UK economic growth and inequality from 1870

Hankui Wang

A thesis submitted for the degree of
Doctor of Philosophy
Cardiff University
September 2024

Acknowledgments

First and foremost, I extend my deepest gratitude to my supervisors, Professor Patrick Minford, Professor Vo Phuong Mai Le and Professor David Meenagh for their unwavering support, invaluable guidance, and consistent encouragement throughout the research and writing process. Your expertise and insights have been instrumental in shaping this work. I am indebted to the members of my thesis committee: Professor Parantap Basu, and Dr Peng Zhou, for their valuable feedback, constructive criticism, and dedication to ensuring this research met the highest standards.

To the faculty and staff Cardiff University, especially those in the CARBS PhD office and IT team, thank you for providing a nurturing environment and the necessary resources that enabled me to focus on my research.Special thanks go to my peers and colleagues, particularly Zhen Liu, Wenjing Wang, Liuyu Yang, Yixuan Mao, Yuqi Zheng, Jiachen Yi, and John Peirce for their camaraderie, collaborative spirit, and for the countless enlightening discussions that enriched this endeavour.

I would be remiss if I didn't acknowledge the unwavering support of my family. To my parents, thank you for your endless patience, love, and belief in my capabilities.

Lastly, I wish to express gratitude to all those who provided technical support, data, or any form of assistance throughout this journey. Every conversation, critique, and word of encouragement played a role in bringing this thesis to fruition.

Without the collective effort and support of each of you, this work would not have been possible.

Abstract

This study introduces a structural model of economic growth and inequality, grounded in the optimizing behaviour of entrepreneurs. The model integrates heterogeneous agents and examines how entrepreneurial incentives, influenced by individual wealth and credit conditions, endogenously impact wealth distribution and economic growth. Unlike traditional methods, this research employs indirect inference to test the model, focusing on its ability to simulate behaviour consistent with observed data rather than predicting specific outcomes. Using comprehensive UK data spanning from 1870 to 2016, the empirical findings reveal a trade-off between growth and inequality. The study's welfare analysis implies that the historical trend towards redistribution is driven by its popularity and potential to balance growth and inequality, providing crucial insights for policy formulation.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This thesis proposes a model of economic growth and inequality, grounded in the optimizing behavior of entrepreneurs. This model uniquely connects growth to entrepreneurial incentives and inequality, differing from existing models (see 2.3 for details). Additionally, the thesis employs an innovative empirical method—indirect inference—to test the model. Unlike traditional methods that rely on the model's predictive ability, indirect inference checks whether the simulated behaviour of the model aligns with the actual data behaviour, providing a robust test of the model's validity.

1.1 Introduction

Embarking on a journey through the annals of economic history, one is inevitably drawn to the intricate interplay between economic inequality and growth, an enduring theme that has captivated scholars and policymakers alike. From the opulent peaks of prosperity to the stark valleys of disparity, the tapestry of human economic endeavour is woven with the threads of inequality. As societies traverse the epochs, transitions in economic growth often herald shifts in the distribution of wealth and opportunity, casting a spotlight on the complex dynamics shaping human welfare.

To dissect the nexus between economic inequality and growth, our exploration ventures into the empirical realm, where data serves as our guiding compass. Three primary series stand as pillars in our quest: real GDP growth, income share, and wealth share. Real GDP growth, a stalwart indicator of economic vitality, traces the fluctuations in a nation's aggregate output over time, offering a lens into the overarching trajectory of prosperity. Complementing this macroeconomic perspective, the income share unveils the distributional nuances within societies, shedding light on how the spoils of growth are apportioned among different segments of the population. Meanwhile, the wealth share delves even deeper, illuminating the concentration of assets and resources among the most affluent echelons, underscoring the intricate interplay between economic growth and the accumulation of wealth. We've chosen to analyze UK data due to its rich economic history, and long-term data availability. The UK's journey offers insights into how economic inequality and growth intersect. With comprehensive data covering GDP growth, income distribution, and wealth concentration, we aim to inspire broader insights beyond national borders.

As shown in Figure 1.1, the UK's real GDP growth from 1870 to 2016 reflects a dynamic economic trajectory characterized by periods of expansion, contraction, and transformation. Over this span of more than a century, the UK experienced significant economic milestones, including the Industrial Revolution, two World Wars, and various shifts in economic policy and global market dynamics. During the late 19th and early 20th centuries, the UK witnessed robust economic growth driven by industrialization, urbanization, and imperial expansion and this period marked the zenith of the British Empire, with the UK serving as a global economic powerhouse. However, the UK's economy faced challenges and setbacks, including the Great Depression of the 1930s and the devastation of World War II. These events led to economic contractions and restructuring efforts as the nation rebuilt its infrastructure and industries. In the post-war period, the UK experienced a period of sustained growth, often referred to as the "post-war economic boom." This era was characterized by government intervention, welfare state expansion, and the rise of industries such as manufacturing and services. However, by the late 20th century, the UK's economic landscape underwent significant transformations. Thatcherite economic policies in the 1980s aimed to deregulate markets, privatize state-owned enterprises and promote free-market principles. While these reforms sparked periods of economic growth and innovation, they also led to socio-economic challenges such as rising inequality and deindustrialization in certain regions. In the 21st century, the UK's economy continued to evolve amidst globalization, technological advancements, and shifting geopolitical landscapes. Economic growth fluctuated in response to global financial crises, and Brexit uncertainties.

Regarding economic inequality, during the late 19th and early 20th centuries, the UK experienced widening income inequality as industrialization accelerated, concentrating wealth among the upper echelons of society. The rise of large-scale industries and a wealthy capitalist class significantly increased income disparity.

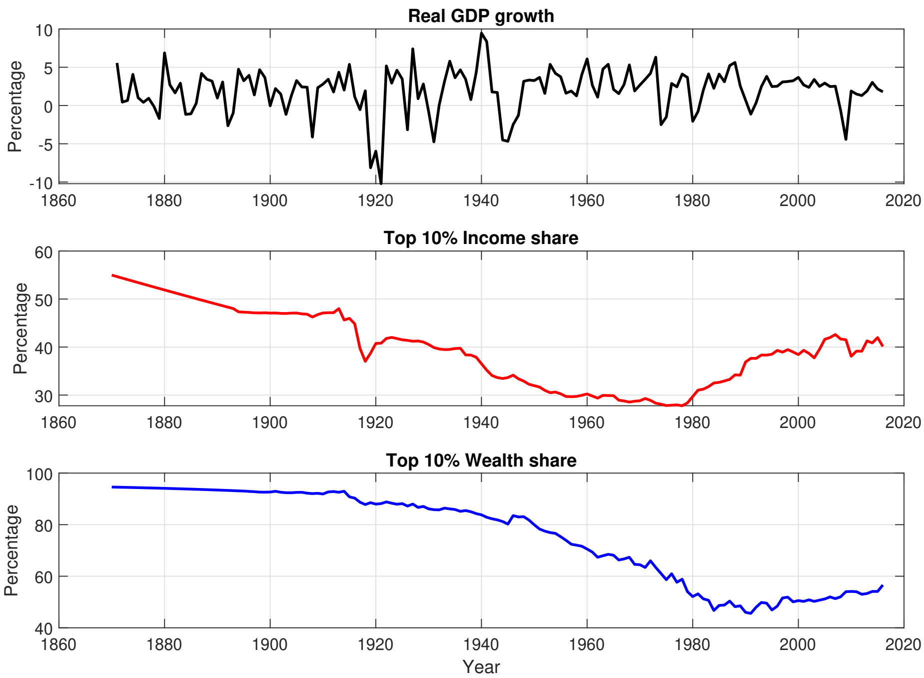


Figure 1.1. Economic inequality and growth in the UK: 1870-2016

However, the interwar period, marked by economic turbulence and the Great Depression, saw some levelling of inequality as the downturn disproportionately affected wealthier individuals and industries. Government interventions through social welfare programs and progressive taxation aimed to mitigate disparities and provide relief to those most impacted. In the post-WWII era, the UK shifted towards greater social equality, driven by welfare state policies, progressive taxation, and labour market reforms that expanded education, healthcare, and social services, improving living standards and redistributing wealth more equitably. From the 1980s onwards, under Margaret Thatcher's government, neoliberal policies such as dereg-

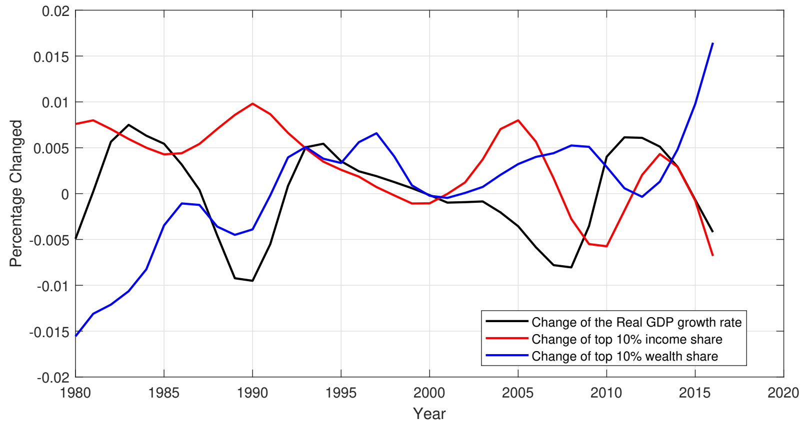


Figure 1.2. Economic inequality and growth change in the UK: 1980-2016

ulation, privatization, and tax cuts favoured the wealthy, leading to a resurgence of income inequality. Globalization, technological advancements, and structural economic changes further contributed to this polarization. In the 21st century, income inequality remained a prominent issue, influenced by financial crises, policy changes, and global economic shifts. Efforts to address inequality through measures such as minimum wage laws, progressive taxation, and social welfare programs continue, reflecting ongoing societal concerns about fairness and social cohesion. From Figure 1.1, the trajectories of both income and wealth shares for the top 10% exhibit a consistent decline from 1870 to 1980, followed by an upward trend in the subsequent four decades, driven by returns on investments. A notable pattern is the decline in the wealth share of the top 10% post-1950, during the 'Golden Age' of capitalism, with wealth becoming more equitably dispersed. Post-1980, the wealth share shows greater volatility than the income share, though both have trended upwards.

While Figure 1.1 depicts the individual dynamics of growth and inequality, discerning their relationship or interplay is challenging. However, by examining the differences or changes in these series, a compelling narrative emerges. We separate the full sample into three subsamples,1873 - 1914, 1946 - 1979, and 1980 - 2016[1]. Britain's economic history reveals intriguing episodes where shifts in economic growth align

[1]We used the HP filtered data here just to show where the motivation comes from, we didn't use the filtered data in the empirical analysis because by filtering one lose useful information. If one were to use HP-filtered data in modelling, the model would only have stationary shocks, and one wouldn't be able to get long-run divergence in inequality.

