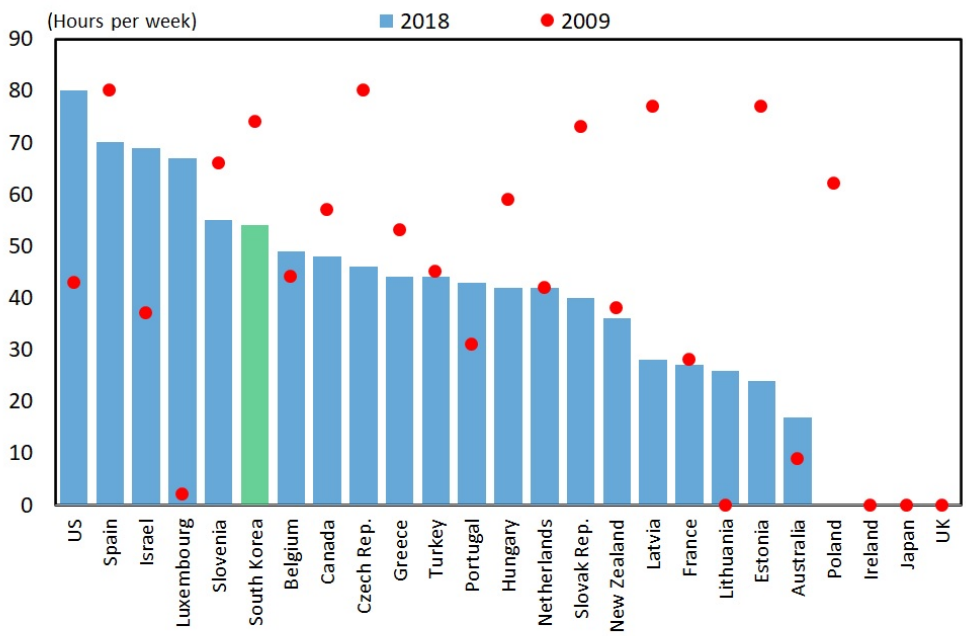
If the increased minimum wage and the reduced hours worked nearly offset each other, the overall earnings of low-paid workers might be unchanged. Nevertheless, raising the minimum wage acted as a financial incentive for employers to avoid the long working hours for which South Korea is notorious.[28] Moreover, the reduced hours worked could have acted to remove exploitation rents (i.e. the gap between labour productivity and real wages) for low-paid workers, gaps that have grown since the 1997–98 Asian financial crisis (as seen in Figure 3.3). In this respect, in South Korea, proponents of raising the minimum wage sharply

[28] In reality, what holds many Korean workers at work so late is the belief that they cannot leave on time because of the perception it gives to their superiors and the belief that the boss wants everyone to stay late and sacrifice as much as they do. Perception reigns supreme in a Korean office. Giving the impression of working hard by staying late is something all Koreans do. In fact, although a contract may state working hours are from 9am to 6pm, working from 8am to 9pm or later (and then receiving dinner drinks) is the norm in South Korea.

would be mostly workers who accept minimum wage jobs for the 'experience' rather than out of necessity because it allows them to attain a better work-life balance. Many of them are expected to be teenage and young adult workers in well-to-do families.

However, the incomes of full-time minimum-wage earners can be well below commonly used poverty lines (i.e. 50% of median net household income), which is the case in South Korea. In these cases, even working very long hours may not enable families to escape poverty as conventionally measured. For example, Figure 3.6 shows that in 2018 in Australia, a half-time minimum wage job can be sufficient to raise a family with two children out of poverty. Moreover, in Japan and the United Kingdom, out-of-work benefits provide income above the poverty threshold, even when no one in a two-parent family works. However, in 2009 in South Korea, a single full-time minimum wage job leaves a two-parent family with two children below the poverty line, so both parents must be employed to ensure that their children do not grow up in poverty. They must work for 74 hours a week to escape poverty. Even though the situation was much better in 2018 (with only 55 hours required to escape

Fig. 3.6 Hours at the minimum wage that must be worked to move above a poverty line



Notes: The data are estimated based on an one-earner couple with two children. The poverty line is 50% of each country's median net household income. Net incomes are calculated by subtracting incomes taxes and mandatory social or private contributions payable by workers, and adding family benefits and minimum-income and other means-tested benefits that are primarily income related and are typically accessible for low-income families.
Source: OECD.stat (2020)

poverty), the working hours required to escape poverty on a minimum wage are unrealistic for single parents. They would need better income support or wages significantly above the minimum wage to work their way out of poverty. In this respect, in South Korea, the opponents of the large increase in the minimum wage (i.e. proponents of modest increases in the minimum wage) would be mostly workers who accept minimum wage jobs for 'to make a living', such as mature adults with several children. They would pursue greater earnings regardless of working overtime or on holidays to escape poverty.

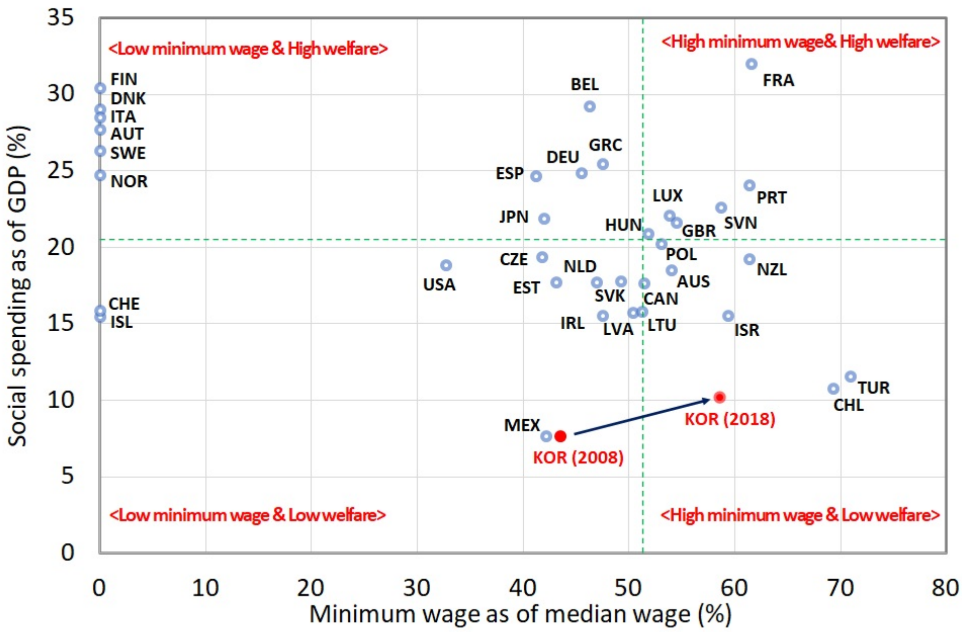
Overall, the minimum wage must be set at a level that prevents adverse effects on employment, with minimal intervention by the government in the bargaining process. It is advisable to maintain the increase in the minimum wage at a moderate level. Otherwise, a rapid increase in the minimum wage may do little to bolster low-paid workers' earnings. In this sense, a large increase in the minimum wage alone is not sufficient as a strategy to alleviate poverty, considering its non-linear and heterogeneous effects by employment status, firm size and age.

Figure 3.7 illustrates how South Korea – since the 2008 financial crisis – has been moving away from being a country of 'low minimum wage and low welfare' (like countries such as Mexico), to one of 'high minimum wage and low welfare' (like countries such as Turkey). It is interesting to note that most of the eight OECD countries that do not have a statutory minimum wage (the Nordic countries, Austria, Italy, and Switzerland) have high levels of social spending relative to GDP. It implies that in South Korea, coordinating the minimum wage policy and the welfare policy is important for alleviating the poverty of low-paid workers.

3.6 Conclusion

This paper has investigated the relationship between increases in the minimum wage and the hours worked, based on the bargained minimum wage model. It has shown that, at modest rates of increase, there is a positive relationship between the two. However, at a much higher rate of increase, the relationship becomes negative. The sudden regime change in South Korea in mid-2017 and the subsequent large rise in the minimum wage was used to investigate the non-linear effects of the increase in the minimum wage on the employment of low-paid workers. To demonstrate these effects, this study used two-step GMM, difference-in-differences, and RD designs, based on a longitudinal survey of labour markets and income activities during the period 2009–2018.

Fig. 3.7 Coordination of the minimum wage and social spending in OECD countries



Notes: The countries on the y-axis are countries without a statutory minimum wage, and the dotted lines imply the OECD averages.
Source: The author utilises the data from OECD.stat (2020)

The estimation results show that the 14.9% rise in the real minimum wage in 2018 was well above the tipping point, which is 5.5% in my estimation. This unexpected double-digit increase has led to a reduction in hours worked of low-paid workers, compared with the modest increases under business-friendly governments between 2009 and 2017. These results confirm our conjecture that an increase in the minimum wage beyond the threshold acts as a financial incentive for employers to reduce working hours. Accordingly, in response to the hike in the minimum wage, employers shortened operating hours and reduced overtime work. This hurt some low-paid workers by reducing their overall earnings, but it has had an unintended positive consequence of promoting better work-life balance in a country that is notorious for over-working.

This study is one of the first, I believe, to examine the non-linear effects of the minimum wage on employment at the intensive margins. However, it is limited in that it does not look at the long-term effects that increases in the minimum wage could have, such as labour productivity. For example, the rapid increase in the minimum wage could serve as a momentum for firms to promote capital investment, thus substituting capital for labour.

Evaluating the impact of raising the minimum wage on long-term employment is a topic for future research.

References

Aaronson, D. and French, E. (2007). Product Market Evidence on the Employment Effects of the Minimum Wage. *Journal of Labor Economics*, 25(1):167–200.

Acemoglu, D. and Robinson, J. A. (2000). Why Did the West Extend the Franchise? Democracy, Inequality, and Growth in Historical Perspective. *Quarterly Journal of Economics*, 115(4):1167–1199.

Acikgoz, O., Hagedorn, M., Holter, H., and Wang, Y. (2018). The Optimum Quantity of Capital and Debt. Discussion Paper 12952, Centre for Economic Policy Research.

Aghion, P., Akcigit, U., Bergeaud, A., Blundell, R., and Hemous, D. (2019). Innovation and Top Income Inequality. *Review of Economic Studies*, 86(1):1–45.

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

Aiyagari, S. R. (1994). Uninsured Idiosyncratic Risk and Aggregate Saving. *Quarterly Journal of Economics*, 109(3):659–684.

Alesina, A. and Perotti, R. (1996). Income distribution, political instability, and investment. *European Economic Review*, 40(6):1203–1228.

Alesina, A. and Rodrik, D. (1994). Distributive Politics and Economic Growth. *Quarterly Journal of Economics*, 109(2):465–490.

Almeida, H., Fos, V., and Kronlund, M. (2016). The real effects of share repurchases. *Journal of Financial Economics*, 119(1):168–185.

Alvaredo, F., Atkinson, A. B., Piketty, T., and Saez, E. (2013). The Top 1 Percent in International and Historical Perspective. *Journal of Economic Perspectives*, 27(3):3–20.

Alvaredo, F., Chancel, L., Piketty, T., Saez, E., and Zucman, G. (2018). The elephant curve of global inequality and growth. *AEA Papers and Proceedings*, 108:103–08.

Alvaredo, F., Chancel, L., Piketty, T., Saez, E., and Zucman, G. (2020). Towards a System of Distributional National Accounts: Methods and Global Inequality Estimates from WID.world. *Economie et Statistique / Economics and Statistics*, (517-518-5):41–59.

Angrist, J. D. and Rokkanen, M. (2015). Wanna Get Away? Regression Discontinuity Estimation of Exam School Effects Away from the Cutoff. *Journal of the American Statistical Association*, 110(512):1331–1344.

Aoki, S. and Nirei, M. (2016). Pareto Distribution of Income in Neoclassical Growth Models. *Review of Economic Dynamics*, 20:25–42.

Aoki, S. and Nirei, M. (2017). Zipf's Law, Pareto's Law, and the Evolution of Top Incomes in the United States. *American Economic Journal: Macroeconomics*, 9(3):36–71.

Atems, B. and Jones, J. (2015). Income inequality and economic growth: A panel VAR approach. *Empirical Economics*, 48(4):1541–1561.

Atkinson, A. B. (2015). *Inequality : what can be done?* Harvard University Press, Massachusetts.

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

Auclert, A. and Rognlie, M. (2018). Inequality and Aggregate Demand. Working Paper 24280, National Bureau of Economic Research.

Auerbach, A. (1989). Tax reform and adjustment costs: The impact on investment and market value. *International Economic Review*, 30(4):939–62.

Autor, D. H., Manning, A., and Smith, C. L. (2016). The Contribution of the Minimum Wage to US Wage Inequality over Three Decades: A Reassessment. *American Economic Journal: Applied Economics*, 8(1):58–99.

Azar, J., Huet-Vaughn, E., Marinescu, I., Taska, B., and von Wachter, T. (2019). Minimum Wage Employment Effects and Labor Market Concentration. Working Paper 26101, National Bureau of Economic Research.

Bach, L., Calvet, L., and Sodini, P. (2016). Rich Pickings? Risk, Return, and Skill in the Portfolios of the Wealthy. Discussion Paper 11734, Centre for Economic Policy Research.

Bachmann, R. and Frings, H. (2017). Monopsonistic competition, low-wage labour markets, and minimum wages: An empirical analysis. *Applied Economics*, 49(51):5268–5286.

Balistreri, E. J., McDaniel, C. A., and Wong, E. V. (2003). An estimation of us industry-level capital–labor substitution elasticities: support for cobb–douglas. *The North American Journal of Economics and Finance*, 14(3):343–356.

Baltagi, B. and Baltagi, P. (2001). *Econometric Analysis of Panal Data*. Wiley.

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

Barkai, S. (2020). Declining Labor and Capital Shares. *Journal of Finance*, 75(5):2421–2463.

Barro, R. and Lee, J. W. (2013). A new data set of educational attainment in the world, 1950–2010. *Journal of Development Economics*, 104(C):184–198.

Barro, R. J. (2000). Inequality and Growth in a Panel of Countries. *Journal of Economic Growth*, 5(1):5–32.

Becker, G. S. and Tomes, N. (1994). Human Capital and the Rise and Fall of Families. In *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education, Third Edition*, NBER Chapters, pages 257–298. National Bureau of Economic Research.

Benabou, R. (1996). Inequality and Growth. Working Paper 5658, National Bureau of Economic Research.

Benhabib, J. and Bisin, A. (2018). Skewed Wealth Distributions: Theory and Empirics. *Journal of Economic Literature*, 56(4):1261–1291.

Benhabib, J., Bisin, A., and Luo, M. (2017). Earnings Inequality and Other Determinants of Wealth Inequality. *American Economic Review*, 107(5):593–597.

Benhabib, J., Bisin, A., and Luo, M. (2019). Wealth Distribution and Social Mobility in the US: A Quantitative Approach. *American Economic Review*, 109(5):1623–1647.

Benhabib, J., Bisin, A., and Zhu, S. (2015). The wealth distribution in Bewley economies with capital income risk. *Journal of Economic Theory*, 159(PA):489–515.

Bertanha, M. and Imbens, G. W. (2020). External Validity in Fuzzy Regression Discontinuity Designs. *Journal of Business & Economic Statistics*, 38(3):593–612.

Bertola, G. and Boeri, T. (2002). EMU Labour Markets Two Years On: Microeconomic Tensions and Institutional Evolution. In Buti, M. and Sapir, A., editors, *EMU and Economic Policy in Europe*, Chapters, chapter 10. Edward Elgar Publishing.

Bewley, T. (1986). Stationary monetary equilibrium with a continuum of independently fluctuating consumers. *Contributions to mathematical economics in honor of Gerard Debreu*, 79.