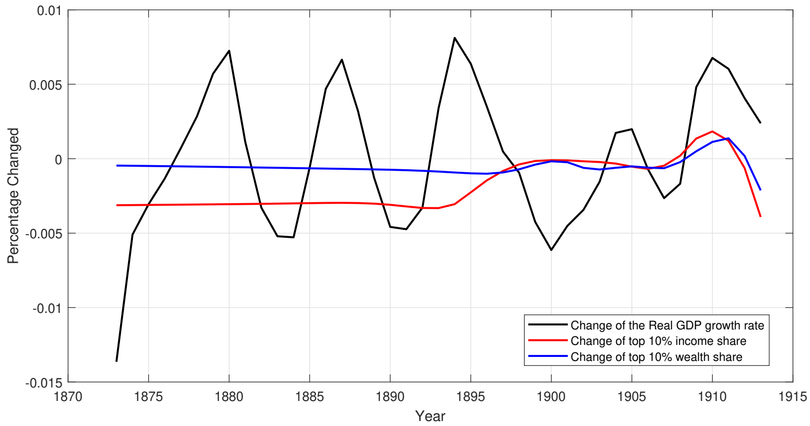


Figure 1.3. Economic inequality and growth change in the UK: 1873-1914

with corresponding changes in economic inequality, suggesting a nuanced and intertwined relationship between the two phenomena. Particularly captivating is the recent period between 1980 and the mid-1990s as shown in Figure 1.2, when the nation experienced a robust economic acceleration driven by a slew of supply-side reforms. Under Margaret Thatcher's leadership, her economic doctrine, known as Thatcherism, unfurled a tapestry of deregulation and supply-side transformations. Deregulation endeavoured to prune state intervention, fostering competition and efficiency in sectors like telecommunications, transportation, and finance. The privatization of state behemoths like British Telecom and British Airways symbolized pivotal milestones, while financial deregulation, epitomized by the Big Bang reforms, globalized the financial realm. Supply-side measures, including tax cuts to spur investment and entrepreneurship, accompanied efforts to temper trade union influence and overhaul welfare. These initiatives catalyzed economic dynamism, moulding a more market-centred economy, and leaving an indelible imprint on both the UK's economic landscape and global economic discourse.

Delving deeper into the annals of pre-WWI history, the period from 1905 to 1914, as shown in Figure 1.3, unveils a striking correlation between economic prosperity and disparity, exemplified by the governance of three influential Liberal Prime Ministers: Henry Campbell-Bannerman, H.H. Asquith, and David Lloyd George (Chancellor in office: 1908-1915, PM in office: 1916-1922). Each of these leaders wielded considerable influence in shaping regulatory policies, particularly in pivotal industries and social spheres. Campbell-Bannerman's administration, inaugurating this era, exhibited a keen focus on social welfare. Initiatives like the Education (Provision of Meals) Act of 1906, providing free meals for underprivileged schoolchildren, underscored a commitment to alleviating social inequalities. Asquith's tenure, characterized by significant regulatory interventions, responded adeptly to the burgeoning social and economic challenges stemming from rapid industrialization. Noteworthy legislations such as the Trade Boards Act of 1909 exemplified this commitment, laying the groundwork for fair labour practices and addressing healthcare and social insurance concerns. The crowning achievement of this era's regulatory landscape was perhaps the National Insurance Act of 1911, a landmark piece of legislation under Asquith's leadership. This act provided workers with comprehensive insurance coverage against sickness and unemployment, heralding a transformative stride towards the establishment of a welfare state. These regulatory endeavours were not only responsive to the pressing needs of the time but also laid the cornerstone for a more equitable and socially inclusive society. In essence, the regulatory policies implemented during the tenure of these Liberal Prime Ministers not only bolstered economic stability but also fostered a more just and compassionate society.

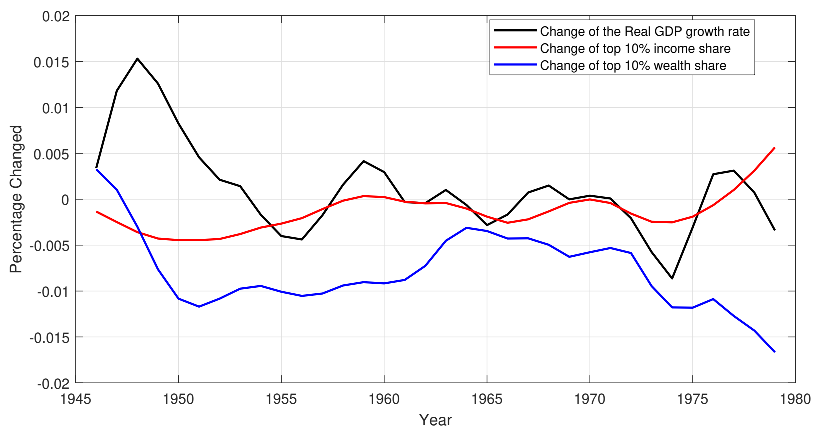


Figure 1.4. Economic inequality and growth change in the UK: 1946-1979

A similar trend was also found in the post-WWII, as shown in the Figure 1.4. Between 1945 and 1979, the UK pursued an economic policy marked by extensive regulation and limited deregulation. Influenced by Keynesian economics and the post-war consensus, the government implemented robust regulatory measures aimed at stabilizing the economy and promoting social welfare. This period witnessed the nationalization of key industries like coal and railways, alongside the establishment of the welfare state, exemplified by the creation of the National Health Service (NHS). Labour market regulations, including minimum wage laws and collective bargaining rights, aimed to safeguard workers' rights and ensure fair employment practices. Although some minor deregulatory efforts were made, particularly under the Conservative government of Harold Macmillan, they were modest in scope compared to the prevailing regulatory framework. Overall, the regulatory approach of this era reflected a commitment to state intervention and collective welfare, prioritizing economic stability and social equity over market-oriented reforms, which was in line with the prevailing economic ideologies of the time.

Table 1.1. Correlations of Inequality-change and Growth-change

Correlations	1873-1914	1946-1979	1980-1999	2000-2016
ΔWealth corr. ΔGrowth	0.1332	0.3315	0.1164	-0.4057
ΔIncome corr. ΔGrowth	0.1210	-0.2771	-0.5083	-0.0112
ΔIncome corr. ΔWealth	0.7744	-0.2760	-0.5940	-0.4301

To refine our understanding further, we compute correlation coefficients between shifts in economic growth and alterations in inequality. Table 1.1 presents the correlation coefficients among the change of wealth share, income share and economic growth. The results show that, before World War I, all three variables—wealth share, income share, and economic growth—were positively correlated. This indicates that increases in wealth and income shares were associated with higher economic growth during this period. However, the dynamics shift significantly after World War II.

While the income share's correlation with economic growth turns negative, indicating that higher income inequality is associated with lower growth, the wealth share of the top 10% maintains a positive relationship with economic growth until 2000. This suggests that wealth concentration among the top 10% continued to drive economic growth in the post-war period, even as income inequality posed a hindrance.

These findings bolster our objective to model the relationship between capital inequality and economic growth in the UK, extending our data back to 1870. By understanding these historical correlations[2], we aim to develop a more nuanced and comprehensive model that captures the intricate dynamics between wealth concentration and economic performance over an extended period. The same behaviour in different regulation periods also triggered the curiosity to explore the welfare implications beyond the growth and inequality nexus.

1.2 Thesis Outline

This thesis investigates the long-term interplay between economic growth and inequality in the United Kingdom by employing an endogenous growth model that incorporates heterogeneous agents. The empirical scope is broadened to encompass a more extensive timeline of UK history, utilizing the Millennium of Macroeconomic Dataset from the Bank of England and inequality data from the World Inequality Database, spanning 1870 to 2016. Through Indirect inference estimation, the structural model adeptly fits the intricate dynamics of the inequality-growth relationship and captures key characteristics of the UK's long-term data. The nuanced trade-off between inequality and growth is meticulously dissected, accompanied by a comprehensive welfare analysis conducted via Monte Carlo simulations.

As for the structure of this thesis: following this introduction, Chapter 2 delves into a literature review. Chapter 3 offers a concise overview of the model. Chapter 4 focuses on data sources and their statistical interpretations. Chapter 5 outlines the

[2]It must be emphasised that our estimation and testing of the causal relationship does not rely on the direct correlation of the time series- such a correlation lacks identification- but rather on the replication by the (identified) causal model indirectly of the relationships found in the data, whatever they may be. This method of Indirect Inference(see Chapter 5.2) is a powerful test of causal models such as we propose here.

primary methodologies employed for empirical exploration. Chapter 6 sheds light on empirical findings, encapsulating estimation, simulation, and testing results. Finally, Chapter 7 provides a conclusion.

Chapter 2

Literature Review

The substantial rise in income inequality in recent decades has reignited debates about its relationship with economic growth (known as the *inequality-growth nexus*), especially following the popularity of Thomas Piketty's book *Capital in the Twenty-First Century* (Piketty, 2014). Policymakers are particularly interested in understanding whether reducing inequality can lead to faster and more sustainable economic growth. This literature review aims to examine the theoretical and empirical evidence on the relationship between inequality and growth, providing insights to inform policy decisions in this crucial area. Broadly, the extant literature on this subject can be categorized into four distinct streams based on the inferred correlation between inequality and growth: positive, negative, synthetic(e.g. hump-shaped), and ambiguous(or inclusive). Methodologically, these studies predominantly adopt one of two approaches from the perspective of modelling: the *structural form*, which investigates different theoretical channels, and the *reduced form*, which relies on varied empirical estimation techniques and datasets.[1]

This literature review[2] is structured in three sections: Section 2.1 provided a brief overview of reduced form empirical research findings related to the inequality-growth nexus. Section 2.2 explores the structural models with different theoretical channels. Section 2.3 gives critical remarks on current structural models and the gaps that this thesis fills.

[1]There is literature combining two methods, e.g. Townsend and Ueda (2006), Basu and Guariglia (2007), Noh and Yoo (2008).

[2]More survey literature, see Aghion et al. (1999), Ehrhart et al. (2009), Mdingi and Ho (2021), Foellmi and Baselgia (2022)

2.1 Reduced form empirical research

The reduced form empirical literature presents varied results, depending on the empirical estimation techniques and datasets used. Cross-country studies reveal differing perspectives: while Perotti (1996), Forbes (2000), and Berg et al. (2018) argue for a positive relationship between inequality and growth, Alesina and Rodrik (1994) suggest that inequality hampers growth, and Bagchi and Svejnar (2015) also find a negative relationship. Barro (2000, 2008) further complicates the picture by identifying a nonlinear relationship, where the impact of inequality on growth varies with income levels. These conflicting findings underscore the complexity of the inequality-growth nexus and suggest that the relationship may be context-dependent. Panel data studies, such as Castelló and Doménech (2002) and Castelló-Climent (2010), emphasize the importance of controlling for country-specific effects and initial conditions, often finding a negative impact of inequality on growth. Similarly, Persson and Tabellini (1994) find a negative relationship between inequality and growth in their study of 56 countries over the period 1960–1985. Chambers and Krause (2010) observe that as returns to human capital rise relative to physical capital, inequality becomes more harmful to growth, though this pattern does not hold in well-educated nations. For a comprehensive meta-analysis of reduced-form empirical studies during 1994–2014, see Neves et al. (2016). Reduced-form empirical research on the inequality-growth nexus faces significant challenges, particularly regarding endogeneity and causality. For example, high growth could lead to increased or decreased inequality, making it difficult to determine the direction of causality. Simultaneous equations models, as discussed by Lundberg and Squire (2003) and Lin et al. (2009), attempt to address this issue, but these approaches also have limitations. Additionally, measurement and data inconsistencies across countries and time periods further complicate comparisons. Inequality and growth data can be inconsistent across countries and time periods, making comparisons difficult. Different measures of inequality (e.g., Gini coefficient, income shares, wealth shares) may yield different results. Especially, in many developing countries, data on inequality is sparse or unreliable, leading to potential measurement errors. Countries also differ significantly in terms of institutions, culture, policies, and stages of development(Scholl and Klasen, 2019). These differences can influence the inequality-growth relationship, making it hard to generalize findings. The relationship between inequality and growth may change over time within the same country, depending on economic, social, and political changes. The divergent findings on the inequality-growth nexus underscore the need for more nuanced theoretical models that account for varying institutional contexts and policy environments. Integrating insights from endogenous growth theory with political economy models could provide a more comprehensive understanding of how inequality influences growth through multiple channels.

2.2 Structural modelling research

In contrast to the extensive reduced-form estimations concerning the relationship between economic inequality and growth, the literature on the structural modelling of this relationship is comparatively limited. They predominantly follow a typical causal chain reasoning structure, exploring how inequality impacts economic growth through specific mechanisms or mediating variables.[3] Additionally, some studies use moderator variables to examine how the strength of the relationship between inequality and economic growth is influenced by specific factors, such as political institutions and redistribution policies.

Early literature (Kaldor, 1955; Bourguignon, 1981) argue that the marginal propensity of ***saving*** is an increasing function of wealth, which means the rich save relatively more than the poor, therefore inequality positively affects ***capital accumulation*** and hence economic growth accordingly.[4] Later, a big stream of literature ex-

[3]There is also literature that focuses on reverse causality, which is not directly related to this thesis and therefore is not extensively analyzed in this literature review. These studies collectively highlight the complex dynamics between entrepreneurship, wealth distribution, and economic growth, offering additional insights that can inspire readers. Notably, much of the research following Piketty's work concentrates on explaining the determinants of wealth and income disparities. For instance, Aoki and Nirei (2017) developed a neoclassical growth model that replicates historical income and firm size distribution patterns, providing insights into how productivity fluctuations at the firm level affect wealth and income distribution. Aghion et al. (2019) shows innovation and entrepreneurial shares can explain the inequality and growth dynamics in the US over the last decades. Atkeson and Irie (2020) focus on the evolution of family-owned businesses in the U.S., illustrating how shocks to family business values contribute to wealth mobility and disparity, they outline a wealth distribution path that aligns closely with earlier research, such as Saez and Zucman (2016) and Gomez (2023).

[4]It also can be argued that the higher marginal propensity to save among wealthier individuals could

pands capital accumulation from physical investment to ***human capital*** investment. The fundamental premise is that human capital is the primary driver of economic growth.[5] These models aim to explore the factors influencing human capital investment or accumulation, e.g., family size (Galor and Zang, 1997; De La Croix and Doepke, 2003) and foreign direct investment (Basu and Guariglia, 2007), providing insights into variations in economic growth. They typically focus on how income or wealth inequality impacts individuals' access to education (Benabou, 1993, 1994, by occupational segregation) and incentives to invest in education and skills. When people face borrowing constraints or fixed costs in investment, these investments are influenced by parental income and intergenerational transfers (Loury, 1981), therefore the initial distribution of human capital plays a crucial role in shaping the long-run correlation between the growth rate and income inequality (Bandyopadhyay and Basu, 2005).

By assuming those ***credit market imperfections***, the poor are unable to sufficiently invest in their human capital, while the rich, who do not need to borrow, can afford to make such investments. Since it is improbable that there is a perfect correlation between ability and wealth, wealth inequality results in poor under-investing in human capital, this under-investment negatively impacts the overall level of human capital and consequently hampers economic growth in both the short and long-term (Galor and Zeira, 1993). When access to borrowing is limited, promising business ideas might not be realized (Foellmi and Oechslin, 2010), or firms may not adopt more productive technologies, thereby reducing long-term growth (Foellmi and Oechslin, 2020).

However, increasing inequality necessitates ***redistribution***[6] policies to mitigate income disparity and enhance social welfare to avoid social unrest and revolution (Acemoglu and Robinson, 2000, 2002). Typically, these policies transfer resources from

lead to over-saving and under-consumption, thereby stifling aggregate demand and growth (Carroll, 1998).

[5]Human capital can also have a dual role in shaping the relationship of the relationship between growth and inequality, see Eicher and Garcıa-Penalosa (2001).

[6]Apart from redistribution, other factors such as financial constraints (Cagetti and De Nardi, 2006) and establishment costs (Restuccia and Rogerson, 2008) faced by the firms and entrepreneurs, can also be ***regulatory barriers*** to impede entrepreneurial motivations, technology adoption (Parente and Prescott, 1994, 2002) and therefore economic growth.

wealthier segments to poorer ones through taxes (e.g., estate taxation in De Nardi and Yang, 2016), social security, and public services. Such measures are crucial for fostering social equity and alleviating poverty, though they also spark debates about economic efficiency—the optimal use of resources to maximize output and welfare. Redistribution can enhance economic efficiency and spur economic growth by boosting aggregate demand. The lower-income households may have a higher marginal propensity to consume than wealthier ones. Thus, redistributing income to these households increases their consumption, which can elevate overall economic demand and potentially drive growth. This perspective underscores the stimulative impact of increased consumption on the broader economy. Moreover, redistribution can promote growth by reducing the poor's need to borrow for investment, thus minimizing distortions in profit-maximizing incentives, promoting equality of opportunity, and accelerating the trickle-down process (Aghion and Bolton, 1997). It is also suggested that raising taxes to subsidize new investments can enhance the growth rate, labour supply, and before-tax income inequality (García-Peñalosa and Turnovsky, 2007; García-Peñalosa and Wen, 2008) while reducing overall inequality and improving human capital accumulation by investing in public infrastructure services[7] (Getachew, 2010). Moreover, public R&D financed by income tax can boost growth and welfare by positively affecting individual savings and effort (Basu and Getachew, 2020; Minford and Meenagh, 2019). However, critics argue that the high taxes and transfers required for significant redistribution can reduce both working and entrepreneurial motivation[8] and create ***disincentives*** for productivity and investment, potentially reducing economic growth(Alesina and Rodrik, 1994; Persson and Tabellini, 1994; Alesina and Perotti, 1996). For example, higher taxes lower the rate of return on working (Mirrlees, 1971) and investment (Rebelo, 1991), which discourages individuals[9] from exerting effort or savings incentives, ultimately reduc-

[7]Government investment on infrastructure can also increase wealth inequality over time, regardless of its financing, see Chatterjee and Turnovsky (2012) for the details.

[8]Most literature in this stream assumes inelastic labour supply and risky entrepreneurship and thus households have to make an occupational choice to be a regular worker or an entrepreneur(Boadway et al., 1991; Banerjee and Newman, 1993), elastic labour supply is explored by García-Peñalosa and Turnovsky (2007) and García-Peñalosa and Wen (2008)

[9]Such effect is also applied to some large firms, see Cagetti and De Nardi (2009).

ing capital accumulation and economic growth. While such policies might promote greater equality, they come at an efficiency cost. Neoclassical theorists often argue that the efficiency losses from redistribution might surpass the benefits derived from heightened consumption among lower-income individuals and even increase wasteful competitive consumption which is used to differentiate oneself on their social status (Hopkins and Kornienko, 2006). From the perspective of human capital, income redistribution may also hinder economic growth if parental nurturing significantly influences human capital formation, as it reallocates resources to less educationally productive families, reducing the overall investment in human capital due to credit market imperfections and increasing returns in intergenerational human capital production (Bandyopadhyay and Tang, 2011). The redistributive effect of taxation on economic growth can be exacerbated by heterogeneous endowments in human and physical capital (Yang and Zhou, 2022). Redistribution might not be fully effective because, as noted by Okun (1975), *"some of the transfer will simply disappear in redistribution, so the poor will not receive all the money taken from the rich"*. In addition, the impact of redistribution could vary under different political and institutional conditions. For instance, high initial inequality reduces growth in democracies but has less clear effects in non-democracies (Deininger and Squire, 1998) and in more inclusive and democratic societies, redistribution can be growth-enhancing by preventing social unrest and creating a more cohesive society (Acemoglu and Robinson, 2002).

Regarding these two effects of redistribution, researchers are exploring optimal income taxation (Saez, 2001) and optimal capital taxation (Chamley, 1986; Judd, 1985; Aiyagari, 1995). Recent studies suggest that effective redistribution policies must navigate these trade-offs, with some research recommending optimal designs—such as productive government spending (Getachew and Turnovsky, 2015) and a combination of flat taxes and lump-sum transfers—that minimize efficiency losses while maximizing equity gains (Boar and Midrigan, 2022).

In addition to the fundamental mechanisms of human capital and redistribution discussed above, other often overlooked factors have been proposed to explain the relationship between inequality and growth. One deep-seated factor, ***cycles of technological progress***, may significantly influence the evolution of earnings inequality and intergenerational earnings mobility, as technological progress increases the relative return to ability, while greater accessibility to technologies decreases it. Earnings mobility may govern the pace of technological progress and output growth (Galor and Tsiddon, 1997). Another ***innovation-related*** factor is that innovative products exhibit both a positive price effect—where increased inequality allows innovators to charge higher prices—and a negative market-size effect—where higher inequality implies smaller markets for new goods and/or a slower transition of new goods into mass markets. It turns out that price effects dominate market-size effects (Foellmi and Zweimüller, 2006). Another often overlooked factor is the ***time dimension***: higher inequality may boost economic performance in the short term but reduce the growth rate of GDP per capita in the long run because the poor prefer direct transfers over public investment (Halter et al., 2014). Income inequality can also lead to disparities in health outcomes, the quality of social arrangements, stress levels, and mortality rates, thereby affecting labour productivity and economic performance (Deaton, 2003).