Bhaskar, V., Manning, A., and To, T. (2002). Oligopsony and Monopsonistic Competition in Labor Markets. *Journal of Economic Perspectives*, 16(2):155–174.

Blinder, A. S. and Deaton, A. (1985). The Time Series Consumption Function Revisited. *Brookings Papers on Economic Activity*, 16(2):465–521.

Bloom, N., Guvenen, F., Smith, B. S., Song, J., and von Wachter, T. (2018). The Disappearing Large-Firm Wage Premium. *AEA Papers and Proceedings*, 108:317–322.

Boeri, T. (2012). Setting the minimum wage. *Labour Economics*, 19(3):281–290.

Boskin, M. J. (1988). Issues in the measurement and interpretation of saving and wealth. Working Paper 2633, National Bureau of Economic Research.

Botta, A., Caverzasi, E., Russo, A., Gallegati, M., and Stiglitz, J. E. (2019). Inequality and finance in a rent economy. *Journal of Economic Behavior & Organization*.

Bradford, D. F. (1989). Market value vs. financial accounting measures of national saving. Working Paper 2906, National Bureau of Economic Research.

Brown, C. (1999). *Minimum wages, employment, and the distribution of income*, volume 3, Part B. Elsevier, 1 edition.

Brown, W. (2009). The Process of Fixing the British National Minimum Wage, 1997-2007. Cambridge Working Papers in Economics 0904, Faculty of Economics, University of Cambridge.

Brueckner, M. and Lederman, D. (2018). Inequality and economic growth: the role of initial income. *Journal of Economic Growth*, 23(3):341–366.

Brun, L. and Gonzalez, I. (2017). Tobin's Q and Inequality. *SSRN Electronic Journal*.

Buchanan, J. M. (1980). *Rent seeking and profit seeking*. Texas A&M University Press, Texas.

Cagetti, M. and De Nardi, M. (2006). Entrepreneurship, Frictions, and Wealth. *Journal of Political Economy*, 114(5):835–870.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014a). Robust Data-Driven Inference in the Regression-Discontinuity Design. *The Stata Journal*, 14(4):909–946.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014b). Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs. *Econometrica*, 82(6):2295–2326.

Cameron, A. and Trivedi, P. (2010). *Microeconometrics Using Stata, Revised Edition*. StataCorp LP.

Cao, D. and Luo, W. (2017). Persistent Heterogeneous Returns and Top End Wealth Inequality. *Review of Economic Dynamics*, 26:301–326.

Card, D. and Krueger, A. (1994). Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania. *American Economic Review*, 84(4):772–93.

Card, D. and Krueger, A. B. (1995). *Myth and Measurement: The New Economics of the Minimum Wage*. Princeton University Press, Princeton.

Carroll, C. D. (1998). Why Do the Rich Save So Much? Working Paper 6549, National Bureau of Economic Research.

Castaneda, A., Diaz-Gimenez, J., and Rios-Rull, J. V. (2003). Accounting for the U.S. Earnings and Wealth Inequality. *Journal of Political Economy*, 111(4):818–857.

Cattaneo, M., Idrobo, N., and Titiunik, R. (2019). *A Practical Introduction to Regression Discontinuity Designs: Foundations*. Elements in Quantitative and Computational Methods for the Social Sciences. Cambridge University Press.

Cattaneo, M., Idrobo, N., and Titiunik, R. (2020a). *A Practical Introduction to Regression Discontinuity Designs: Extensions*. Elements in Quantitative and Computational Methods for the Social Sciences. Cambridge University Press.

Cattaneo, M., Titiunik, R., and Vazquez-Bare, G. (2020b). *The Regression Dscontinuity Design*. Handbook of Research Methods in Political Science and International Relations. Sage Publications.

Cattaneo, M. and Vazquez-Bare, G. (2016). The Choice of Neighborhood in Regression Discontinuity Designs. *Observational Studies*, 2:134–146.

Cavalcanti, T. V. V. and Giannitsarou, C. (2017). Growth and Human Capital: A Network Approach. *Economic Journal*, 127(603):1279–1317.

Cengiz, D., Dube, A., Lindner, A., and Zipperer, B. (2019). The Effect of Minimum Wages on Low-Wage Jobs. *Quarterly Journal of Economics*, 134(3):1405–1454.

Chamley, C. (1986). Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives. *Econometrica*, 54(3):607–622.

Chang, H. J. (1994). *The Political Economy of Industrial Policy*. Macmillan, London and Basingstoke.

Chirinko, B. (2008). σ: The long and short of it. *Journal of Macroeconomics*, 30(2):671–686.

Chirinko, R. S. and Mallick, D. (2017). The Substitution Elasticity, Factor Shares, and the Low-Frequency Panel Model. *American Economic Journal: Macroeconomics*, 9(4):225–253.

Choi, I. (2001). Unit root tests for panel data. *Journal of International Money and Finance*, 20(2):249–272.

Clemens, J. (2015). The Minimum Wage and the Great Recession: Evidence from the Current Population Survey. Working Paper 21830, National Bureau of Economic Research.

Collard-Wexler, A. and De Loecker, J. (2016). Production function estimation and capital measurement error. Working Paper 22437, National Bureau of Economic Research.

Costantini, M. and Paradiso, A. (2018). What do panel data say on inequality and GDP? New evidence at US state-level. *Economics Letters*, 168(C):115–117.

Couch, K. A. and Wittenburg, D. C. (2001). The Response of Hours of Work to Increases in the Minimum Wage. *Southern Economic Journal*, 68(1):171–177.

Currie, J. and Fallick, B. C. (1996). The minimum wage and the employment of youth evidence from the nlsy. *The Journal of Human Resources*, 31(2):404–428.

Cutler, D. M. and Katz, L. F. (1992). Rising Inequality? Changes in the Distribution of Income and Consumption in the 1980's. *American Economic Review*, 82(2):546–551.

Davis, M. A. and Heathcote, J. (2007). The price and quantity of residential land in the United States. *Journal of Monetary Economics*, 54(8):2595–2620.

De Nardi, M. (2004). Wealth Inequality and Intergenerational Links. *Review of Economic Studies*, 71(3):743–768.

De Nardi, M. (2015). Quantitative Models of Wealth Inequality: A Survey. Working Paper 21106, National Bureau of Economic Research.

De Nardi, M. and Fella, G. (2017). Saving and Wealth Inequality. *Review of Economic Dynamics*, 26:280–300.

den Haan, W. J. (2009). Solving Dynamic Models with Heterogeneous Agents and Aggregate Uncertainty with Dynare or Dynare++. Meeting Paper 776, Society for Economic Dynamics.

den Haan, W. J. and Rendahl, P. (2010). Solving the incomplete markets model with aggregate uncertainty using explicit aggregation. *Journal of Economic Dynamics and Control*, 34(1):69–78.

Dickens, R., Riley, R., and Wilkinson, D. (2015). A Re-examination of the Impact of the UK National Minimum Wage on Employment. *Economica*, 82(328):841–864.

Doucouliagos, H. and Stanley, T. D. (2009). Publication Selection Bias in Minimum-Wage Research? A Meta-Regression Analysis. *British Journal of Industrial Relations*, 47(2):406–428.

Draca, M., Machin, S., and Van Reenen, J. (2011). Minimum wages and firm profitability. *American Economic Journal: Applied Economics*, 3(1):129–51.

Drutman, L. (2015). *The Business of America Is Lobbying: How Corporations Became Politicized and Politics Became More Corporate*. Oxford University Press, New York.

Dube, A., Lester, T. W., and Reich, M. (2010). Minimum wage effects across state borders: Estimates using contiguous counties. *Review of Economics and Statistics*, 92(4):945–964.

Eggertsson, G. B., Robbins, J. A., and Wold, E. G. (2018). Kaldor and Piketty's Facts: The Rise of Monopoly Power in the United States. Working Paper 24287, National Bureau of Economic Research.

Eisner, R. (1988). Extended accounts for national income and product. *Journal of Economic Literature*, 26(4):1611–1684.

Esteban, J. and Ray, D. (2006). Inequality, Lobbying, and Resource Allocation. *American Economic Review*, 96(1):257–279.

Fagereng, A., Guiso, L., Malacrino, D., and Pistaferri, L. (2016). Heterogeneity in Returns to Wealth and the Measurement of Wealth Inequality. *American Economic Review*, 106(5):651–55.

Fagereng, A., Guiso, L., Malacrino, D., and Pistaferri, L. (2020). Heterogeneity and Persistence in Returns to Wealth. *Econometrica*, 88(1):115–170.

Fagereng, A., Holm, M. B., Moll, B., and Natvik, G. (2019). Saving Behavior Across the Wealth Distribution: The Importance of Capital Gains. Working Paper 26588, National Bureau of Economic Research.

Feenberg, D. and Coutts, E. (1993). An introduction to the TAXSIM model. *Journal of Policy Analysis and Management*, 12(1):189–194.

Fidrmuc, J. and Tena, J. D. (2018). UK national minimum wage and labor market outcomes of young workers. *Economics - The Open-Access, Open-Assessment E-Journal*, 12:1–28.

Flavin, M. and Yamashita, T. (2002). Owner-Occupied Housing and the Composition of the Household Portfolio. *American Economic Review*, 92(1):345–362.

Flinn, C. (2006). Minimum Wage Effects on Labor Market Outcomes under Search, Matching, and Endogenous Contact Rates. *Econometrica*, 74(4):1013–1062.

Flinn, C. J. (2010). *The Minimum Wage and Labor Market Outcomes*. The MIT Press.

Forbes, K. J. (2000). A Reassessment of the Relationship between Inequality and Growth. *American Economic Review*, 90(4):869–887.

Fox, D. and McCully, C. (2018). *NIPA Handbook: Concepts and Methods of the U.S. National Income and Product Accounts*. Bureau of Economic Analysis.

Frank, M. W. (2009). Inequality and Growth in The United States: Evidence from a New State-level Panel of Income Inequality Measures. *Economic Inquiry*, 47(1):55–68.

Gabaix, X., Lasry, J.-M., Lions, P.-L., and Moll, B. (2016). The dynamics of inequality. *Econometrica*, 84(6):2071–2111.

Galor, O. and Moav, O. (2004). From Physical to Human Capital Accumulation: Inequality and the Process of Development. *Review of Economic Studies*, 71(4):1001–1026.

Galor, O. and Tsiddon, D. (1997). The Distribution of Human Capital and Economic Growth. *Journal of Economic Growth*, 2(1):93–124.

Galor, O. and Zeira, J. (1993). Income Distribution and Macroeconomics. *Review of Economic Studies*, 60(1):35–52.

Gelman, A. and Imbens, G. (2019). Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs. *Journal of Business & Economic Statistics*, 37(3):447–456.

George, H. (1879). *Progress and Poverty*. Kegan Paul, Tench & Co., London.

Giuliano, L. (2013). Minimum Wage Effects on Employment, Substitution, and the Teenage Labor Supply: Evidence from Personnel Data. *Journal of Labor Economics*, 31(1):155–194.

Goes, C. (2017). Testing Piketty's Hypothesis on the Drivers of Income Inequality: Evidence from Panel VARs with Heterogeneous Dynamics. *IMF Staff Papers*.

Goldsmith, R. W. (1982). *The National Balance Sheet of the United States, 1953–1980*. National Bureau of Economic Research, Inc.

Gruber, J. W. and Kamin, S. B. (2017). Corporate buybacks and capital investment: An international perspective. IFDP Notes 2017-04-11, Board of Governors of the Federal Reserve System (U.S.).

Gutierrez, G. and Philippon, T. (2017). Investmentless Growth: An Empirical Investigation. *Brookings Papers on Economic Activity*, pages 89–169.

Gutiérrez, G. and Philippon, T. (2017). Declining competition and investment in the u.s. Working Paper 23583, National Bureau of Economic Research.

Guvenen, F. (2009). An Empirical Investigation of Labor Income Processes. *Review of Economic Dynamics*, 12(1):58–79.

Hadri, K. (2000). Testing for stationarity in heterogeneous panel data. *Econometrics Journal*, 3(2):148–161.

Hahn, J., Todd, P., and der Klaauw, W. V. (2001). Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design. *Econometrica*, 69(1):201–209.

Hall, R. E. and Mishkin, F. S. (1982). The Sensitivity of Consumption to Transitory Income: Estimates from Panel Data on Households. *Econometrica*, 50(2):461–481.

Halter, D., Oechslin, M., and Zweimüller, J. (2014). Inequality and growth: The neglected time dimension. *Journal of Economic Growth*, 19(1):81–104.

Hamermesh, D. S. (1995). Labour Demand and the Source of Adjustment Costs. *Economic Journal*, 105(430):620–634.

Hamermesh, D. S. (2014). Do labor costs affect companies' demand for labor? *IZA World of Labor*, pages 1–3.

Hansen, G. and Imrohoroglu, A. (1992). The Role of Unemployment Insurance in an Economy with Liquidity Constraints and Moral Hazard. *Journal of Political Economy*, 100(1):118–42.

Harasztosi, P. and Lindner, A. (2019). Who Pays for the Minimum Wage? *American Economic Review*, 109(8):2693–2727.

Harris, R. D. and Tzavalis, E. (1999). Inference for unit roots in dynamic panels where the time dimension is fixed. *Journal of Econometrics*, 91(2):201–226.

Heathcote, J. (2005). Fiscal Policy with Heterogeneous Agents and Incomplete Markets. *Review of Economic Studies*, 72(1):161–188.

Heathcote, J., Storesletten, K., and Violante, G. L. (2010). The Macroeconomic Implications of Rising Wage Inequality in the United States. *Journal of Political Economy*, 118(4):681–722.

Hendricks, L. (2007). How important is discount rate heterogeneity for wealth inequality? *Journal of Economic Dynamics and Control*, 31(9):3042–3068.

Hicks, J. R. (1975). *Value and Capital: An Inquiry into some Fundamental Principles of Economic Theory*. Number 9780198282693 in OUP Catalogue. Oxford University Press.

Hirsch, B., Kaufman, B., and Zelenska, T. (2015). Minimum Wage Channels of Adjustment. *Industrial Relations: A Journal of Economy and Society*, 54(2):199–239.

Hirsch, F. (1977). *Social Limits to Growth*. Routledge and Kegan Paul, London.

Holmquist, E. B. and McIntosh, S. H. (2015). U.S. Net Wealth in the Financial Accounts of the United States. Feds notes, Board of Governors of the Federal Reserve System.

Hubmer, J., Krusell, P., and Smith, A. J. (2016). The Historical Evolution of the Wealth Distribution: A Quantitative-Theoretic Investigation. Working Paper 23011, National Bureau of Economic Research.

Huggett, M. (1993). The risk-free rate in heterogeneous-agent incomplete-insurance economies. *Journal of Economic Dynamics and Control*, 17(5-6):953–969.

ILO (2014). *General Survey on Minimum Wage Systems*. International Labour Organization.

Im, K. S., Pesaran, M., and Shin, Y. (2003). Testing for unit roots in heterogeneous panels. *Journal of Econometrics*, 115(1):53–74.

Imbens, G. W. and Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, 142(2):615–635.

Imrohoroglu, A. (1992). The welfare cost of inflation under imperfect insurance. *Journal of Economic Dynamics and Control*, 16(1):79–91.

Jones, C. I. and Kim, J. (2018). A Schumpeterian Model of Top Income Inequality. *Journal of Political Economy*, 126(5):1785–1826.

Judd, K. L. (1985). Redistributive taxation in a simple perfect foresight model. *Journal of Public Economics*, 28(1):59–83.

Kaas, L. and Madden, P. (2008). Holdup in oligopsonistic labour markets – a new role for the minimum wage. *Labour Economics*, 15(3):334–349.

Kaldor, N. (1955). Alternative Theories of Distribution. *Review of Economic Studies*, 23(2):83–100.