2.3 Conclusion remarks

The literature on the inequality-growth nexus reveals a complex and multifaceted relationship. Empirical studies offer mixed results, with some indicating a positive impact of inequality on growth, while others highlight the negative effects. Theoretical models provide various mechanisms through which inequality can influence growth and therefore significantly advance our understanding of how economic disparities impact overall economic performance. However, the conclusions drawn from these structural models can be sensitive to slight changes in these assumptions, raising questions about their robustness and applicability in diverse real-world scenarios. Additionally, to maintain analytical tractability, models often simplify complex real-world phenomena. For instance, factors such as political power dynamics, cultural influences, and historical contexts are often pared down or entirely omitted. This simplification can lead to models that fail to capture critical aspects that influence the relationship between inequality and growth. Many models do not adequately address the dynamic feedback mechanisms between inequality and growth. While some endogenous growth theories incorporate these dynamics to an extent, the intertemporal aspects of inequality—such as the cumulative effects of wealth accumulation and distribution—are often inadequately explored. Theoretical models often propose mechanisms that are difficult to validate empirically, data limitations often hinder the ability to test these models thoroughly. Theoretical models suggest various policy measures to mitigate inequality and promote growth. For example, policies that improve access to education and healthcare can enhance human capital, as highlighted by Galor and Zeira (1993). Additionally, reforms that reduce credit market imperfections, as discussed by Banerjee and Duflo (2003), can facilitate entrepreneurial activities and investment among less affluent individuals. However, implementing these policies effectively requires careful consideration of political feasibility and potential unintended consequences.

Despite extensive studies on the role of entrepreneurial incentives in driving economic growth which could be impacted by redistribution policies, there remains a significant gap in the empirical examination of the relationship between inequality and growth via direct entrepreneurship channel, this channel includes factors such as entrepreneurship incentives—financial and policy-driven motivations that encourage individuals to start and grow businesses—and regulatory barriers—legal and bureaucratic obstacles that hinder entrepreneurial activities. More specifically, while existing literature such as Parente and Prescott (1994, 2002) and Cagetti and De Nardi (2006, 2009) has extensively modelled the impact of technology adoption barriers and borrowing constraints, our model differs in that inequality in our model directly impacts the incentives to undertake innovative activities rather than focusing on the accumulation of skills. While the human capital literature emphasizes how disparities in income influence investment in education and skill development, our model highlights how wealth inequality affects the motivation and capacity for entrepreneurial and innovative endeavours by using the entrepreneurship penalty rate which further impacts productivity growth. The role of entrepreneurship in our model is an extension of Minford and Meenagh (2019, 2020), Yang et al. (2021).

Empirically, there is a paucity of empirical analyses, especially using a long-term dataset to test the role of entrepreneurial incentives on wealth distribution and economic growth within a specific developed country context. This thesis addresses this gap by investigating the long-term interplay between economic growth and inequality in the United Kingdom from 1870 to 2016. Using an endogenous growth model with heterogeneous agents, incorporating entrepreneurial incentives tied to individual wealth and credit conditions, provides a nuanced understanding of how wealth distribution influences growth. Our theory here is chosen for its potential relevance to a developed economy like the UK experiencing reforms designed specifically to benefit entrepreneurs; it focuses on an asymmetry between entrepreneurs according to wealth, an asymmetry that both causes inequality to boost growth and growth to boost inequality. However, as noted in the introduction, this asymmetry could also spring from some of the wealth-related advantages proposed in these earlier theories. This thesis instead considers the deterministic cost of entrepreneurship such as market regulatory barriers and government barriers like taxes, and the dynamic interactions of entrepreneurship, wealth, and growth.

In addition, this thesis offers critical insights into the welfare implications for policy formulation, emphasizing the trade-offs between growth and inequality. The welfare calculations further explain the tendency of governments to engage in redistribution. Due to political competition, governments are incentivized to please the average voter, who often benefits from redistribution policies. This political motivation helps to explain the historical trend in the UK towards increasing redistribution. As political parties vie for votes, they adopt policies that promise to reduce inequality and enhance social welfare, thereby appealing to the median voter who typically favours such measures. Our model implies this trade-off between growth and inequality due to redistribution. This will allow it to shed light on the political economy of the UK over history- which showed a long trend towards redistribution.

Chapter 3

Model

Diverse income brackets exhibit varied preferences in allocating their time between traditional employment and innovative endeavours, more precisely termed "entrepreneurship". The economic burden associated with entrepreneurship tends to be less for individuals with superior income or capital standings. Notably, the amplification of capital disparity often parallels economic growth in industrialized nations. The subsequent structural model draws upon these established observations to delineate the interrelationship between inequality and growth. This is achieved by postulating a mechanism: as capital inequality intensifies, with an increasingly narrow segment of the population amassing wealth, these individuals encounter diminished costs and augmented incentives to drive productivity growth. Subsequent sections will provide an in-depth elucidation of this mechanism.

3.1 Individual behavior

Let us consider an economy that consists of two distinct groups of individuals, each with fixed proportions represented by $\mu_i (i = 1, 2)$. It is assumed that both the capital market and the respective segmented labour markets operate under conditions of perfection. Each group, in this setting, selects their consumption $C_{i,t}$, labour input $N_{i,t}$, and time dedicated to innovation $Z_{i,t}$ in a manner that optimizes their utility over an infinite horizon. By assuming an infinite time horizon, the model focuses on the behaviour of individuals over their entire lives and even across generations. This allows for the examination of long-term decision-making, where individuals care not only about their current well-being but also about the future, giving a stationary decision-making process.

$$U = \max E_0 \left[\sum_{t=0}^{\infty} \beta^t U \left(C_{i,t}, N_{i,t}, Z_{i,t} \right) \right] \tag{3.1}$$

where

$$U \left(C_{i,t}, N_{i,t}, Z_{i,t} \right) = \Phi \frac{\left(v_{i,t} C_{i,t} \right)^{1-\Psi_1}}{1 - \Psi_1} + (1 - \Phi) \frac{\left(1 - u_{i,t} N_{i,t} - Z_{i,t} \right)^{1-\Psi_2}}{1 - \Psi_2} \tag{3.2}$$

where Ψ_1 and Ψ_2 are coefficients of relative risk aversion, Φ is the consumption preference; $v_{i,t}$ and $u_{i,t}$ are idiosyncratic shocks to consumption and labor respectively. The utility function takes the form of constant relative risk aversion (CRRA), which is a widely used representation of preferences in economics, particularly in models of intertemporal choice, consumption-savings decisions, and portfolio choice. It has been slightly changed to the additive form of consumption and leisure with the preference parameters Φ over these two choices. Relative risk aversion is a measure of how risk-averse a person is and how willing they are to accept a gamble that might increase or decrease their wealth. In the CRRA utility function, relative risk aversion is constant, meaning it does not depend on the level of wealth or consumption. The elasticity of intertemporal substitution (EIS), which measures how willing individuals are to substitute consumption and leisure across different periods, is given by $1/\Psi_1$ and $1/\Psi_2$ respectively. The CRRA utility function has the property of diminishing the marginal utility of consumption/leisure, meaning that additional consumption/leisure becomes less valuable as consumption increases. Also, the CRRA utility function remains unchanged under positive scaling of consumption and leisure, meaning that proportional changes in consumption/leisure levels have no effect on preferences.

The individual divides the time among three activities: leisure $L_{i,t}$, labour supply $N_{i,t}$ for real wage $w_{i,t}$, and innovation time $Z_{i,t}$ that is unpaid at period t but expected to have future returns. The time endowment of individuals is normalized as

$$L_{i,t} + N_{i,t} + Z_{i,t} = 1 \tag{3.3}$$

Normalizing the time endowment of individuals to 1 simplifies the calculations, offers flexibility in interpretations, and maintains generality by focusing on proportional allocations rather than absolute hours, allowing the model to be applied across various scenarios and time periods. The individual chooses $\{L_{i,t}, N_{i,t}, Z_{i,t}\}$ subject to the real terms budget constraint as

$$(1-\tau)Y_{i,t} + (1+r_{t-1})b_{i,t} - \pi_t Z_{i,t} + T_t = C_{i,t} + b_{i,t+1} + K_{i,t} - (1-\delta)K_{i,t-1} \tag{3.4}$$

The unit cost of entrepreneurship is $\pi_t, \pi_t Z_{i,t}$ is the individual's total cost of entrepreneurship, this cost measures the indirect effects of government taxes or regulations such as licensing and permits, access to financing, labour laws and trade policies on entrepreneurial activities.

Compared to the literature on human capital, productivity growth is driven by innovation in production processes or products in this model. This implies that the key determinant of wealth accumulation is the ability of agents to innovate, rather than their initial endowment of capital. The credit market in the model allows individuals to borrow for capital investments, physical or human. However, it does not extend to financing productivity growth, which in this context is driven by innovation time $(Z_{i,t})$ rather than changes in capital. The idiosyncratic risk associated with the payoffs from innovation is uninsurable; furthermore borrowing to offset it would require the posting of collateral in the absence of insurance, which is unavailable to poor entrepreneurs. Undertaking this risk personally is the only possibility for entrepreneurs; for poor entrepreneurs this is relatively unattractive in that the marginal utility of the expected returns may not exceed that of the costs involved, remembering that the poor have a higher marginal utility of current consumption from regular salary. This effect of wealth inequality lies at the heart of the relationship between growth and inequality in this model. Furthermore, the persistence of wealth disparities in the model arises endogenously due to the differential ability of agents to innovate and generate productivity growth. While there is an initial allocation of wealth, subsequent dynamics within the model, driven by innovation and productivity growth, lead to the emergence of persistent differences in wealth over time. Given that productivity growth is primarily driven by innovation rather than investments in human capital, it is plausible to assume that a segment of the population may lack the resources or capabilities to engage in productive innovation, leading to persistent poverty. The government levies a general income tax at a uniform rate, T. This could be the tax on wages, profits, interests, rents, etc. Instead of having different tax brackets or variable rates based on sources or amounts, the rate is a fixed percentage for all. Assume that these revenues are offset by lump sum transfer payments T_t to households. These are payments made by the government to households, which are not contingent on any action taken by the recipient. This means that it doesn't matter if the individual is working, unemployed, rich, or poor; everyone gets the same amount. The assumption here is that the total amount collected through income tax is equal to the total amount given back to households as lump sum transfers. This implies that the government isn't saving any money, investing it, or using it to finance its expenditure. The only fiscal activity they're engaging in is redistribution: collecting money through taxes and then returning it as transfers. Both the entrepreneurial cost and the lump sum transfers are indexed to general income, i.e. overall GDP. For households, the net effect of the tax and transfer could be neutral for the average person (they pay a tax but then get a transfer back). However, it would be redistributive: higher-income individuals or households would pay more in taxes than they receive in transfers, while the opposite would be true for lower-income individuals or households. This effect will be discussed in the empirical policy implications. Furthermore, the entrepreneurial cost indexed to GDP ensures that as the economy grows, the absolute cost or burden faced by entrepreneurs grows, but its relative cost (as a percentage of potential profits or revenue) remains the same. This fiscal policy acts as an automatic stabilizer. In times of economic downturns, when GDP shrinks, the government would automatically collect less in taxes and pays out less in transfers, providing some automatic fiscal stimulus. The individual has a Cobb-Douglas production function, it relates the output $Y_{i,t}$ to the inputs: physical capital $K_{i,t-1}$, labour $N_{i,t}$, and a productivity term $A_{i,t}$. The exponents α and $1-\alpha$ represent the output elasticity of capital and labour, respectively. The term $A_{i,t}$ is the "Total Factor Productivity" (TFP) and captures the efficiency with which inputs are transformed into outputs.

$$Y_{i,t} = A_{i,t} \left(K_{i,t-1}\right)^{\alpha} \left(N_{i,t}\right)^{1-\alpha} \tag{3.5}$$

where individual productivity growth is set to depend on the entrepreneurship time, which origins from Meenagh et al. (2007), $Z_{i,t}$,

$$\frac{A_{i,t+1}}{A_{i,t}} = \theta_1 + \theta_2 Z_{i,t} + v_{A,t} \tag{3.6}$$

This suggests that the more time an individual spends on entrepreneurial activities, the higher their productivity growth. An error term $v_{A,t}$ captures other unobserved factors affecting productivity growth at time t. See Appendix A.3 for more explanations on the characterization of endogenous balanced growth path. Then, to solve the problem, setting up the Lagrangian function,

$$L = E_0 \sum_{t=0}^{\infty} \beta^t \left\{ U\left(C_{i,t}, N_{i,t}, Z_{i,t}\right) - \lambda_t \left[C_{i,t} + \pi_t Z_{i,t} - (1-\tau) Y_{i,t} + b_{i,t+1} - \left(1 + r_{t-1}\right) b_{i,t} \right. \right.$$
$$\left. \left. + K_{i,t} \ - (1-\delta) K_{i,t-1} - T_t \right] \right\}$$

Take the first order conditions,

$$C_{i,t}: \quad \Phi \left(v_{i,t} C_{i,t}\right)^{-\Psi_1} = \lambda_t \tag{1a}$$

$$N_{i,t}: \quad (1-\Phi)\left(1 - u_{i,t} N_{i,t} - Z_{i,t}\right)^{-\Psi_2} = \lambda_t (1-\alpha)(1-\tau) A_{i,t} \left(K_{i,t-1}\right)^{\alpha} \left(N_{i,t}\right)^{-\alpha} \tag{2b}$$

$$Z_{i,t}: \quad (1-\Phi)\left(1 - u_{i,t} N_{i,t} - Z_{i,t}\right)^{-\Psi_2} = -\lambda_t \pi_t + (1-\tau) E_t \sum_{s=1}^{\infty} \beta^s \lambda_s \frac{\partial Y_{i,t+s}}{\partial Z_{i,t}} \tag{3c}$$

$$K_{i,t}: \quad \beta E_t \lambda_{t+1} \left[(1-\tau) \frac{\partial Y_{i,t+1}}{\partial K_{i,t}} + 1 - \delta \right] = \lambda_t \tag{4d}$$

$$b_{i,t+1}: \quad \beta \left(1 + r_t\right) E_t \lambda_{t+1} = \lambda_t \tag{5e}$$

Equation (1a) and (5e) gives,

$$C_{i,t}: \left(v_{i,t} C_{i,t}\right)^{-\Psi_1} = \left(1 + r_t\right) \beta E_t \left[\left(v_{i,t+1} C_{i,t+1}\right)^{-\Psi_1} \right] \tag{3.7}$$

Substituting out $(1 + r_t)$ gives the individual capital, see Appendix A.1 for the proof of the substitution.

$$K_{i,t} : (v_{i,t}C_{i,t})^{-\Psi_1} = \beta \left\{ E_t \left[(v_{i,t+1}C_{i,t+1})^{-\Psi_1} \right] \left[\alpha(1-\tau) \frac{E_t Y_{i,t+1}}{K_{i,t}} + 1 - \delta \right] \right\} \tag{3.8}$$

Equation (2b) and (1a) gives,

$$N_{i,t} : \quad (1-\Phi)\left(1 - u_{i,t}N_{i,t} - Z_{i,t}\right)^{-\Psi_2} = \Phi \left(v_{i,t}C_{i,t}\right)^{-\Psi_1} (1-\tau)(1-\alpha)\frac{Y_{i,t}}{N_{i,t}} \tag{3.9}$$

Equation (3c) and (1a) gives,

$$\frac{(1-\Phi)}{\left(1 - u_{i,t}N_{i,t} - Z_{i,t}\right)^{\Psi_2}} + \Phi \frac{\pi_t}{\left(v_{i,t}C_{i,t}\right)^{\Psi_1}} = (1-\tau)E_t \sum_{s=1}^{\infty} \beta^s \lambda_s \frac{\partial Y_{i,t+s}}{\partial Z_{i,t}} \tag{3.10}$$

Assuming $v_{i,t} = 1$ for simplicity, given individual production equation 3.5 and individual productivity equation 3.6, the differential of $Y_{i,t+s}$ in terms of $Z_{i,t}$ is

$$\frac{\partial Y_{i,t+s}}{\partial Z_{i,t}} = \frac{\partial Y_{i,t+s}}{\partial A_{i,t+s}} \frac{\partial A_{i,t+s}}{\partial A_{i,t+s-1}} \cdots \frac{\partial A_{i,t+1}}{\partial Z_{i,t}} = \frac{Y_{i,t+s}}{A_{i,t+s}} \frac{A_{i,t+s}}{A_{i,t+s-1}} \cdots \frac{A_{i,t+2}}{A_{i,t+1}} \frac{\partial A_{i,t+1}}{\partial Z_{i,t}}$$

where the partial derivative of $Y_{i,t+s}$ with respect to $A_{i,t+s}$:

$$\frac{\partial Y_{i,t+s}}{\partial A_{i,t+s}} = \left(K_{i,t+s-1}\right)^{\alpha} \left(N_{i,t+s}\right)^{1-\alpha} = \frac{Y_{i,t+s}}{A_{i,t+s}}$$

and the partial derivative of $A_{i,t+1}$ with respect to $Z_{i,t}$:

$$\frac{\partial A_{i,t+1}}{\partial Z_{i,t}} = \theta_2 A_{i,t}$$

Therefore,

$$\frac{\partial Y_{i,t+s}}{\partial Z_{i,t}} = Y_{i,t+s} \frac{A_{i,t}}{A_{i,t+1}} \theta_2 \tag{3.11}$$

This expression shows how $Z_{i,t}$ affects future output through the productivity channel. To find the first-order condition for $Z_{i,t}$, we relate the marginal utility from $Z_{i,t}$ to the expected future utility.

Now, we need to incorporate the dynamic response of $Y_{i,t+s}$ to $Z_{i,t}$. Substituting equation 3.11 into equation 3.10, Rewriting the first-order condition to include the marginal effect of $Z_{i,t}$:

$$\frac{(1-\Phi)}{(1-u_{i,t}N_{i,t}-Z_{i,t})^{\Psi_2}}+\Phi\frac{\pi_t}{(v_{i,t}C_{i,t})^{\Psi_1}}=(1-\tau)\frac{A_{i,t}}{A_{i,t+1}}\theta_2 E_t\sum_{s=1}^{\infty}\beta^s\lambda_s Y_{i,t+s}$$

Subtituting λ_s out by using equation (1a),

$$\frac{(1-\Phi)}{(1-u_{i,t}N_{i,t}-Z_{i,t})^{\Psi_2}}+\Phi\frac{\pi_t}{(v_{i,t}C_{i,t})^{\Psi_1}}=\Phi(1-\tau)\frac{A_{i,t}}{A_{i,t+1}}\theta_2 E_t\sum_{s=1}^{\infty}\beta^s\frac{Y_{i,t+s}}{(C_{i,t+s})^{\Psi_1}}$$

Substitute equation 3.9 multiply both sides with $(C_{i,t})^{\Psi_1}/\Phi$ yields

$$(1-\alpha)(1-\tau)\frac{Y_{i,t}}{N_{i,t}}+\pi_t=(1-\tau)\theta_2\frac{A_{i,t}}{A_{i,t+1}}(C_{i,t})^{\Psi_1}E_t\left[\sum_{s=1}^{\infty}\beta^s\frac{Y_{i,t+s}}{(C_{i,t+s})^{\Psi_1}}\right]$$

Set Ψ_1 to unity for simplification and $E_t\left[\frac{Y_{i,t+s}}{C_{i,t+s}}\right]\approx\frac{Y_{i,t}}{C_{i,t}}$, then Approximating $Y_{i,t}/C_{i,t}$ as a random walk around steady state. See Appendix A.2 for proof. Then the final first order condition of $Z_{i,t}$ can be rewritten as:

$$(1-\tau)(1-\alpha)\frac{Y_{i,t}}{N_{i,t}}+\pi_t=(1-\tau)\frac{A_{i,t}}{A_{i,t+1}}Y_{i,t}\frac{\beta\theta_2}{1-\beta} \tag{3.12}$$

Rearrange it as

$$\frac{A_{i,t+1}}{A_{i,t}}=\frac{(1-\tau)Y_{i,t}\frac{\beta\theta_2}{1-\beta}}{(1-\tau)(1-\alpha)\frac{Y_{i,t}}{N_{i,t}}+\pi_t} \tag{3.13}$$

Divided by $w_{i,t}=(1-\alpha)Y_{i,t}/N_{i,t}$ on the right-hand side,

$$\frac{A_{i,t+1}}{A_{i,t}}=\frac{\beta\theta_2(1-\tau)Y_{i,t}/w_{i,t}}{(1-\beta)(1-\tau+\pi'_{i,t})} \tag{3.14}$$

where $\pi'_{i,t}=\pi_t/w_{i,t}$ is denoted as the individual entrepreneurship penalty rate, whose determination will be discussed in the next section. See Appedix A.4 for the details of the log-linearization.