Kaldor, N. (1957). A Model of Economic Growth. *Economic Journal*, 67(268):591–624.

Kaldor, N. (1961). *Capital Accumulation and Economic Growth*. Palgrave Macmillan UK, London.

Karabarbounis, L. and Neiman, B. (2013). The Global Decline of the Labor Share. *Quarterly Journal of Economics*, 129(1):61–103.

Karabarbounis, L. and Neiman, B. (2018). Accounting for Factorless Income. Working Paper 24404, National Bureau of Economic Research.

Kim, C. J. and Nelson, C. R. (1999). Has the U.S. Economy Become More Stable? A Bayesian Approach Based on a Markov-Switching Model of the Business Cycle. *Review of Economics and Statistics*, 81(4):608–616.

Kim, K. M. (2010). Changes in korean labor market regimes. *Seoul Journal of Korean Studies*, 23:247–269.

Kim, S. (1999). Do monetary policy shocks matter in the G-7 countries? Using common identifying assumptions about monetary policy across countries. *Journal of International Economics*, 48(2):387–412.

Kindermann, F. and Krueger, D. (2014). High Marginal Tax Rates on the Top 1%? Lessons from a Life Cycle Model with Idiosyncratic Income Risk. Working Paper 20601, National Bureau of Economic Research.

Kitao, S. (2008). Entrepreneurship, taxation and capital investment. *Review of Economic Dynamics*, 11(1):44–69.

Klump, R., McAdam, P., and Willman, A. (2007). The long-term success of the neoclassical growth model. *Oxford Review of Economic Policy*, 23(1):94–114.

Krueger, A. O. (1974). The Political Economy of the Rent-Seeking Society. *American Economic Review*, 64(3):291–303.

Krusell, P. and Smith, A. J. (1998). Income and Wealth Heterogeneity in the Macroeconomy. *Journal of Political Economy*, 106(5):867–896.

Kuhn, M., Schularick, M., and Steins, U. I. (2020). Income and wealth inequality in america, 1949–2016. *Journal of Political Economy*, 128(9):3469–3519.

Kuznets, S. (1955). Economic Growth and Income Inequality. *American Economic Review*, 45(1):1–28.

Kydland, F. E. and Prescott, E. C. (1982). Time to Build and Aggregate Fluctuations. *Econometrica*, 50(6):1345–1370.

Lee, D. S. (1999). Wage Inequality in the United States During the 1980s: Rising Dispersion or Falling Minimum Wage? *Quarterly Journal of Economics*, 114(3):977–1023.

Lee, J. W. and Lee, H. (2018). Human capital and income inequality. *Journal of the Asia Pacific Economy*, 23(4):554–583.

Lee, J. Y. (2020). Rentier Premium and Wealth Inequality. Conference Paper 226, Royal Economic Society.

Lemos, S. (2008). A survey of the effects of the minimum wage on prices. *Journal of Economic Surveys*, 22(1):187–212.

Levin, A., Lin, C. F., and Chu, J. (2002). Unit root tests in panel data: Asymptotic and finite-sample properties. *Journal of Econometrics*, 108(1):1–24.

Love, I. and Zicchino, L. (2006). Financial development and dynamic investment behavior: Evidence from panel VAR. *Quarterly Review of Economics and Finance*, 46(2):190–210.

Lucas, Robert E, J. (1990). Why Doesn't Capital Flow from Rich to Poor Countries? *American Economic Review*, 80(2):92–96.

Lusardi, A., Michaud, P. C., and Mitchell, O. S. (2017). Optimal Financial Knowledge and Wealth Inequality. *Journal of Political Economy*, 125(2):431–477.

Lutkepohl, H. (2007). *New Introduction to Multiple Time Series Analysis*. Springer Publishing Company, Incorporated.

MaCurdy, T. (2015). How Effective Is the Minimum Wage at Supporting the Poor? *Journal of Political Economy*, 123(2):497–545.

Malthus, T. (1798). *An Essay on the Principle of Population*. Macmillan, London.

Manning, A. (2003). *Monopsony in Motion: Imperfect Competition in Labor Markets*. Princeton University Press.

Manning, A. (2011). Minimum wages and wage inequality. In Marsden, D., editor, *Labour Market Policy for the Twenty-First Century*, chapter 10. Oxford University Press.

Manning, A. (2016). The Elusive Employment Effect of the Minimum Wage. Discussion Paper 1428, Centre for Economic Performance, LSE.

Marglin, S. A. and Schor, J. B., editors (1992). *The Golden Age of Capitalism: Reinterpreting the Postwar Experience*. Oxford University Press.

McConnell, M. M. and Perez-Quiros, G. (2000). Output Fluctuations in the United States: What Has Changed since the Early 1980's? *American Economic Review*, 90(5):1464–1476.

Meer, J. and West, J. (2016). Effects of the Minimum Wage on Employment Dynamics. *Journal of Human Resources*, 51(2):500–522.

Mehra, R. and Prescott, E. C. (1985). The equity premium: A puzzle. *Journal of Monetary Economics*, 15(2):145–161.

Metcalf, D. (2008). Why has the British National Minimum Wage had Little or No Impact on Employment? *Journal of Industrial Relations*, 50(3):489–512.

Michl, T. R. (2000). Can Rescheduling Explain the New Jersey Minimum Wage Studies? *Eastern Economic Journal*, 26(3):265–276.

Mill, J. J. (1821). *Elements of Political Economy*. Baldwin, London.

Mill, J. S. (1848). *Principles of Political Economy with Some of their Applications to Social Philosophy*. John W. Parker, London.

Mukand, S. W. and Rodrik, D. (2020). The Political Economy of Liberal Democracy. *Economic Journal*, 130(627):765–792.

Neumark, D. and Corella, L. F. M. (2019). Do Minimum Wages Reduce Employment in Developing Countries? A Survey and Exploration of Conflicting Evidence. Working Paper 26462, National Bureau of Economic Research.

Neumark, D. and Wascher, W. (2000). Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania: Comment. *American Economic Review*, 90(5):1362–1396.

Neumark, D. and Wascher, W. (2008). *Minimum Wages*. The MIT Press.

Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica*, 49(6):1417–1426.

Nickell, S. J. and Andrews, M. (1983). Unions, Real Wages and Employment in Britain 1951-79. *Oxford Economic Papers*, 35:183–206.

Nirei, M. and Souma, W. (2007). A two factor model of income distribution dynamics. *Review of Income and Wealth*, 53(3):440–459.

OECD (2015). FOCUS on Minimum wages after the crisis: Making them pay. Technical report, Organisation for Economic Cooperation and Development (OECD).

Pagano, U. (2014). The crisis of intellectual monopoly capitalism. *Cambridge Journal of Economics*, 38(6):1409–1429.

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.

Partridge, M. D. (1997). Is Inequality Harmful for Growth? Comment. *American Economic Review*, 87(5):1019–1032.

Pedroni, P. (2013). Structural Panel VARs. *Econometrics*, 1(2):1–27.

Peek, J. (1986). Household wealth composition: the impact of capital gains. *New England Economic Review*, (Nov):26–39.

Perotti, R. (1996). Growth, Income Distribution, and Democracy: What the Data Say. *Journal of Economic Growth*, 1(2):149–187.

Perri, F. and Krueger, D. (2009). How does Household Consumption Respond to Income Shocks? Meeting Paper 14, Society for Economic Dynamics.

Persson, T. and Tabellini, G. (1994). Is Inequality Harmful for Growth? *American Economic Review*, 84(3):600–621.

Pesaran, M. H. (2012). On the interpretation of panel unit root tests. *Economics Letters*, 116(3):545–546.

Pesaran, M. H. and Smith, R. (1995). Estimating long-run relationships from dynamic heterogeneous panels. *Journal of Econometrics*, 68(1):79–113.

Philippon, T. and Reshef, A. (2012). Wages and Human Capital in the U.S. Finance Industry: 1909–2006. *Quarterly Journal of Economics*, 127(4):1551–1609.

Piketty, T. (2014). *Capital in the Twenty-First Century*. Harvard University Press.

Piketty, T., Saez, E., and Stantcheva, S. (2014). Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities. *American Economic Journal: Economic Policy*, 6(1):230–71.

Piketty, T., Saez, E., and Zucman, G. (2018). Distributional National Accounts: Methods and Estimates for the United States. *Quarterly Journal of Economics*, 133(2):553–609.

Piketty, T. and Zucman, G. (2014). Capital is Back: Wealth-Income Ratios in Rich Countries 1700 to 2010. *Quarterly Journal of Economics*, 129(3):1255–1310.

Preston, B. and Roca, M. (2007). Incomplete Markets, Heterogeneity and Macroeconomic Dynamics. Working Paper 13260, National Bureau of Economic Research.

Quadrini, V. (2000). Entrepreneurship, Saving and Social Mobility. *Review of Economic Dynamics*, 3(1):1–40.

Raurich, X., Sala, H., and Sorolla, V. (2012). Factor shares, the price markup, and the elasticity of substitution between capital and labor. *Journal of Macroeconomics*, 34(1):181–198.

Robbins, J. (2019). Capital Gains and the Distribution of Income in the United States. Meeting Paper 202, Society for Economic Dynamics.

Rognlie, M. (2015). Deciphering the Fall and Rise in the Net Capital Share: Accumulation or Scarcity? *Brookings Papers on Economic Activity*, 46(1):1–69.

Rowthorn, R. (1999). Unemployment, wage bargaining and capital—labour substitution. *Cambridge Journal of Economics*, 23(4):413–425.

Rowthorn, R. (2014). A note on Piketty Capital in the Twenty-First Century. *Cambridge Journal of Economics*, 38(5):1275–1284.

Saez, E. and Zucman, G. (2016). Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data. *Quarterly Journal of Economics*, 131(2):519–578.

Schmitt, J. (2013). Why Does the Minimum Wage Have No Discernible Effect on Employment? CEPR Reports and Issue Briefs 2013-04, Center for Economic and Policy Research.

Schumpeter, J. (1942). *Capitalism, Socialism, and Democracy*. Harper and Row, New York.

Shakhnov, K. (2014). The allocation of talent: finance versus entrepreneurship. Economics Working Paper ECO2014/13, European University Institute.

Sheikh, A. and Qiao, H. (2009). Non-Normality of Market Returns: A Framework for Asset Allocation Decision Making. *Journal of Alternative Investments*, 12:8–35.

Shin, K. Y. (2013). Economic crisis, neoliberal reforms, and the rise of precarious work in south korea. *American Behavioral Scientist*, 57(3):335–353.

Simon, H. (1982). *Models of bounded rationality*. The MIT Press.

Sims, C. A. and Zha, T. (2006). Does Monetary Policy Generate Recessions? *Macroeconomic Dynamics*, 10(2):231–272.

Smith, A. (1776). *An Inquiry into the Nature and Causes of the Wealth of Nations*. W. Strahan and T. Cadell, London.

Stachurski, J. and Toda, A. (2019). An impossibility theorem for wealth in heterogeneous-agent models with limited heterogeneity. *Journal of Economic Theory*, 182:1–24.

Stewart, M. B. and Swaffield, J. K. (2008). The Other Margin: Do Minimum Wages Cause Working Hours Adjustments for Low-Wage Workers? *Economica*, 75(297):148–167.

Stiglitz, J. E. (2015a). New Theoretical Perspectives on the Distribution of Income and Wealth among Individuals: Part I. The Wealth Residual. Working Paper 21189, National Bureau of Economic Research.

Stiglitz, J. E. (2015b). New Theoretical Perspectives on the Distribution of Income and Wealth among Individuals: Part II: Equilibrium Wealth Distributions. Working Paper 21190, National Bureau of Economic Research.

Stiglitz, J. E. (2015c). New Theoretical Perspectives on the Distribution of Income and Wealth among Individuals: Part III: Life Cycle Savings vs. Inherited Savings. Working Paper 21191, National Bureau of Economic Research.

Stiglitz, J. E. (2015d). New Theoretical Perspectives on the Distribution of Income and Wealth among Individuals: Part IV: Land and Credit. Working Paper 21192, National Bureau of Economic Research.

Stiglitz, J. E. (2015e). The Origins of Inequality, and Policies to Contain It. *National Tax Journal*, 68(2):425–448.

Straub, L. and Werning, I. (2020). Positive Long-Run Capital Taxation: Chamley-Judd Revisited. *American Economic Review*, 110(1):86–119.

Tinbergen, J. (1974). Substitution of graduate by other labour. *Kyklos*, 27(2):217–226.

Weber, C. E. (2000). 'Rule-of-Thumb' Consumption, Intertemporal Substitution, and Risk Aversion. *Journal of Business & Economic Statistics*, 18(4):497–502.

Zavodny, M. (2000). The effect of the minimum wage on employment and hours. *Labour Economics*, 7(6):729–750.

Appendix A

Appendix for Chapter 1

Table A.1 Panel unit root tests

	Growth		Inequality		Wealth residual	
	Intercept	Plus trend	Intercept	Plus trend	Intercept	Plus trend
LLC	-31.282	-28.299	-28.757	-25.793	-38.767	-38.213
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
HT	-98.440	-60.580	-97.563	-58.968	-86.664	-52.705
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
IPS	-31.949	-30.373	-32.138	-29.270	-36.935	-36.295
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ADF	-16.436	-13.928	-12.124	-7.860	-16.616	-13.846
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Hadri	0.387	1.120	-2.627	0.247	0.152	-1.835
	(0.350)	(0.131)	(0.996)	(0.403)	(0.440)	(0.967)

Notes: The LLC, HT, IPS, and ADF test the null hypothesis of a unit root for each series in the panel, while the Hadri tests the null hypothesis that there is no unit root in any series. The t-statistics are reported and p-values are in parentheses.

Table A.2 Political regime changes of the US, 1978-2015

State	78	80	82	84	85	88	92	94	95	96	00	02	04	06	08	10	12	14	15
Alabama	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0
Alaska	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0
Arizona	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
Arkansas	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
California	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Colorado	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1
Connecticut	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Delaware	1	1	1	1	0	0	0	1	1	1	1	0	0	0	1	1	1	1	1
District of Columbia	1	1	0	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
Florida	1	1	1	1	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0
Georgia	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0
Hawaii	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Idaho	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Illinois	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1
Indiana	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	1	0	0	0
Iowa	0	0	1	1	1	1	1	1	1	0	0	0	0	1	1	1	1	1	1
Kansas	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Kentucky	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	1	1	1	1
Louisiana	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0
Maine	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1
Maryland	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Massachusetts	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Michigan	1	1	1	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0
Minnesota	1	1	1	0	0	1	1	1	1	1	1	0	0	1	1	1	0	1	1
Mississippi	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0
Missouri	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0
Montana	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
Nebraska	1	1	0	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
Nevada	1	1	1	0	0	0	0	1	1	1	1	0	0	0	1	1	1	1	0
New Hampshire	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	0
New Jersey	1	1	1	0	0	1	0	0	0	0	0	0	1	1	1	1	1	1	1
New Mexico	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1
New York	1	1	0	0	0	0	0	1	1	1	1	0	0	0	1	1	1	1	1
North Carolina	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	0	0
North Dakota	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ohio	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
Oklahoma	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0
Oregon	1	1	1	1	1	1	0	0	0	0	0	0	0	1	1	1	1	1	1
Pennsylvania	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
Rhode Island	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
South Carolina	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0
South Dakota	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Tennessee	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0
Texas	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0
Utah	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Vermont	0	0	0	0	0	0	0	1	1	1	1	0	1	1	1	1	1	1	1
Virginia	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	1	1	1	0
Washington	1	0	1	1	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1
West Virginia	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0
Wisconsin	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	0	0	0
Wyoming	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Notes: The Republican Party = 0 and the Democratic party = 1. If there were no changes, I omit reports of the data.
Source: The author utilises the National Conference of State Legislatures

Table A.3 Statistics for labour-friendly and rentier-friendly governments, 1978-2015

(%p,%)

	Rent-related policies						Outcomes		
	ΔTAXw	ΔTAXk	ΔMOR	ΔLTV	ΔHOU	ΔSC	$\bar{Y}$	ΔINQ	ΔWR
Republic (A)	-19.4	+9.2	-8.2	+4.1	-3.9	-8.8	4.7	9.9	19.4
Democrat (B)	-33.8	+12.1	-3.6	+3.5	-0.9	-11.9	5.7	9.2	1.3
Diff. (A-B)	+14.4	-2.9	-4.6	+0.6	-3.0	+3.1	-0.7	+0.7	+18.1

Notes: Variable description is given in 1.4. The growth rate of real per capita income is an averaged value and the other explanatory variables are the changes of the values during its term in the office.

Fig. A.1 Phase space graph of the dynamic model

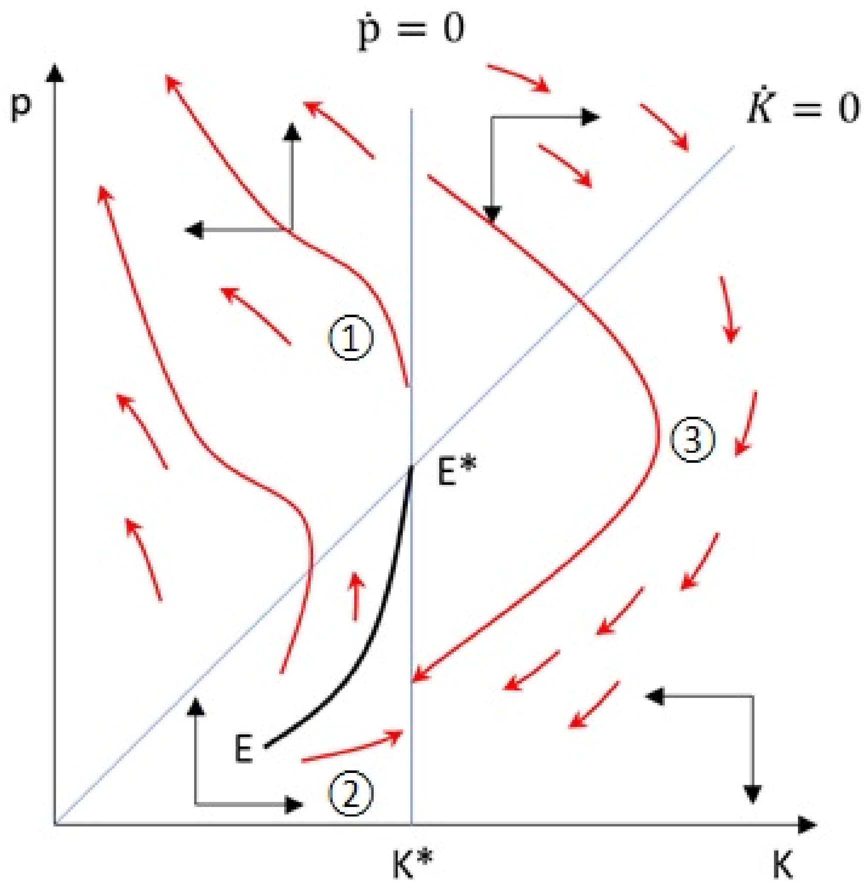
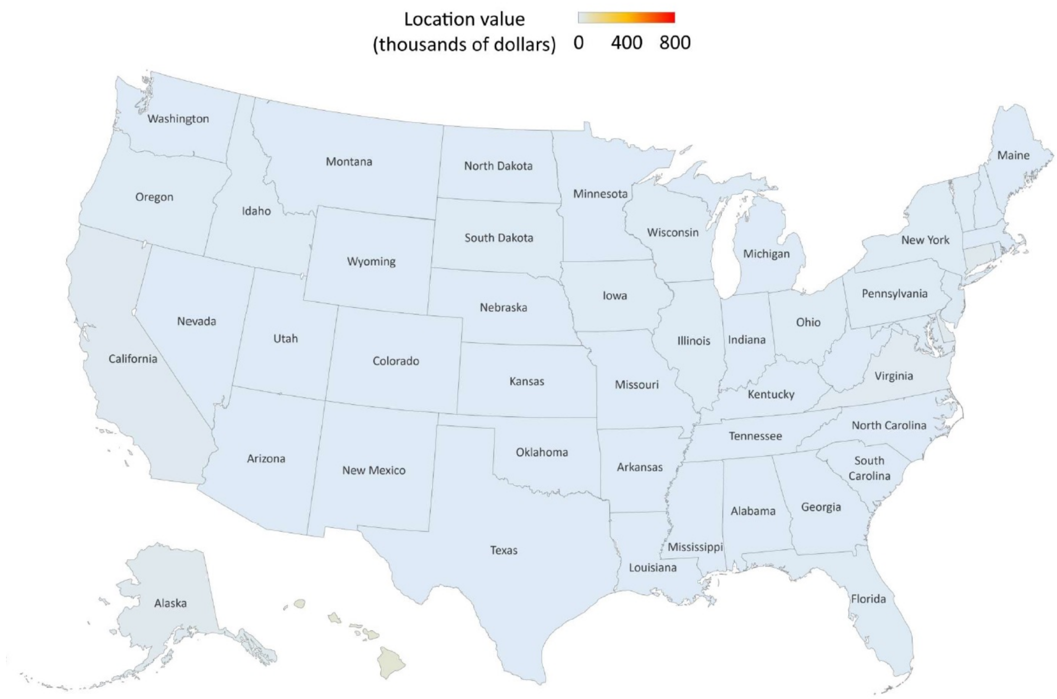
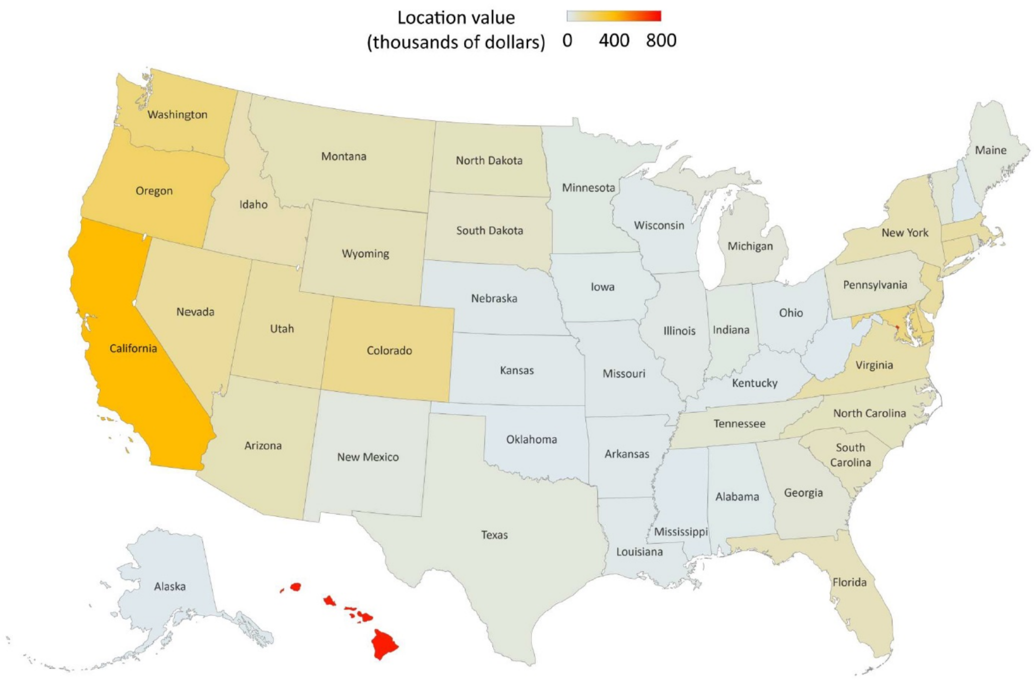


Fig. A.2 Changes in the geographic distribution of wealth residual



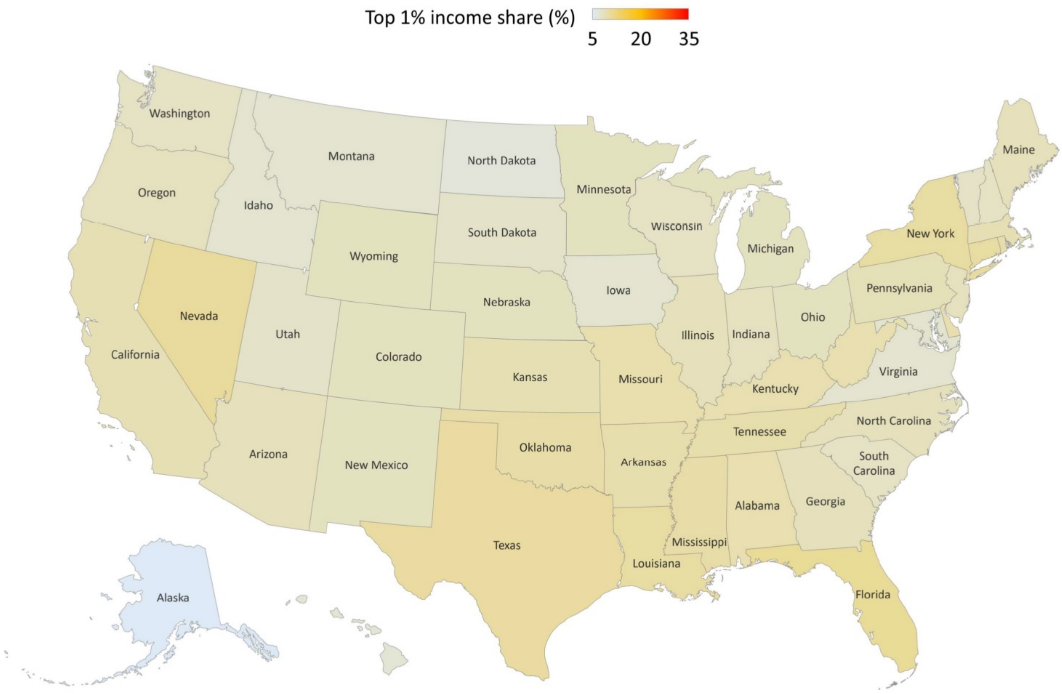
(a) At the end of 1978



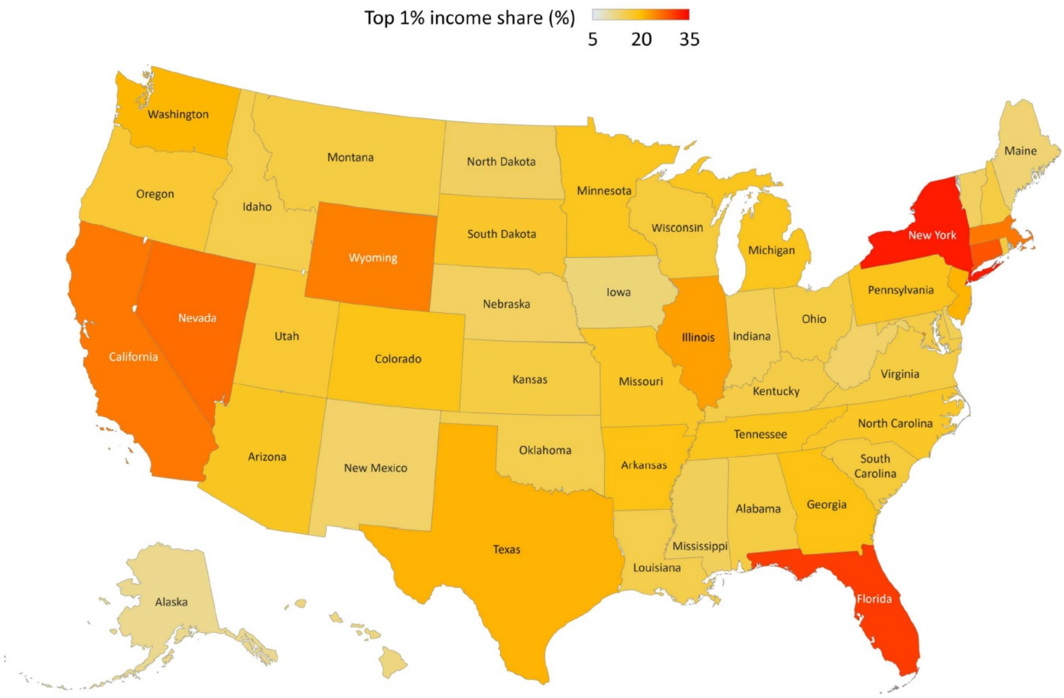
(b) At the end of 2015

Note: Wealth residual refers to the market value of real estate minus the replacement-cost value of structures
Source: The author utilises information from the Lincoln Institute of Land Policy

Fig. A.3 Changes in the geographic distribution of inequality



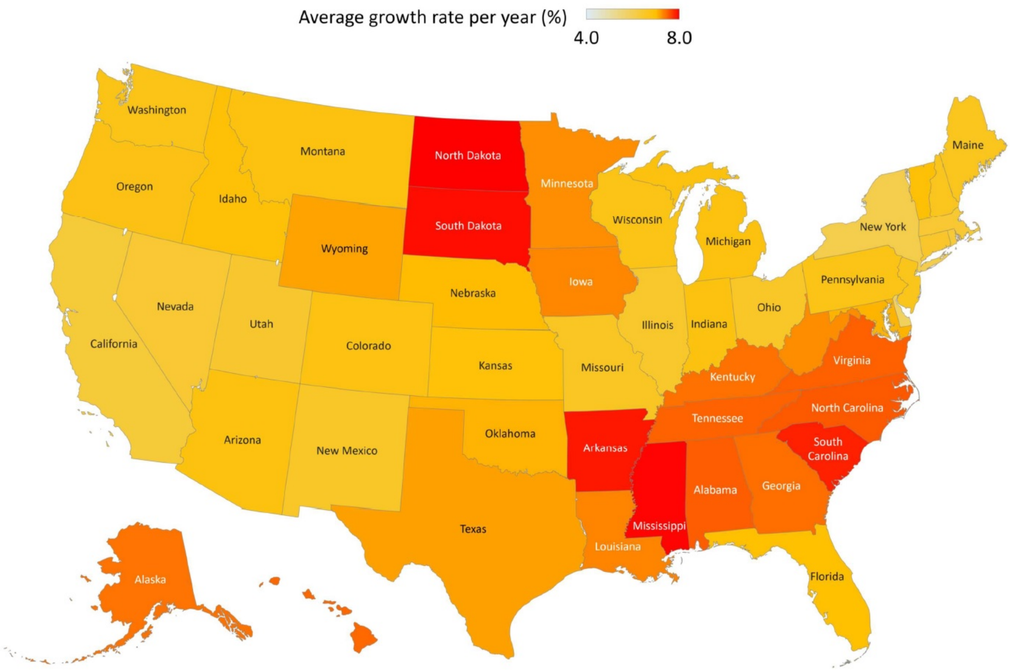
(a) At the end of 1978



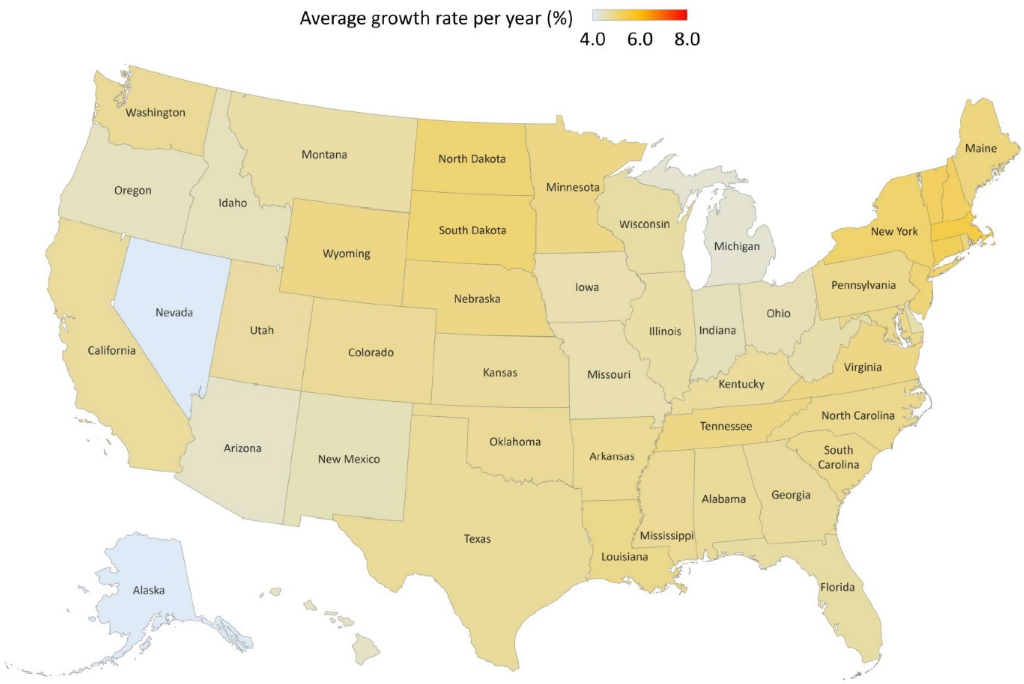
(b) At the end of 2015

Source: The author utilises information from the US State-level Income Inequality Database