3.2 Entrepreneurship penalty rate

To link the barriers to entrepreneurship with the long-term growth rate, we adopted the mechanism of entrepreneurship penalty rate, which is explored by Minford and Meenagh (2019, 2020) and Yang et al. (2021). The individual entrepreneurship penalty rate, $\pi'_{i,t}$, and the wage-relative entrepreneurship penalty rate $\pi'_{i,t} = \frac{\pi_t}{w_{i,t}}$, measure the wage-relative costs of individuals on entrepreneurial activities. It falls with the rising of income or capital. The entrepreneur cost is deflated by $w_{i,t}$, which is used to introduce heterogeneity in entrepreneurship cost. While it is true that higher wages may lead individuals to seek more leisure, it doesn't necessarily imply causality between wage levels and the entrepreneurship penalty rate. The model focuses on the cost-benefit analysis of entrepreneurship, independent of leisure preferences induced by wage levels. The entrepreneurship penalty rate reflects the cost associated with entrepreneurship relative to potential earnings. Entrepreneurs evaluate this cost against the expected benefits of entrepreneurship.

We assume individuals can observe penalty rates but do not know exactly how it is set, thus, $\pi'_{i,t}$ is predetermined for optimizing individuals. Rather than coming from the first-order conditions directly, $\pi'_{i,t}$ is modelled as an exogenous process. Overall, the individual's entrepreneurship penalty rate, $\pi'_{i,t}$, evolves over time based on various factors including lagged capital ratio.[1], credit conditions, and its own lagged value.

$$\ln \pi'_{i,t} = \rho_0^{\pi} + \rho_1^{\pi} \ln \pi'_{i,t-1} - \rho_2^{\pi} \cdot Q\frac{K_{i,t-2}}{K_{t-2}} + \rho_3^{\pi}\,\mathrm{Cre}_t + \epsilon_t^{\pi} \tag{3.15}$$

The empirical findings of Banerjee and Duflo (2003) and Kolev and Niehues (2016) indicate that inequality has significant nonlinear effects on growth particularly when considering the squared term of inequality. Hence, we set $Q\frac{K_{i,t}}{K_t} \equiv \frac{\mu_i}{w_{Y,i}}\left(\frac{K_{i,t}}{K_t}\right)^2$. The coefficient $\frac{\mu_i}{w_{Y,i}}$ aims to avoid the penalty policy being too beneficial to the rich as the poor generally have a greater population weight relative to their average income share, for example, this relative ratio for poor is 0.9/0.7 while 0.1/0.3 for the poor.

[1]Note that capital level at the beginning of period t is denoted by $K_{i,t-1}$, so lag one-period capital inequality will take the form of $K_{i,t-2}/K_{t-2}$.

The credit conditions, Cre_t, show that easy access to credit can encourage individuals to start businesses. When individuals can borrow easily, they might undertake projects with potentially high returns, even if they have high upfront costs. These credit conditions could be indexed by the credit availability or the cost of credit.

3.3 Aggregate economy

To complete the model formulation, aggregate variables $\{Y_t, K_t, C_t\}$ are defined as the weighted sum of individual ones since the individual variable measures the representative value in each group, plus aggregate uncertainties $\epsilon_t^Y, \epsilon_t^K, \epsilon_t^C$,

$$Y_t = \mu_1 Y_{1,t} + \mu_2 Y_{2,t} + \epsilon_t^Y \tag{3.16}$$

$$K_t = \mu_1 K_{1,t} + \mu_2 K_{2,t} + \epsilon_t^K \tag{3.17}$$

$$C_t = \mu_1 C_{1,t} + \mu_2 C_{2,t} + \epsilon_t^C \tag{3.18}$$

where the weights of the two groups $\{\mu_1, \mu_2\}$ are set as $\{0.1, 0.9\}$. We didn't assume a time-varying μ_i, the aggregate uncertainties are therefore set to account for the potential changes in the weights.[2]

Market clearing in goods,

$$(1-\tau)Y_t = K_t + C_t - (1-\delta)K_{t-1} + \epsilon_t^M \tag{3.19}$$

where the left-hand side is the total production of goods after accounting for a tax (τ). It's the net output available for consumption and investment after the government takes its share, τY_t, with residual term ϵ_t^M representing the credit shocks. When this condition holds, the goods market is in equilibrium.

An aggregate production function here is a weighted average of individual output determined by the Cobb Douglas production function 3.5. This weighted average aggregate output was also used to back out aggregate productivity by $A_t = Y_t/(K_t N_t)$

[2]It would be possible to let the weights evolve, e.g. as the lagged weights, but it would make little difference to the model simulations. The shocks include any resulting measurement errors.

which follows an I(1) process. The model introduces 10 shocks in total, the productivity shocks (ϵ_t^A) are backout from the productivity growth $\Delta \ln A_t$ which is set to be an AR(1) process,

$$\Delta \ln A_t = A_0^A + \rho_1^A \Delta \ln A_{t-1}^A + \epsilon_t^A \tag{3.20}$$

The credit shocks (ϵ_t^M) are also backout from an AR(1) process of Cre_t:

$$\mathrm{Cre}_t = A_0^M + \rho_1^M \, \mathrm{Cre}_{t-1} + \epsilon_t^M \tag{3.21}$$

The aggregate penalty shock (ϵ_t^π) is also back-out from the AR(1) process.

$$\ln \pi_t' = A_0^\pi + \rho_1^\pi \ln \pi_{t-1}' + \epsilon_t^\pi \tag{3.22}$$

All the other 7 residuals in the model, $\{\epsilon_t^Y, \epsilon_t^K, \epsilon_t^C, \epsilon_t^{C1}, \epsilon_t^{C2}, \epsilon_t^{N1}, \epsilon_t^{N2}\}$, are assumed to follow independent AR(1) processes with a deterministic trend, which gives the other 7 innovations, $\boldsymbol{\eta_t}$.

$$\boldsymbol{\epsilon_t} = \boldsymbol{A_0} + \boldsymbol{\rho_2} t + \boldsymbol{\rho_1 \epsilon_{t-1}} + \boldsymbol{\eta_t} \tag{3.23}$$

See Appendix A.4 for the details of derivations of the full log linearized model.

Chapter 4

Data

4.1 Actual data

Most of the actual data comes from the A Millennium of Macroeconomic Data Version 3.1, blended by the team in Bank of England which has now been updated to 2016. The selected 147-year annual data starts from 1870 to 2016, totally covering 15 series. The inequality data mainly comes from the World Inequality Database maintained by the team led by Thomas Piketty. See Table 4.2 for more details on the raw data.

Figure 4.1 presents the real interest rate alongside its 10-year moving average, highlighting three notable periods of rising real interest rates in modern British history: 1918-1922 (post-WWI), 1945-1963 (post-WWII), and 1979-1984 (during Thatcher's first term). Each of these periods followed significant global events that influenced economic policies. The importance of the real interest rate in the growth-inequality nexus lies in its influence on investment and consumption decisions, which are critical drivers of economic growth. For example, higher real interest rates typically reduce borrowing and investment, slowing down economic growth. Conversely, lower real interest rates can stimulate borrowing, investment, and consumption, thereby fostering growth. This relationship can be easily derived from the Euler equation 3.7, which links consumption growth to the real interest rate. Additionally, the real interest rate clears the goods market by equating demand with production supply as equation 3.19. However, while it influences current investment and consumption levels, it does not drive steady-state growth, as it does not impact productivity growth directly. Consequently, the real interest rate has a temporary effect on capital, con-

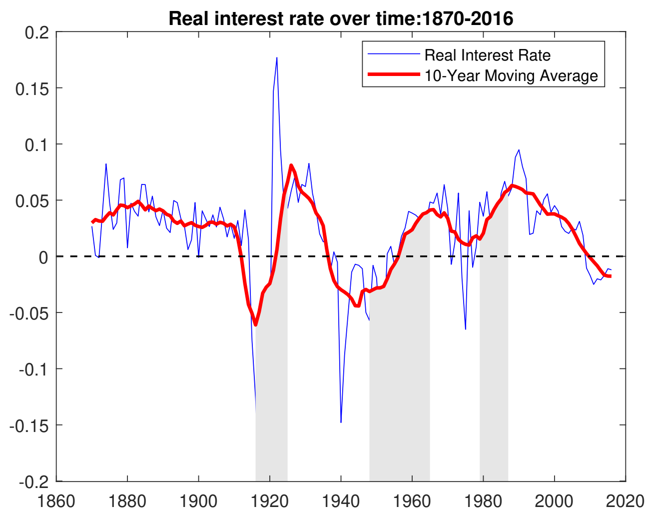


Figure 4.1. UK's real interest rate

sumption, and labour supply, leading to only a temporary impact on growth.

In the post-WWI period, inflation spikes necessitated higher interest rates to manage inflation and finance national rebuilding and debt. The UK's return to the pre-war gold parity in 1925 required deflationary policies, including elevated interest rates, to support the pound's value. Similarly, post-WWII, the British pound faced multiple crises, prompting interest rate hikes to stabilize the currency. During Margaret Thatcher's tenure, monetarist policies aimed at curbing inflation led to significant interest rate increases to control the money supply. Fiscal deficit reduction efforts and external factors like oil price shocks further influenced these rates. Each of these periods of increased real interest rates is associated with substantial economic shifts. As illustrated in Figures 4.2 and Figure 1.1, periods of post-war and post-crisis reconstruction coincide with rises in real GDP growth, capital stock, and economic inequality. This suggests that while high real interest rates might initially dampen growth, the subsequent economic adjustments and policies aimed at stabilization and rebuilding can lead to significant growth and changes in economic inequality. Understanding these dynamics helps to uncover the intricate relationship between real interest rates, growth, and inequality within the broader economic context.

By shading the periods with consecutive increases in three years, the UK economy experienced 10 episodes of expansion over the last two centuries and the capital stock

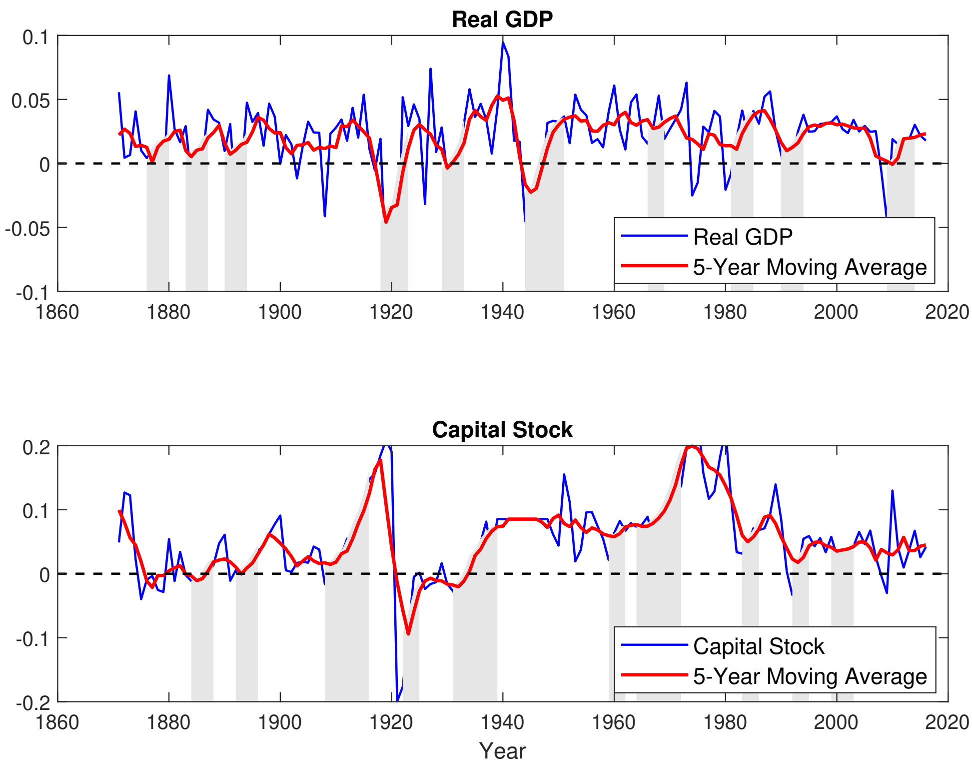


Figure 4.2. UK's real GDP and Capital Stock

almost follows the same trend. Both capital stock(approximated by non-dwellings capital of the whole economy) and real GDP are indicators of economic growth. As an economy grows and real GDP increases, businesses often invest more in capital goods to increase production, leading to an increase in capital stock, which is an essential part of the production function. However, the UK's capital stock obviously fluctuated with higher volatility than its real GDP growth, (S.D. 0.067 vs. 0.024). This can be impacted by the investment climate, where the real investment data has a maximum of 92% increase but 37% decrease over the two centuries from Table 1, technological progress, economic crises, etc. In a favourable investment environment, firms are more likely to invest in new capital goods, increasing capital stock for example the 1980s. Simultaneously, this investment contributes to economic growth and, thus, real GDP. During an economic downturn, businesses may reduce investment in capital goods, leading to a decline in capital stock, while real GDP may also fall but not necessarily at the same rate, e.g., during the world wars and financial

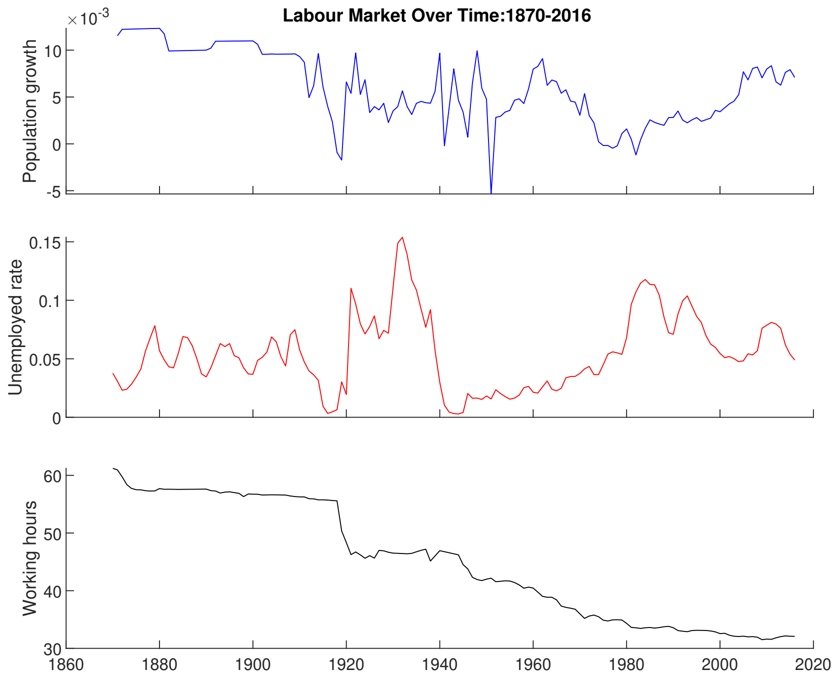


Figure 4.3. UK's labour statistics

crises. See Chadha et al. (2000) for more details of the UK business cycle properties.

Figure 4.3 provides a comprehensive view of the UK's labour market from 1870 to 2016, illustrating three key variables: population growth, unemployment rate, and working hours. The top panel shows population growth rates, which exhibit significant fluctuations over the period. Early years are marked by relatively high and volatile growth rates, with notable peaks and troughs. During the World Wars, there is a pronounced decline in population growth, reflecting the demographic impacts of the wars. Post-World War II, the growth rate stabilizes somewhat but continues to show variability, especially around the late 20th and early 21st centuries.

The middle panel displays the unemployment rate, revealing important historical shifts in the UK's economic landscape. Prior to the mid-1910s, unemployment rates predominantly hovered around the 5% mark. During the World Wars, the rates saw a significant rise, averaging around 10% and peaking at 15% at times. However, there was a swift decline post-war. After World War II, the unemployment rate experienced a gradual incline, moving from 2% to 5%. The onset of the Thatcher era marked a pronounced surge in unemployment, reaching highs of 13%. While there was a dip to 5% during Thatcher's second term, it rebounded to 9% in the early 1990s. Subsequent years witnessed a steady decline, only to be disrupted by the 2008 financial crisis.

The bottom panel illustrates the trend in working hours, showing a notable decline from the late 1800s to the mid-1900s, likely reflecting improvements in labour regulations and productivity. Several potential reasons contribute to this decline in working hours. Firstly, the introduction of labour laws such as the Factory Acts reduced maximum working hours and improved working conditions, promoting shorter workweeks. Secondly, increased productivity due to technological advancements meant that workers could produce more in less time, reducing the need for long working hours. Additionally, the rise of the service sector, which often requires fewer hours compared to manufacturing, and the growing emphasis on work-life balance, supported by both societal changes and government policies, contributed to the reduction in working hours. Following this decline, the mid-1900s to the early 2000s witnessed a period of relative stability in working hours, suggesting an equilibrium in the labour market or the absence of impactful societal shifts. However, the early 2000s onward marked a resurgence in working hours, with several years exhibiting increased working hours. These trends highlight the dynamic nature of the UK labour market over the past century and a half, showcasing the interplay between economic conditions, policy changes, and labour market dynamics.

To clarify, the labour supply functions in our model, as in equations A49 and A55, do not account for all potential drivers of labour supply. These omitted factors are captured in the error term. Our model focuses on elucidating the fundamental causes of growth and inequality and their effects on labour supply, rather than comprehensively addressing all determinants of labour supply trends. In the model, when entrepreneurs allocate more time to innovation, labour supply decreases. However, this mechanism alone does not explain the broader, ongoing decline in labour hours.

Figure 4.4 illustrates the tax revenue as a percentage of GDP and the trade union membership ratio from 1870 to 2016. The tax revenue as a percentage of GDP shows a significant increase starting around the early 20th century, peaking during the World

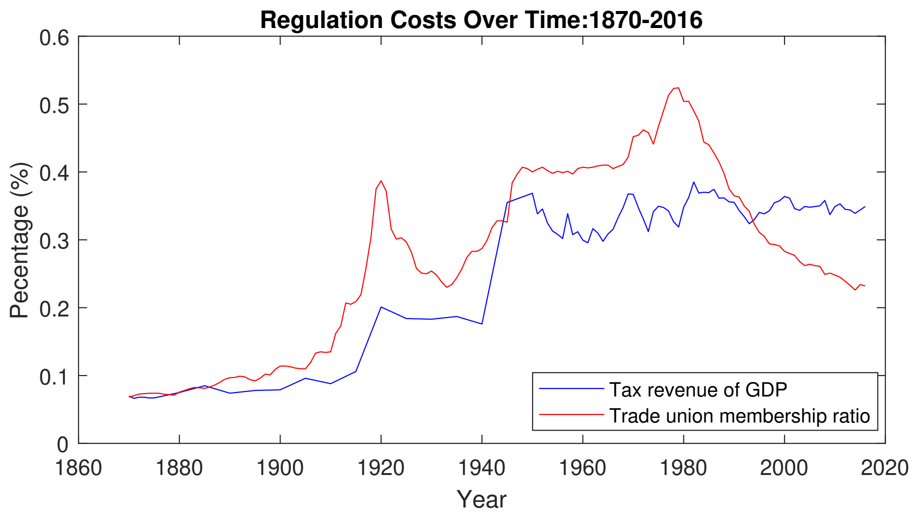


Figure 4.4. UK's regulation costs

Wars, and then stabilizing at higher levels post-World War II. Trade union membership saw a sharp rise in the early 20th century, peaking around the mid-20th century, followed by a steady decline from the late 20th century onwards.

Collectively Figure 1.1, provides insights into the economic inequality and regulatory landscape of the UK over a century and a half. Prior to the 1970s, both the income and wealth shares of the top 10% were on a declining trend, which correlates with the increasing tax revenue as a percentage of GDP shown in Figure 4.4. This period was characterized by progressive taxation and a stronger role of government in wealth redistribution, as evidenced by rising tax revenues. Post-1970s, there is a notable divergence where the income share of the top 10% starts to increase while the wealth share stabilizes. This period coincides with a significant reduction in trade union membership, indicating weaker collective bargaining power and potentially less equitable income distribution. See similar results and more details in Allen (2009) and (Lindert, 2000).

The impact of wars and economic policies is evident in both figures. During the World Wars, significant shifts are observed. Tax revenue peaks during these periods due to increased government spending and economic mobilization, while the income and wealth shares of the top 10% continue to decline, reflecting the broader economic impacts of the wars and post-war policies aimed at rebuilding and redistribution.

The stabilization of tax revenues at higher levels post-World War II reflects the establishment of the welfare state and ongoing redistributive policies. Concurrently, the continued decline in the wealth share of the top 10% suggests the effectiveness of these policies in reducing wealth concentration.

From the 1980s onward, the Thatcher era marked significant policy shifts, including deregulation and tax reforms aimed at stimulating economic growth. Figure 1.1 shows a corresponding increase in the income share of the top 10%, while Figure 4.4 highlights the decline in trade union membership, suggesting a weakening of labour's bargaining power. The simultaneous trends of increasing income inequality and decreasing regulatory costs indicate a shift towards more market-oriented policies.

These figures illustrate the intricate relationship between economic inequality and regulatory policies in the UK over time. The data suggest that periods of higher taxation and stronger labour unions correlate with reduced income and wealth concentration, while deregulation and reduced union influence are associated with increased income inequality. Understanding these historical trends is crucial for informing current and future economic policies aimed at addressing inequality and promoting sustainable growth.

The descriptive statistics of the raw data series are provided in Table 4.1.