Fig. A.4 Changes in the geographic distribution of growth



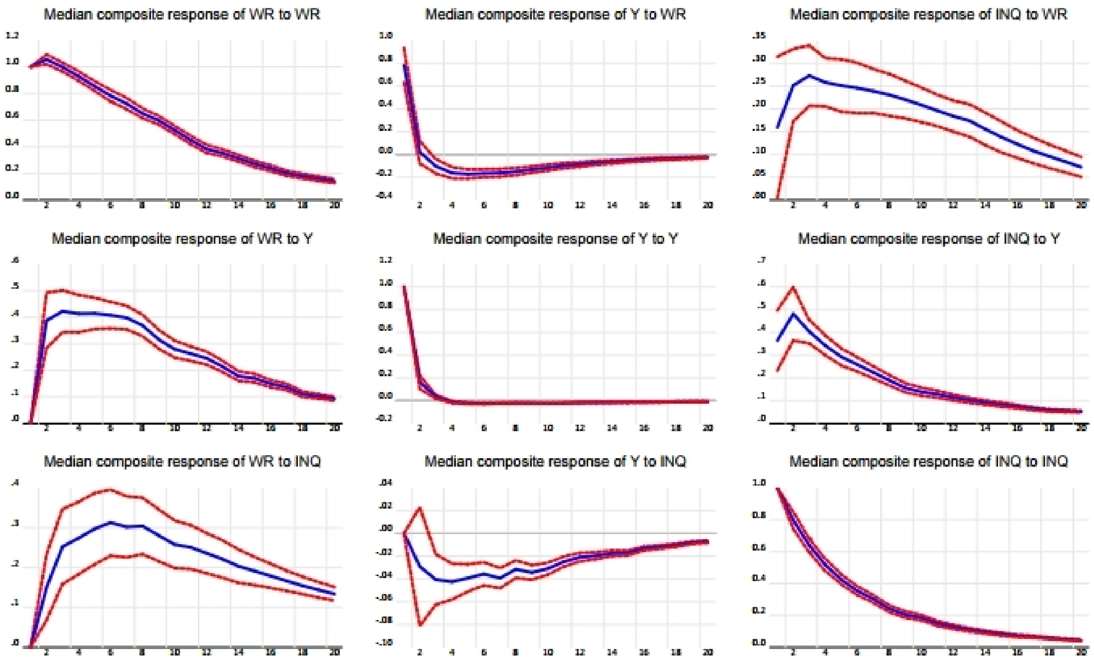
(a) Over the period of 1959-1977



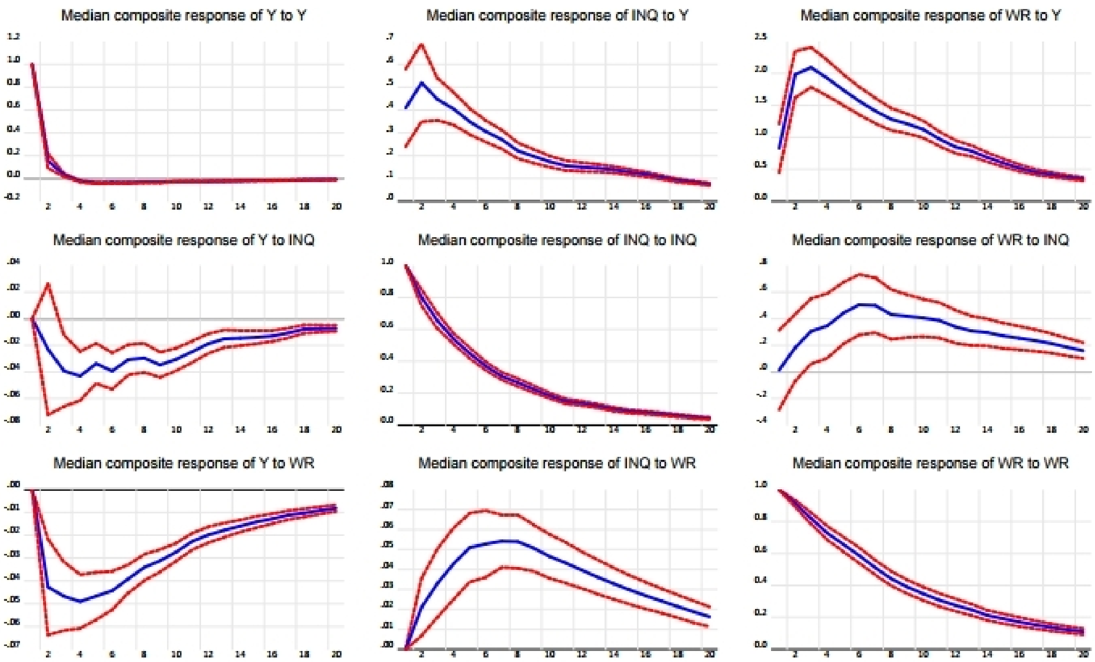
(b) Over the period of 1978-2015

Note: Growth is measured by the growth rate of real per capita personal income
Source: The author utilises information from the Bureau of Economic Analysis

Fig. A.5 Marginal impulse responses- different orderings



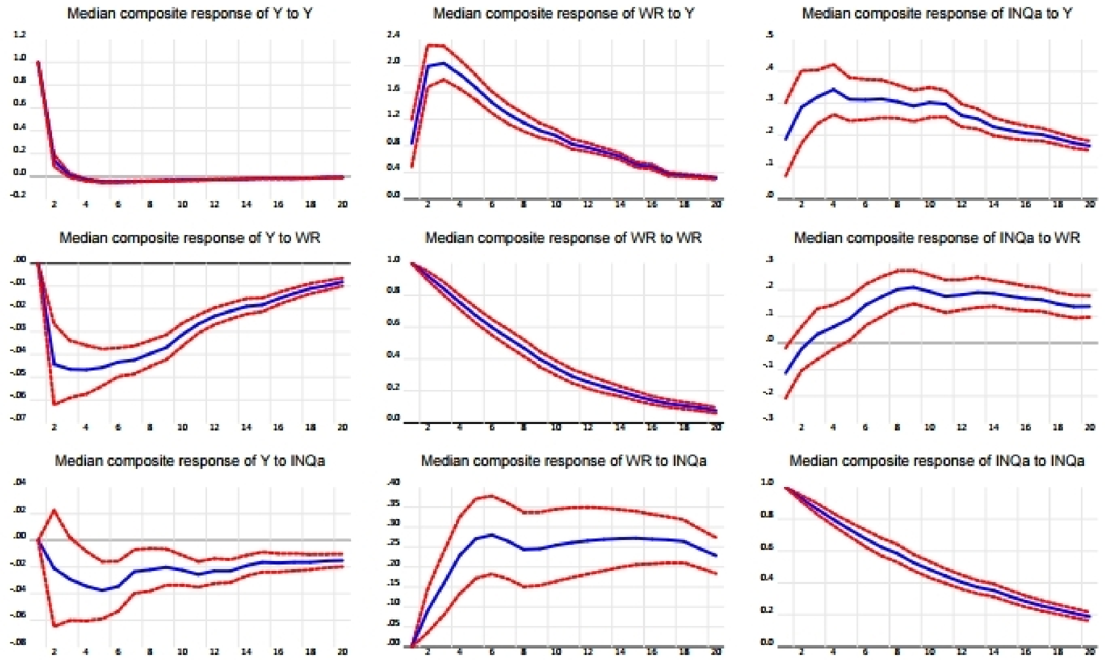
(a) [WR, Y, INQ]



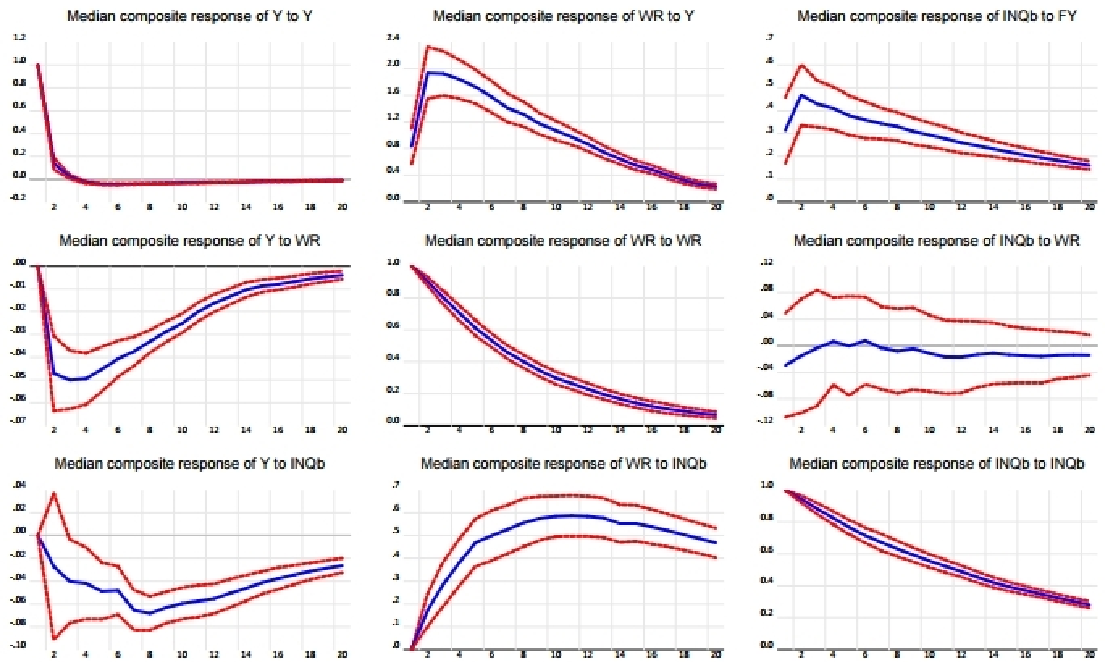
(b) [Y, INQ, WR]

Notes: Each variable denotes the following data: Y, growth rate of real per capita personal income; INQ, share of the top 1% income; and WR, share of land in the market value of real estate. Dotted lines denote 5% confidence intervals calculated from a resampling simulation with 500 repetitions.

Fig. A.6 Marginal impulse responses- alternative measures of inequality



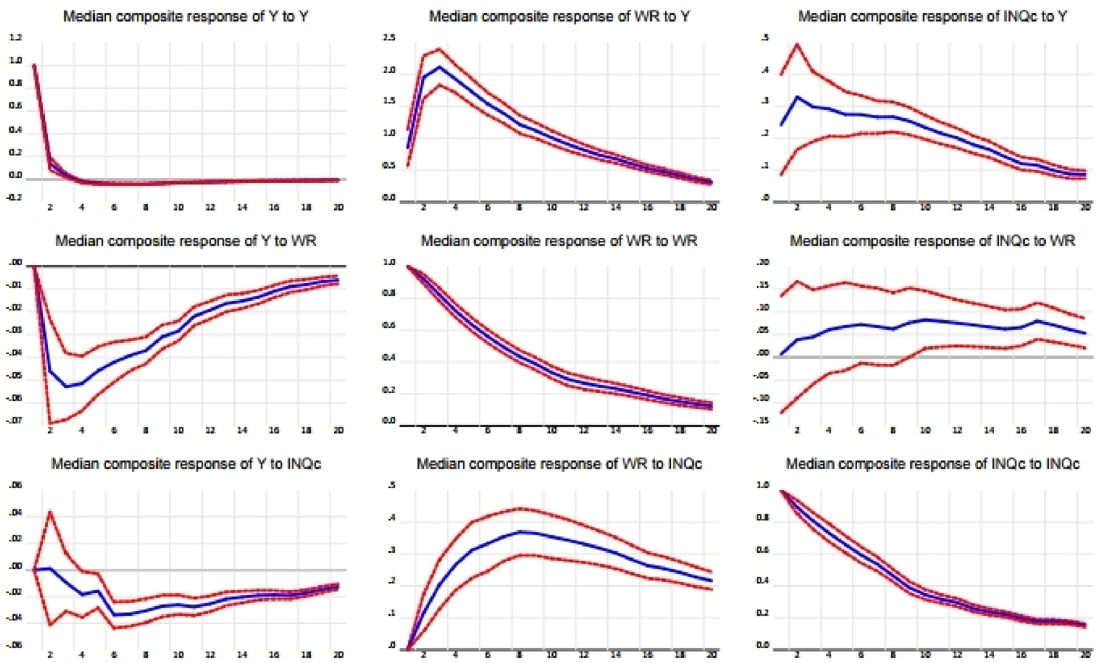
(a) Gini index



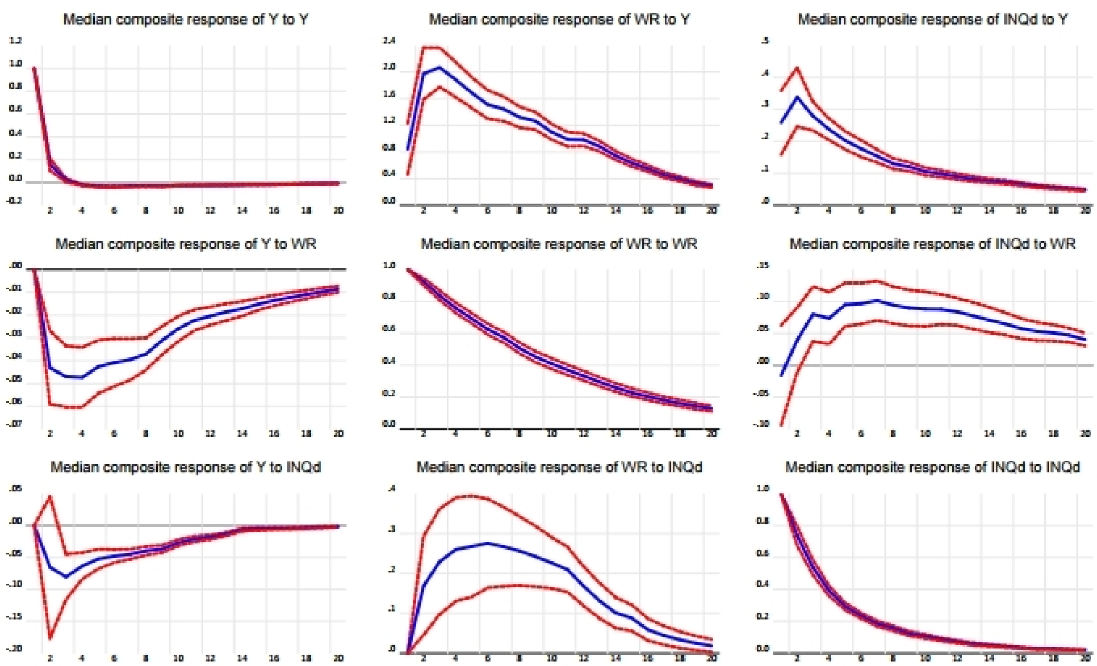
(b) Atkinson index

Notes: Each variable denotes the following data: Y, growth rate of real per capita personal income; INQa, Gini index; INQb, Atkinson index; and WR, share of land in the market value of real estate. Dotted lines denote 5% confidence intervals calculated from a resampling simulation with 500 repetitions.

Fig. A.6 Marginal impulse responses- alternative measures of inequality (cont.)



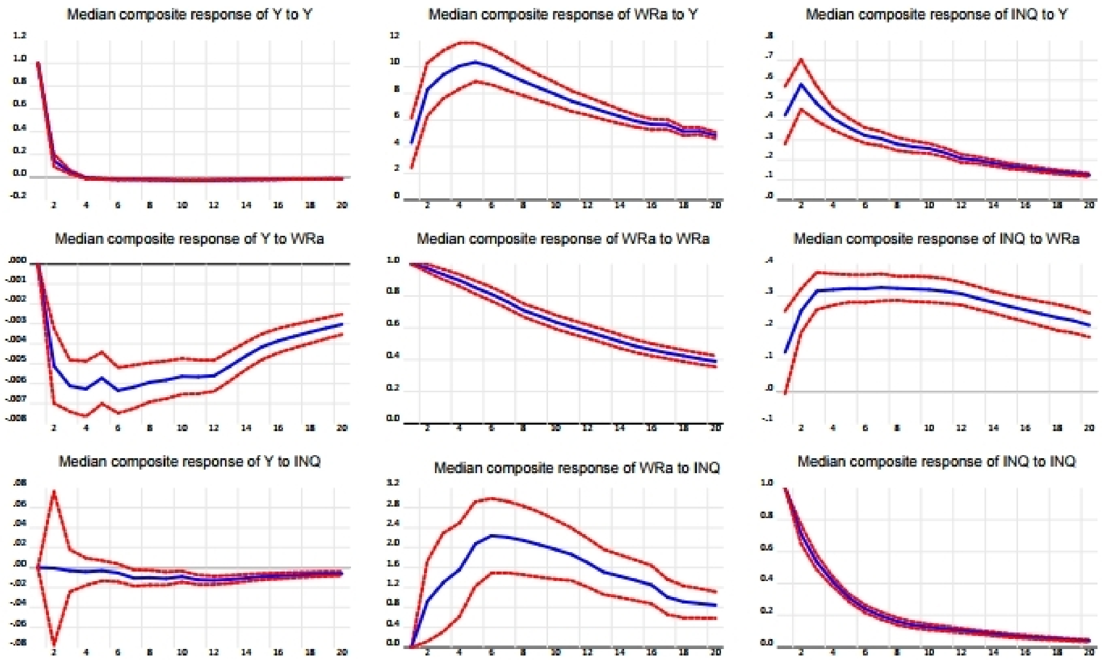
(c) Share of the top 10% income



(d) Share of the top 0.1% income

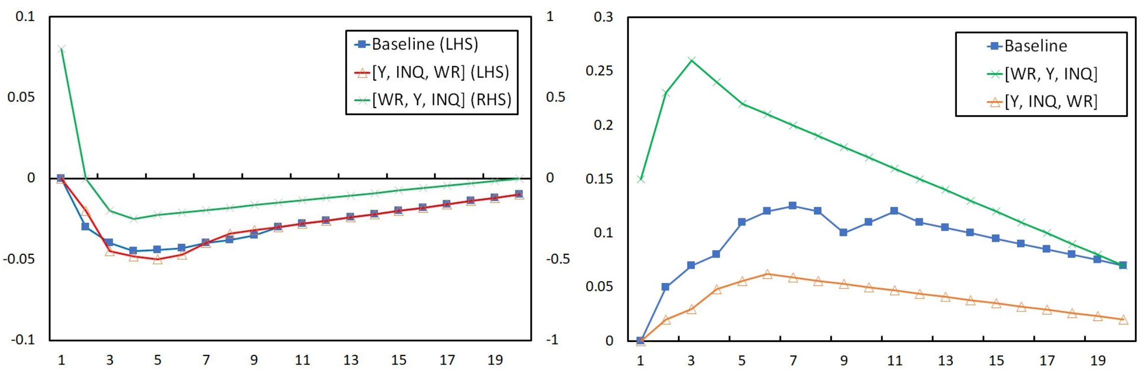
Notes: Each variable denotes the following data: Y, growth rate of real per capita personal income; INQc, share of the top 10% income; INQd, share of the top 0.1% income; and WR, share of land in the market value of real estate. Dotted lines denote 5% confidence intervals calculated from a resampling simulation with 500 repetitions.

Fig. A.7 Marginal impulse responses: an alternative measure of wealth residual



Notes: Each variable denotes the following data: Y, growth rate of real per capita personal income; INQ, share of the top 1% income; and WRa, land price index. Dotted lines denote 5% confidence intervals calculated from a resampling simulation with 500 repetitions.

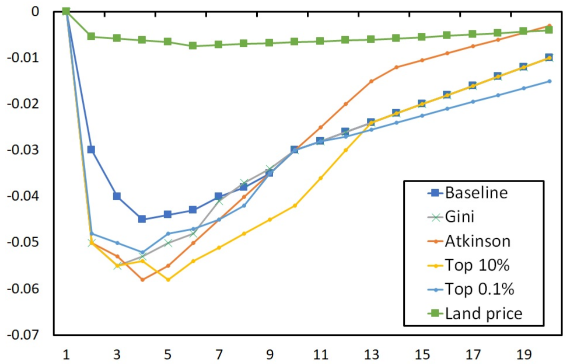
Fig. A.8 Robustness check- different orderings



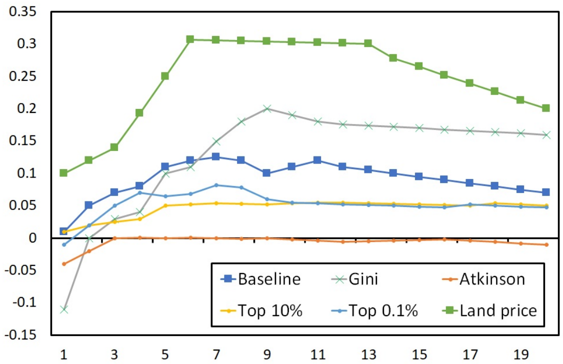
(a) Response of Y to an WR shock (b) Response of INQ to an WR shock

Notes: Y, a measure of growth; INQ, a measure of inequality; and WR, a measure of wealth residual.

Fig. A.9 Robustness check- alternative measures



(a) Response of Y to an WR shock



(b) Response of INQ to an WR shock

Notes: Y, a measure of growth; INQ, a measure of inequality; and WR, a measure of wealth residual.

Appendix B

Appendix for Chapter 2

B.1 Optimal behaviour and equilibrium

Given single markets for labour and capital, wage and rental rates are determined by the aggregate production function (2.15) and the aggregate quantities of the two inputs,

$$l_t \equiv \lambda \int_{i\in R} l_{r,i,t} di + (1-\lambda) \int_{j\in W} l_{w,j,t} dj = \mu^e \bar{l} + \frac{(1-\lambda)\rho^{ez}}{1-\rho^e}(z_t - \mu^z) \tag{B.1}$$

$$k_t \equiv \lambda \int_{i\in R} a_{i,t} di \tag{B.2}$$

where the final equality in equation (B.1) follows the law of large numbers so that $\int_{i\in R} \zeta^e_{r,i,t} di \simeq \int_{j\in W} \zeta^e_{w,j,t} dj \simeq 0$. It implies that all idiosyncratic labour productivity shocks are cancelled out in sum, and thus, the aggregate labour supply is positively correlated with the deviation of the aggregate TFP from its steady state. In other words, the aggregate labour supply behaves procyclically, as in the real-business-cycle literature (notably, Kydland and Prescott, 1982). As usual, wage and rental rates equal

$$w(k_t, l_t, z_t) = (1-\alpha) z_t (\frac{k_t}{l_t})^{\alpha}, \tag{B.3}$$

$$r^k(k_t, l_t, z_t) = \alpha z_t (\frac{k_t}{l_t})^{\alpha-1}. \tag{B.4}$$

To solve the optimisation problem, agents must forecast future prices. Under the assumption that $\{l_{i,t}, \theta_{i,t}, z_t\}$ are governed by stochastic processes, as in equations (2.10), (2.11), (2.14) and (2.16), to forecast future wage and rental rates, agents need to know the processes that determine the evolution of the aggregate capital stock. However, the stochastic properties of the aggregate capital stock depend on the distribution of capital holdings across the population. As a consequence, the whole capital distribution itself becomes a state variable. However, in a setup with a continuum of agents, capital distribution is an infinite dimensional object, which cannot be used as an argument of the individual policy rule. Thus, the standard solution methods (e.g. den Haan, 2009; Krusell and Smith, 1998; Preston and Roca, 2007) propose to summarise this distribution by discrete and finite set of moments by assuming that agents make decisions based on 'bounded rationality' (Simon, 1982).[1] More specifically, if we consider only the first-order moment, the law of motion for aggregate

[1] Simon (1982) argues that people try to be rational, but their ability to be so is very limited, especially given the complexity of the world or given the prevalence of uncertainty. This implies that often, the main constraint on our decision-making is not the lack of information but our limited capability to process the information we have.

capital k_{t+1}, is given by

$$k_{t+1} = \zeta_0 + \zeta_1 k_t + \zeta_2 z_t. \tag{B.5}$$

Generally, if we denote capital distribution as Γ_t, the associated law of motion becomes

$$\Gamma_{t+1} = \mathbb{H}(\Gamma_t, z_t). \tag{B.6}$$

In sum, a rentier i's maximisation problem can be represented as a dynamic programming problem in which $a_{i,t}$, $e_{r,i,t}$, $\theta_{i,t}$, Γ_t and z_t are the state variables, while $c_{r,i,t}$ and $a_{i,t+1}$ are the decision variables. A rentier i's optimality equation is then given by

$$v(a_{i,t}, e_{r,i,t}, \theta_{i,t}; \Gamma_t, z_t) = \max_{c_{r,i,t}, a_{i,t+1}} \{u(c_{r,i,t}) + \beta E[v(a_{i,t+1}, e_{r,i,t+1}, \theta_{i,t}; z_{t+1}, \Gamma_{t+1}) - P(a_{i,t+1})]\} \tag{B.7}$$

subject to

$$c_{r,i,t} + a_{i,t+1} = \{1 + (1-\tau_r)(r^k(k_t, l_t, z_t) + (1-\lambda)\theta_{i,t}) - \delta\} a_{i,t} + w(k_t, l_t, z_t) e_{r,i,t} \tag{B.8}$$

$$z_{t+1} = (1-\rho^z)\mu^z + \rho^z z_t + \zeta^z_{t+1} \tag{B.9}$$

$$e_{r,i,t+1} = (1-\rho^e)\mu^e + \rho^e e_{r,i,t} + \zeta^e_{i,t+1} \tag{B.10}$$

$$\Gamma_{t+1} = \mathbb{H}(\Gamma_t, z_t) \tag{B.11}$$

where $v(\cdot)$ is a rentier i's value function. An equilibrium for this model then consists of the following:

1. Given price functions $\{r^k(\Gamma, z),\ w(\Gamma, z)\}$ and an aggregate function $\mathbb{H}(\Gamma, z)$, optimal decision rules $\{a' = h_a(a, e, \theta; \Gamma, z), c = h_c(a, e, \theta; \Gamma, z)\}$ are associated with a value function $v(a, e, \theta; \Gamma, z)$.

2. Given price functions $\{r^k(\Gamma, z)$ and $w(\Gamma, z)\}$, a firm maximises profits.

3. Markets clear:

   Labour market: $l = \lambda \int_{i \in R} e_{r,i} \bar{l} di + (1-\lambda) \int_{j \in W} e_{w,j} \bar{l} dj$

   Capital market: $k = \lambda \int_{i \in R} h_a(a, e, \theta; \Gamma, z) d\Gamma$

   Good market: $Y = \lambda \int_{i \in R} h_c(a, e, \theta; \Gamma, z) d\Gamma + (1-\lambda) \int_{j \in W} e_{w,j} \bar{l} dj + \delta k$

4. A transition law for the cross-sectional distribution of capital, $\mathbb{H}(\Gamma, z)$, is consistent with the individual policy function.

B.2 Solution methods

This study combines two solution methods: (i) the first-order perturbation method[2] that solves the individual policy function and (ii) the explicit aggregation algorithm that derives the aggregate law of motion for capital.

First, the perturbation method allows us not to be restricted in the low-dimensional-discrete-state Markov process for the stochastic components of the model, but to give a quick solution thanks to the presence of the standardised computational software (e.g. the Dynare). Second, the explicit aggregation algorithm, as developed by den Haan and Rendahl (2010), relies on the decision rules for directly evaluating the moments of the distribution without simulating the entire distribution.[3] This is done by explicitly aggregating the economy conditional on the properties of the approximated decision rule. As a consequence, the moments of the distribution that are included in the solution algorithm depend on the properties of the polynomial that approximates the decision rule. For instance, the aggregation of the first-order approximation of the economy will only account for the first-order moments of the cross-sectional distribution. The equation thus exists such that

$$a_{i,t+1} = \gamma_0 + \gamma_1 a_{i,t} + \gamma_2 e_{r,i,t} + \gamma_3 z_t + \gamma_4 k_t + \gamma_5 \theta_{i,t}. \tag{B.12}$$

This function determines the optimal allocation of individual assets for the next period based on current variables. Note that the coefficient γ_1 is the same for all possible values of $a_{i,t}$, which correspond to the property of approximate aggregation, as in Krusell and Smith (1998). To derive the aggregate law of motion for capital, we integrate equation (B.12) so that

$$\int_{i \in R} a_{i,t+1} di = \gamma_0 + \gamma_1 \int_{i \in R} a_{i,t} di + \gamma_2 \int_{i \in R} e_{r,i,t} di + \gamma_3 z_t + \gamma_4 k_t + \gamma_5 \int_{i \in R} \theta_{i,t} di. \tag{B.13}$$

[2] Krusell and Smith (1998) argues that it is sufficient for households to use only the mean of capital distribution to predict future prices because their marginal propensities to save are almost independent of wealth levels, except at the lowest wealth percentile. But even though the very poor have much high marginal savings propensities, their portions in total wealth will be negligible. Furthermore, Preston and Roca (2007) shows that the coefficients on the second-order terms are very small, such that the improvement in accuracy is only around two percent.

[3] According to den Haan (2009), this approach is simpler than the methods that rely on the parameterisation of the cross-sectional distribution, and it is not as computationally intensive as the methods that rely on simulation and regression.

Since $\lambda \int_{i \in R} a_{i,t+1} di = k_{t+1}$, $\lambda \int_{i \in R} a_{i,t} di = k_t$, $\int_{i \in R} e_{r,i,t} di = \mu^e \bar{l}$, and $\int_{i \in R} \theta_{i,t} di = \mu^{rp}$, we can rewrite equation (B.13) as

$$k_{t+1} = \lambda \underbrace{(\gamma_0 + \gamma_2 \mu^e \bar{l} + \gamma_5 \mu^{rp})}_{\zeta_0} + \underbrace{(\gamma_1 + \lambda \gamma_4)}_{\zeta_1} k_t + \underbrace{\lambda \gamma_3}_{\zeta_2} z_t \tag{B.14}$$

which gives us the aggregate law of motion for capital that is identical to equation (B.5).

Table B.1 Parameter values of the baseline model

Parameter	Value	Description
Structure		
β_r	0.985	subjective discount factor for the rentier class such that the average risk-free rate of interest is 4.67% at an annual frequency
β_w	0.960	subjective discount factor for the working class
γ_r	2	relative risk aversion coefficient for the rentier class
γ_w	5	relative risk aversion coefficient for the working class
α	0.33	average share of capital income based on the BLS
δ	0.022	depreciation rate such that average share of investment is 22%
λ	0.75	population share of the rentier class based on the DNAs
b	0	borrowing limit to prevent negative asset holdings
ϕ	0.05	barrier parameter based on Preston and Roca (2007)
ω	0.33	portion of positional competition based on the NIPAs
τ	0	tax rate on capital income
Prior stochastic processes		
μ^z	1	steady state of the aggregate TFP
μ^e	1	steady state of labour productivity
μ^{rp}	0	steady state of rentier premium
ρ^z	0.750	persistency of aggregate TFP shocks
ρ^e	0.700	persistency of labour productivity shocks
ρ^{rp}	0.900	persistency of rentier premium shocks
ρ_w^{ez}	0.450	cyclicality of labour productivity shocks for a worker
σ^z	0.013	volatility of aggregate TFP shocks
σ_r^e	0.100	volatility of labour productivity shocks for a rentier
σ_w^e	0.200	volatility of labour productivity shocks for a worker
σ^{rp}	0.100	volatility of rentier premium shocks

Notes: BLS, Bureau of Labour Statistics; DNAs, Distributional National Accounts; NIPAs, National Income and Product Accounts; TFP, Total Factor Productivity.

Table B.2 Statistical prior, posterior distribution and mode

parameter	prior mean	post mean	90% HPD interval	prior distribution	prior s.d.
ρ^z	0.750	0.755	0.619-0.911	beta	0.1
ρ^e	0.700	0.703	0.565-0.862	beta	0.1
ρ^{rp}	0.900	0.903	0.882-0.922	beta	0.1
ρ^{ez}	0.450	0.459	0.306-0.623	beta	0.1
σ^z	0.013	0.012	0.004-0.023	inverse gamma	infinity
σ^e_r	0.100	0.069	0.026-0.115	inverse gamma	infinity
σ^e_w	0.200	0.140	0.052-0.240	inverse gamma	infinity
σ^{rp}	0.100	0.094	0.089-0.099	inverse gamma	infinity

Note: HPD, highest posterior density.

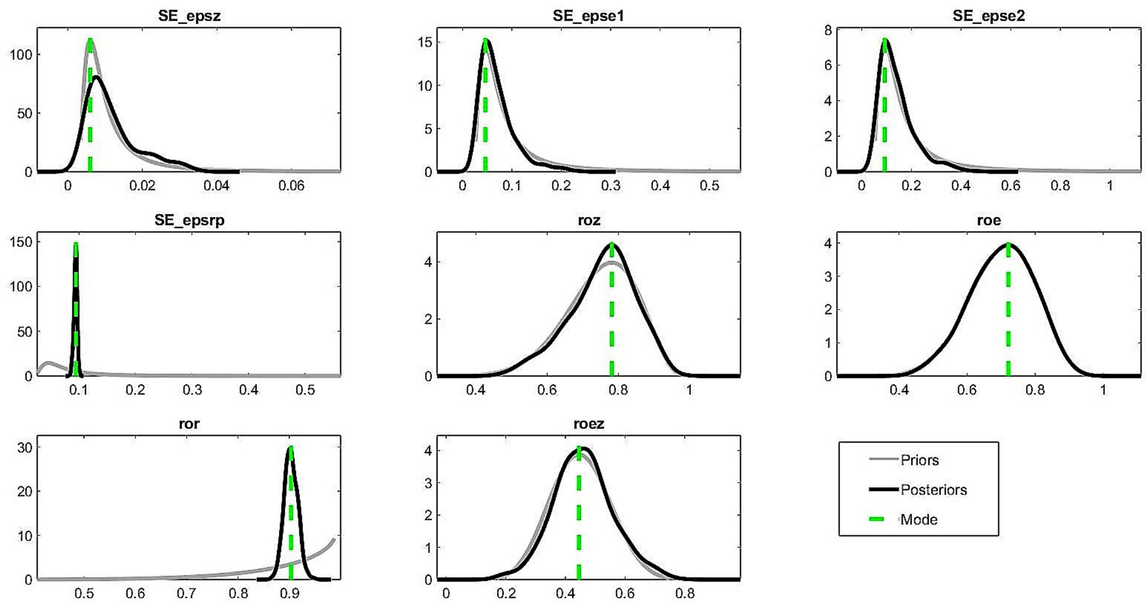
Table B.3 Distribution of returns on wealth: models and data

(%, %p)

				percentile of returns			fraction of	
	median	s.d.	skewness	10	25	90	r >0.3	r <0
US data	0.4	39.2	2.3	-20.2	0.0	39.2	13.4	44.2
Benchmark	3.3	-	1.0	3.3	3.3	3.3	-	-
Model 1	3.4	1.5	-1.2	0.0	0.0	3.4	-	-
Model 2	0.0	4.8	0.4	-3.7	0.0	8.5	-	24.9
Model 3	-1.5	19.9	-0.4	-32.9	-18.9	12.9	2.7	51.8

Notes: US data – the Panel Study of Income Dynamics recompiled by Cao and Luo (2017); Benchmark – heterogeneity in labour productivity; Model 1 – plus innate difference in asset ownership; Model 2 – plus heterogeneity in rentier premium; Model 3 – plus non-normality of asset market returns

Fig. B.1 Priors, posteriors and mode



Notes: SE, standard error; epsz, σ^z; epse1, σ^e_r; epse2, σ^e_w; epsrp, σ^{rp}; roz, ρ^z; roe, ρ^e; ror, ρ^{rp}; roez, ρ^{ez}

Appendix C

Appendix for Chapter 3

C.1 Simple model of rescheduling

As in Michl (2000), for simplicity, let us assume a production function in Cobb-Douglas form having no other inputs except hours worked and workers such that:

$$y = h^{\alpha} N^{\beta}, \qquad 0 < \alpha < \beta < 1 \tag{C.1}$$

where y is the output of the representative firm which hires N workers and schedules them to work h hours per week; and $\alpha < \beta$ implies that the proportional increase from hiring more workers exceeds the increase in output from working longer hours.

Total payroll hours, E, are the product of hours worked per worker and the number of workers such that:

$$E = hN. \tag{C.2}$$

The firm will minimises its costs, C, to produce its output level, $\bar{y}$, using the number of workers and their hours worked. The firm's optimisation problem is then described as:

$$\begin{aligned} & min\ C = M + F \cdot N + w \cdot h \cdot N \\ s.t \qquad & \bar{y} = h^{\alpha} N^{\beta} \qquad \text{and} \qquad y = \bar{y} \end{aligned} \tag{C.3}$$

where F is a fixed cost (e.g. recruiting costs, non-wage benefits, and training costs) for each worker; M is other fixed cost (e.g. a franchise or incorporation fee) to make the cost function consistent with an initial long run equilibrium having a finite number of firms; and w is hourly wage. From the first-order conditions for the equation (C.3), we can derive the following conditional demand functions for hours worked per worker and the number of workers such that:

$$h^* = \frac{\alpha F}{(\beta - \alpha) w} \tag{C.4}$$

$$N^* = (\bar{y}^{1/\beta})(h^*)^{-\alpha/\beta}. \tag{C.5}$$

Equations (C.4) and (C.5) imply that (i) an increase in the fixed cost of hiring workers will induce firms to lengthen the working week (i.e. $\partial h / \partial F > 0$) and shed workers (i.e. $\partial N / \partial F < 0$), and (ii) an increase in the cost of employing hours per worker will induce firms to reduce the working week (i.e. $\partial h / \partial w < 0$) and to take on additional workers (i.e. $\partial N / \partial w > 0$). These are pure subsitution effects that move the firms along a scheduling isoquant.

Using equations (C.2), (C.4), and (C.5), we can also derive a demand curve for total payroll hours. Taking logs gives:

$$lnE^* = (\beta - \alpha)/\beta lnh^* + 1/\beta ln\bar{y}. \tag{C.6}$$

Differentiating equation (C.6) with respect to the log wage yields the wage elasticity of total payroll hours, δ, such as

$$\delta = (\alpha - \beta)/\beta. \tag{C.7}$$

This implies that total payroll hours would decline after a wage increase because $\alpha < \beta$. Therefore, this simple model predicts that payroll hours declined after an increase in the minimum wage, as argued by Neumark and Wascher (2000), and (ii) the number of employees may have increased, as argued by Card and Krueger (1995), at the same time.

Furthermore, introducing a scale effect could explain why the minimum wage does have not have a statistically discernible effect on the number of workers. If the minimum wage increase affects an individual firm in a competitive industry, the increase in marginal costs with normal inputs will reduce the firm's optimal level of output, given that the product price remains constant. If, as is more relevant, raising the minimum wage affects all the firms in the industry, the increase in their marginal cost schedules will push prices up and reduce the demand for output. Therefore, these scale effects reduce the derived demand for total payroll hours. More specifically, from equations (C.4) and (C.5), we can see that this will be effected exclusively through a reduced demand for workers since from equation (C.4) it is apprarent that scale does not affect hours per week. That is, the substitution effect (i.e. increasing the demand for workers) and scale effect (decreasing the demand for workers) are offset. If the substitution and scale effects leave the number of workers unchanged, total payroll hours can decline because the substitution effect reduces the scheduled workweek of the existing workers. Nonetheless, the existing workforce can be working fewer hours per week, earning the same monetary income, and enjoying more leisure.

C.2 The description of the KLIPS

The Korean Labour and Income Panel Study (KLIPS) is the nation's only labour-related panel survey in South Korea, one that comprises the best of cross-sectional data and time-series data. It is conducted annually on a sample of 5,000 urban households and all members of the 5,000 households. It is held once a year, starting from Wave 1 in 1998 with the latest Wave 21 completed in 2018. The same set of questions are repeatedly asked annually to the same set of households and members to retain the original sample households.

There are largely two types of the KLIPS dataset: the household dataset derived from Household Questionnaires and the individual dataset compiled from Individual Questionnaires administered on the household members aged 15 or older. On the one hand, the household dataset includes demographics, changes in household members, family relations, children's education and childcare, household income and expenditure, assets and debts, financial status, and consumption requirements that put pressure on the household finance. On the other hand, the individual dataset includes a wide array of categories such as the person's state of economic activity, income-earning activities and consumption, education and vocational training, employment characteristics, work hours, professional and life satisfaction, job-seeking activities, and labour market mobility.

The survey method is the interviewer-administered in principle, for which the interviewer verbally asks the questions to the interviewee and records the responses. But those who are not available for interview owing to serving in the military or placed in social protection facilities, they are excluded from the Individual Questionnaires.

The interviews have been conducted by the same professional survey agency from 1998 onwards. The survey peiod is around six months from April to September each year, administered by around 100 professional interviewers. Training for interviewers is provided in a systematic manner in each region based on the content of the questionnaires before beginning the annual survey. About 20–30% more interviewers are trained than actually needed to create a buffer for those who drop out due to the challenges of conducting a panel study. To ensure the highest possible accuracy for the already-collected data, all of the surveys gathered have been reviewed. For incomplete responses, excessively high rate of nonresponse, and logical inconsistencies between responses, the interviewer was required to readminister the survey.

Table C.1 Descriptive statistics for socio-economic characteristics of each group

(%)		Affected workers	Unaffected workers	Non-target workers
Gender	Male	42.5	61.6	72.2
	Female	57.5	38.4	27.8
Marital Status	Unmarried	26.9	25.4	18.7
	Married	73.1	74.6	81.3
Age	15-29	19.6	14.9	8.7
	30-59	59.3	71.8	85.4
	60+	21.1	13.3	5.9
Education attainment	High school	67.8	71.9	42.4
	College	12.7	15.3	19.0
	University+	19.5	12.8	38.5
Employment Status	Regular	55.3	68.0	84.7
	Non-regular	44.7	32.0	15.3
Firm size	1-9	45.1	42.8	22.9
	10-299	41.8	46.2	49.6
	300+	13.1	11.0	28.5
Industry classification	Agriculture	2.5	1.0	0.6
	Manufacture	17.2	21.8	25.2
	Utilities	6.7	6.4	13.0
	Service	73.7	69.7	61.2

Notes: Affected workers are those whose current wage is below the new minimum wage. Unaffected workers are those whose current wage is between the new minimum wage and the living wage (1.22 times the minimum wage). Non-target workers are those whose current wage is greater than the living wage. Agriculture includes forestry, fishing, and mining, and Utilities stand for electricity, gas, water, and construction.

Table C.2 Results of fuzzy RD designs

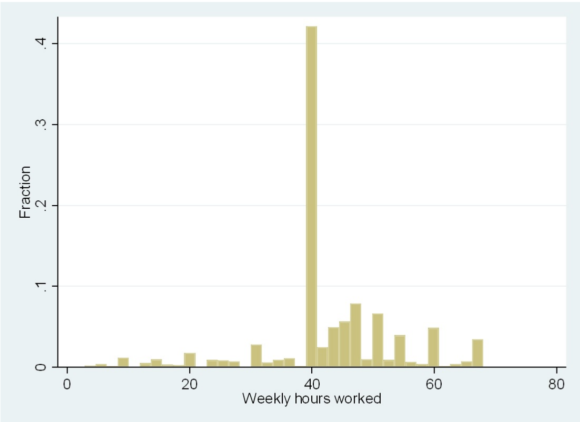
RD effect	Robust standard errors	Observations
-31.813	98.624	355

Table C.3 Results of local randomisation RD designs

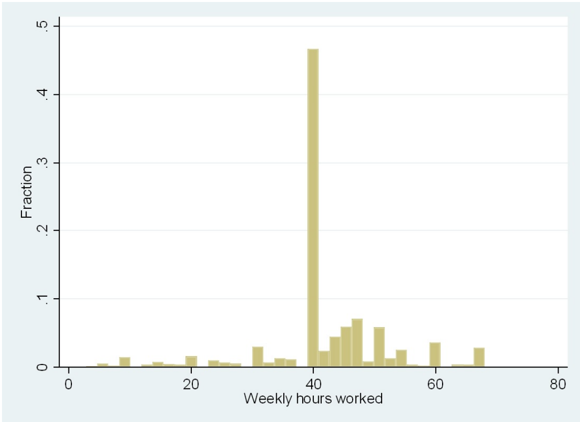
RD effect	(Fisher) p-value	N_c^-	N_c^+
-17.932	0.078	19	16

Note: N_c^- (N_c^+) implies the number of observations at the left side (right side) of the cut-off

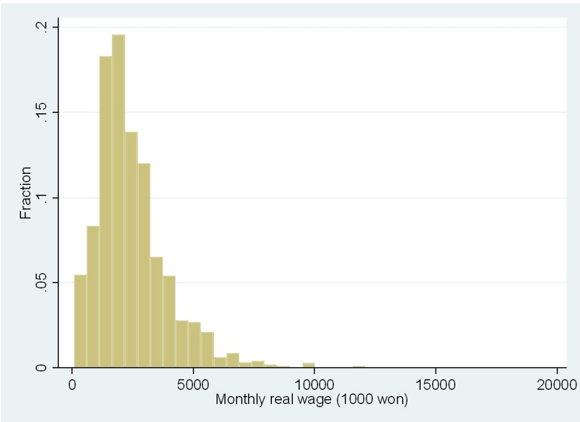
Fig. C.1 Descriptive statistics before and after the 2018 minimum wage increase



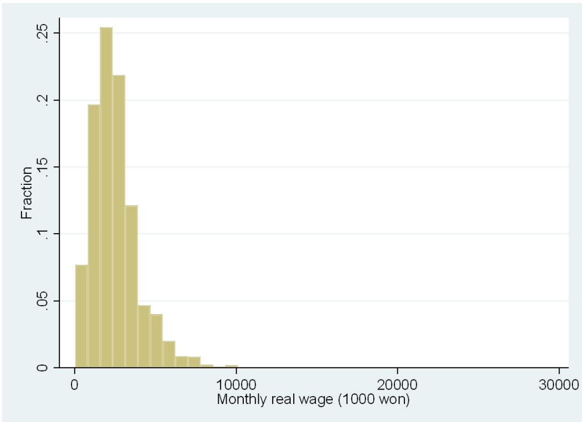
(a) Distribution of hours worked in 2017



(b) Distribution of hours worked in 2018

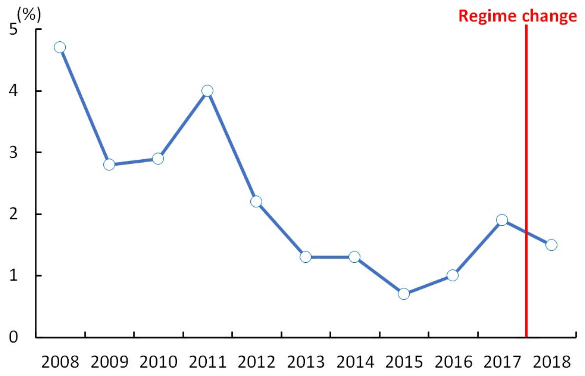


(c) Distribution of real wages in 2017

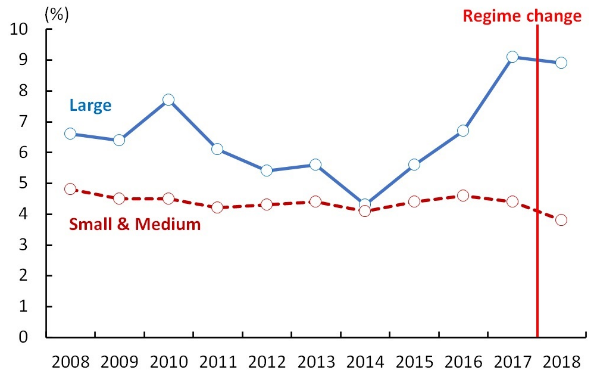


(d) Distribution of real wages in 2018

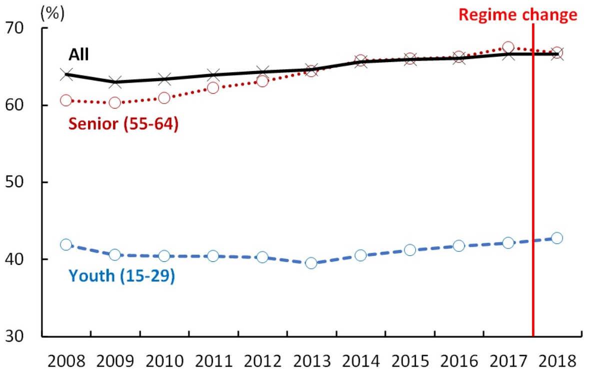
Fig. C.2 Macroeconomic trends in South Korea



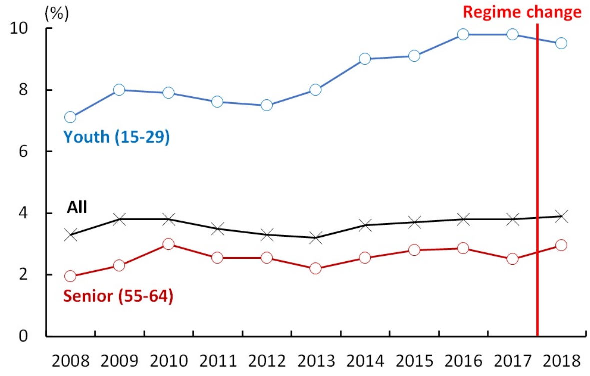
(a) Inflation rate



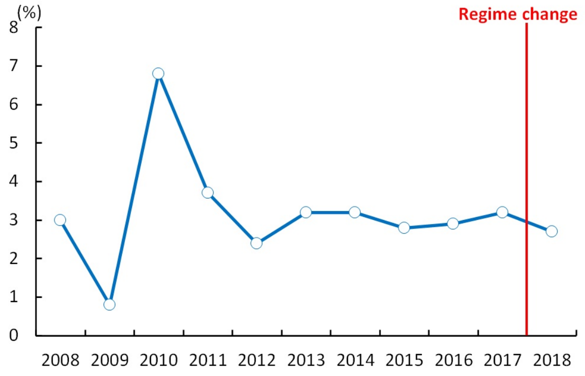
(b) Profits – return on assets



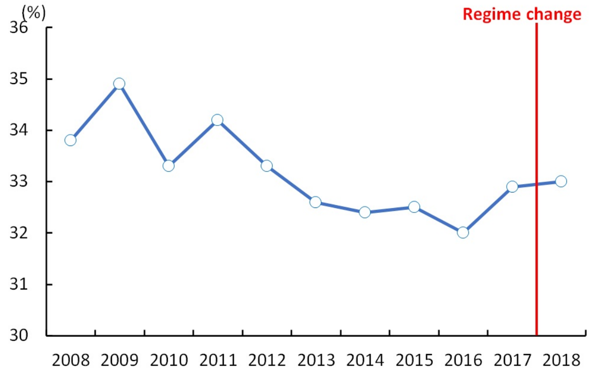
(c) Employment rate by age



(d) Unemployment rate by age



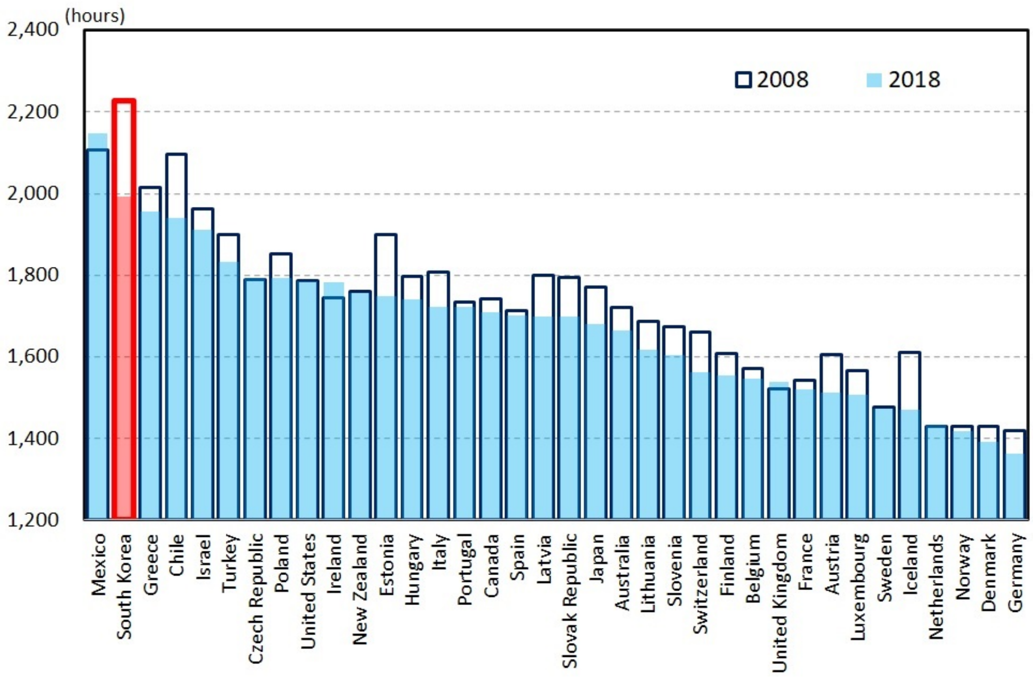
(e) Real GDP growth rate



(f) Non-regular employment ratio

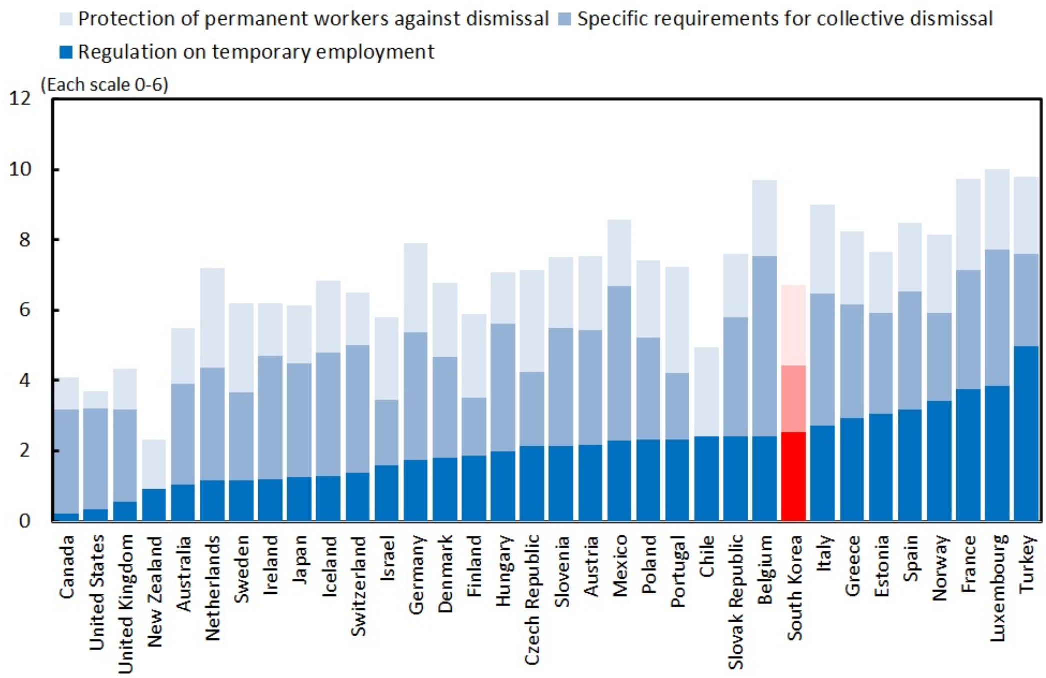
Source: Statistics Korea (2020)

Fig. C.3 Average annual hours worked per worker in OECD countries



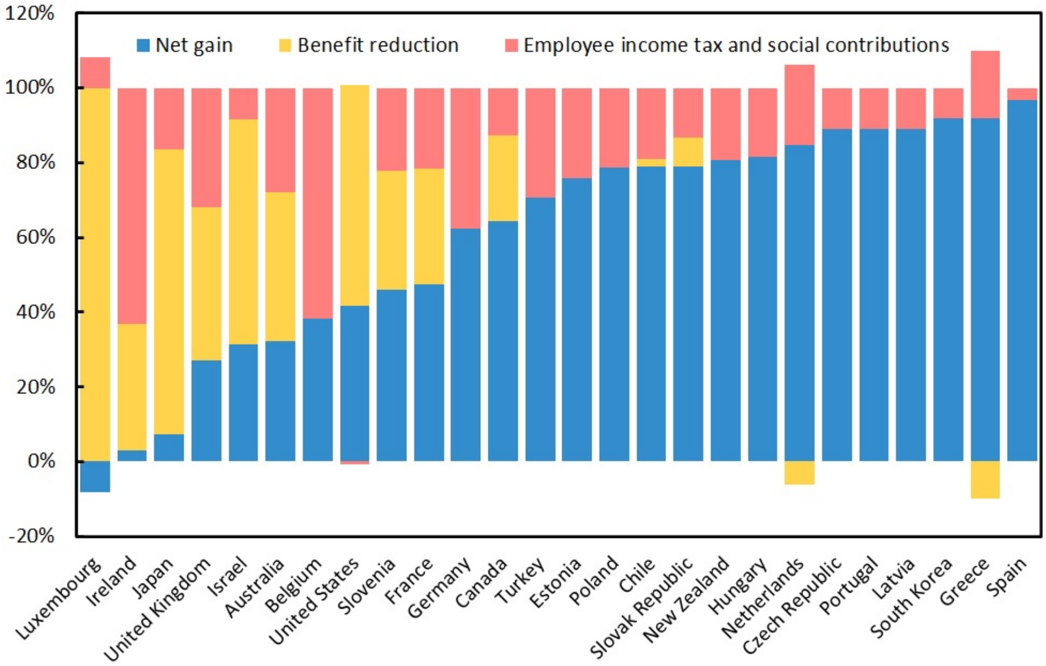
Source: OECD.stat (2020)

Fig. C.4 Strictness of employment protection legislation in OECD countries



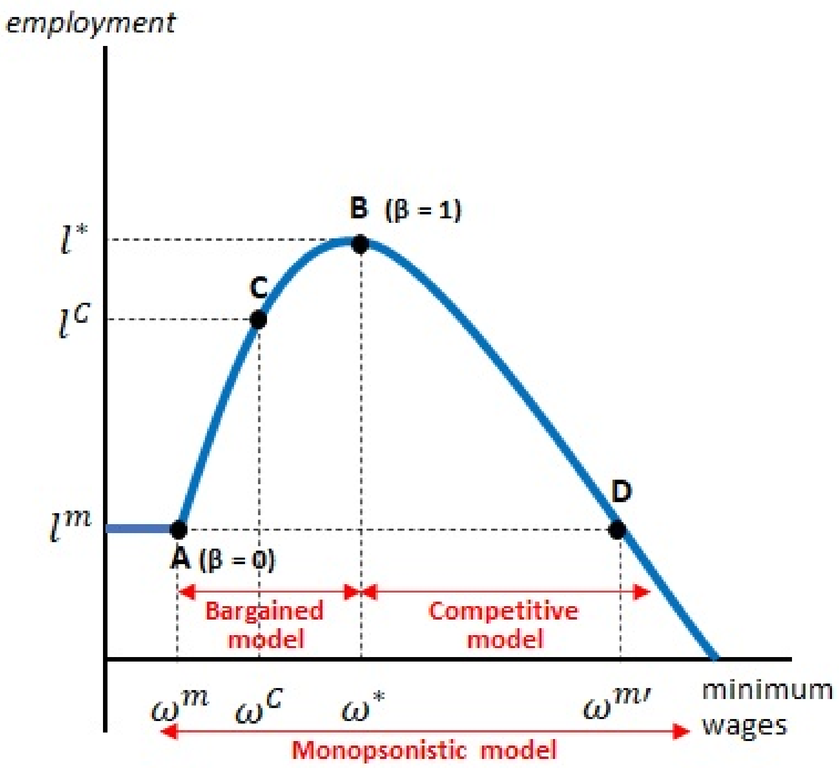
Note: Each scale is ranged from 0 (least restrictions) to 6 (most restrictions) in 2013.
Source: OECD.stat (2020)

Fig. C.5 Share of a minimum wage increase that adds to net income in OECD countries



Notes: Calculations refer to a 5% minimum wage increase and a single-adult household with two children in 2013, assuming that all tax and benefit provisions remain as they were before the increase, and accounting for minimum income and other means-tested benefits that are primarily income related and are typically accessible for low-income families. The minimum wage in Germany is for 2015.
Source: OECD (2015), utilising OECD tax-benefit models and minimum wage database

Fig. C.6 Prediction of employment effects under the bargained minimum wage model



Notes: Employment implies total payroll hours (the number of workers × average workweek) and β represents the bargaining power of employees.