Table 4.1. Descriptive statistics of the main raw data series

Series	Mean	Median	S.D.	Min.	Max.
Real interest rate	0.02021	0.02912	0.04958	-0.20048	0.17693
Real GDP	0.02000	0.02476	0.02900	-0.10215	0.09461
Capital stock	0.05160	0.04594	0.06710	-0.19926	0.25821
Consumption	0.04979	0.04724	0.05458	-0.17710	0.23089
Investment	0.02876	0.03341	0.11831	-0.37469	0.92676
Population	0.00603	0.00564	0.00379	-0.00535	0.01233
Inc10	0.39785	0.39505	0.07784	0.27780	0.54990
Wea10	0.75659	0.82978	0.17384	0.45589	0.94580
Tau	0.23097	0.28340	0.11820	0.06640	0.38514
TUMR	0.27168	0.27700	0.13430	0.06832	0.52400
Unemployment	0.05460	0.05188	0.03139	0.00283	0.15387
Working hours	44.829	45.624	9.9676	31.507	61.22

4.2 Calculated data

Given the unobservability of individual series such as consumption and labour supply, we rely on income/wealth distribution data from the World Inequality Database and aggregate data from the Millennium of Macroeconomic Data. Specifically, we utilize the top 10% income group (Group 1) and the bottom 90% income group (Group 2) to demarcate heterogeneous-agent groups. While some may advocate for treating the top 1% income group as representative of the wealthy, several considerations guided our choice. Firstly, the top 10% income group offers a more representative sample of the entrepreneurial class, encompassing a broader range of professions and sectors than the top 1%. Secondly, the availability and reliability of data play a pivotal role; the World Inequality Database, a commonly used dataset, provides income shares for both top 1% and top 10%. As the top 10% data is derived from fiscal data, it is often considered more dependable than the top 1%. Additionally, the accessibility and comprehensiveness of the top 10% data facilitate robust analyses and informed conclusions. Lastly, from a policy standpoint, focusing on the top 10% allows for the exploration of interventions that can address inequality and promote entrepreneurship with broader societal implications. By prioritizing the top 10%, we can devise policies that are not only more feasible but also resonate with a wider segment of the population. History provides numerous examples of successful entrepreneurs who were born into wealthy or middle-class families, such as John Pierpont Morgan, Rupert Murdoch, Warren Edward Buffett, William Henry Gates III, and Steven Paul Jobs. Supporting this notion, Levine and Rubinstein (2017) presents evidence from NLSY79 data in the US, showing that a higher proportion of entrepreneurs come from well-educated and high-income families.

The subsequent section delineates the methodological intricacies involved in computing the derived data. Initially, individual income and capital are estimated through the utilization of corresponding income and capital shares derived from the aggregate. Subsequently, individual consumption is derived via the application of the budget constraint. The equation governing individual consumption is expressed as:

$$c_{i,t} = C_t \times ICR_{i,t} \tag{4.1}$$

where $ICR_{i,t}$ denotes the individual consumption ratio, defined as the ratio between individual consumption from the budget $c^b_{i,t}$ and total individual consumption $c^A_{i,t}$.

$$ICR_{i,t} = c^b_{i,t}/c^A_{i,t} \tag{4.2}$$

Individual consumption from the budget ($c^b_{i,t}$) is further back-out from the individual budget constraint:

$$c^b_{i,t} = y_{i,t} - k_{i,t+1} + (1-\delta)k_{i,t} \tag{4.3}$$

The total individual consumption ($c^A_{i,t}$) is calculated as the summation across both income groups:

$$c^A_{i,t} = \sum_{i=1}^{2}(y_{i,t} - k_{i,t+1} + (1-\delta_t)k_{i,t}) \tag{4.4}$$

The depreciation rate (δ_t) is determined by the ratio of the difference between future capital K_{t+1} and investment I_t to current capital K_t.

$$\delta_t = 1 - (K_{t+1} - I_t)/K_t \tag{4.5}$$

Individual labour supply is estimated by using composite weekly working hours (H), unemployed workforce(U), and total workforce (W). Composite series of Average Weekly Hours worked adjusted for part-time work, sickness, holidays, and stoppages and assume unemployment only occurs in Group 2,

$$N_{1,t} = H/(24 \times 5) \tag{4.6}$$

$$N_{2,t} = \frac{0.9W - U}{0.9W}\frac{H}{24 \times 5} \tag{4.7}$$

where (24×5) gives the total theoretical working hours in a week except for the weekend, $\frac{0.9W-U}{0.9W}$ adjusted the labor supply by considering the unemployment in Group 2.

The aggregate penalty rate, π, is proxied by the trade union membership ratio, a higher ratio means the employee has more power to bargain with the entrepreneur, which induces a higher penalty rate. Trade union membership ratio also provides insights into the broader institutional and regulatory environment, which could be relevant to the barriers to entrepreneurship in this model. Especially in the UK, trade unions have played a particularly significant role in influencing economic policies, labour laws, and regulations that could affect entrepreneurial activities in history. The trade union membership data were originally archived by the Department for Business, Innovation and Skills of the UK, which dates from 1892 to 2016. To complete the data from 1870-1891, the exponential approximation was used based on the number of employees. Individual penalty rates are also approximated by the aggregate penalty rate based on the share of income.

Some may argue that taxes, as a typical regulatory cost, should be incorporated into the penalty rate. However, several considerations address this perspective. Firstly, the tax data utilized in this study represents tax revenue as a percentage of GDP, as illustrated in Figure 4.4. This data closely mirrors the trend of the Total Unobserved Mark-up (TUM) until the 1990s, which encompasses the majority of our sample period. Secondly, tax data has been relatively stable compared to TUM data, especially after 1960, making it an unsuitable approximation for real regulation costs during that period. Furthermore, it is important to clarify that tax data has not been integrated into the penalty rate computation in our analysis. Consequently, tax rates do not directly influence productivity growth within our model. Instead, the impact of taxes on growth is considered solely through their redistribution effects. This distinction ensures that our analysis accurately reflects the role of regulatory costs and their influence on economic dynamics without conflating them with the effects of taxation.

The credit ratio, an essential metric for gauging financial conditions, is calculated based on a common approximation:

$$Cre_t = \ln \frac{M4\,\text{stock}}{\text{Nominal GDP}} \tag{4.8}$$

The original data for the M4 stock to Nominal GDP ratio, sourced from the Bank of England (BoE), spans the period from 1880 to 2016. To extend this series back to 1870, a linear approximation method was employed for the years 1870 to 1879, ensuring a continuous and comprehensive dataset. This extended dataset provides a robust foundation for analyzing historical trends in credit conditions.

The forecast of expected individual consumption was derived using a Vector Autoregression (VAR) model of order 1, estimated from the individual consumption data discussed earlier. This methodological approach allows for capturing the dynamic relationships between consumption and other economic variables, providing more accurate predictions. All nominal series were deflated using the 2013 market price index to maintain consistency and comparability across time periods.

Aggregate productivity (A) was approximated by the Solow residuals, which reflect the portion of output growth not explained by the accumulation of capital and labour, thus serving as a measure of technological progress. Individual productivity was calculated using the same methodology to ensure consistency in our analysis.

The final dataset comprises data for 16 endogenous variables within the model. All endogenous variables, except for the real interest rate, were found to be integrated of order one, indicating that they contain a unit root and are non-stationary in levels but stationary in first differences. This stationarity analysis is detailed in Table 4.3. We use the unfiltered data in the following evaluation and estimation as we want to utilize the unobservable information that is contained in the long-run dataset. See Appendix C.3 for a detailed justification of the use of unfiltered and non-stationary data for empirical analysis of the structural model..

Table 4.2. Data Description

Series	Description	Definition / Calculation	Source
R	Real interest rate	Annual average of 3-month rates minus 1-year head expectation at the start of the year	BOE
Y	GDP	Real GDP per capita deflator at market price 2013, log-demean	BOE
K	Capital Stock	Approximated by Non-dwellings whole economy, at market price 2013, log-demean of output	BOE
C	Consumption	Household consumption, at market price 2013, log-demean of output	BOE
A	Productivity	Approximated by Solow Residuals ($= Y/(K^{\alpha} * N^{1-\alpha})$)	Calculated
P’	Aggregate Penalty	Approximated by Trade union members ratio ($= Tradeunionmembers/AllEmployers$)	BOE
Tau	Tax rate	Tax Revenue of GDP	WID
CRE	Credit condition	Ratio of broad money over GDP ($= \ln(M4$ stock/GDP$)$)	BOE
Y1, Y2	Individual Income	Estimated by WID income distribution with interpolation ($= GDP \times$ Income share)	WID
K1, K2	Individual Capital	Estimated by WID wealth distribution with interpolation ($= Capitalstock \times$ Wealth share)	WID
C1, C2	Individual Consumption	Backed out by Individual budget constraint	Calculated
N1, N2	Individual Labour	Estimated by using Composite weekly working-hours, Total workforce/unemployed	BOE
P1’, P2’	Individual Penalty	Calculated by $= P' \times (Y/Y_i)$	Calculated
A1’, A2’	Individual Productivity	Approximated by Solow Residuals ($= Y_i/(K_i^{\alpha} * N_i^{1-\alpha})$)	Calculated
EC1, EC2	Expected individual consumption	Calculated from a VAR model	Calculated
TUM	Trade union members	Department for Employment (1892-1973); and the Certification Office (1974-2015)	BOE
EE	Total Employment in Heads	Survey data, e.g., Labour Force Survey	BOE
TUMR	TUM Ratio	$TUMR = TUM/EE$	Calculated
P	Market prices	GDP deflator at market prices, 2013=100	BOE
Pop	Population	Census data	BOE
Inc10	Top 10 income share	Post-tax income of top 10% population aged over 20	WID
Wea10	Top 10 wealth share	Net personal wealth of top 10% population aged over 20	WID
I	Real investments	Chained Volume measure, 2013 prices	BOE
W	Total workforce	Total number of people who are physically able to do a job and are available for work	BOE
U	Total unemployed	Total number of people of working age who are without work	BOE
H	Composite weekly working hours	Average Weekly Hours worked adjusted for part-time work, sickness, holidays, and stoppages	BOE

Table 4.3. Stationarity the endogenous variables

		Augmented Dickey-Fuller t-stats				Phillips-Perron t-stats				KPSS LM-stats			
		Level(c)	Level(c,t)	Diff.(c)	Concl.	Level(c)	Level(c,t)	Diff.(c)	Concl.	Level(c)	Level(c,t)	Diff.(c)	Concl.
R	Real interest rate	-4.89***	-4.88***	-12.91***	S(1%)	-4.98***	-4.98***	-18.17***	S(1%)	0.0976	0.0976	0.0946	S(1%)
Y	Real GDP	1.013	-2.129	-8.33***	I(1)(1%)	1.177	-1.737	-8.00***	I(1)(1%)	1.377+++	0.343+++	0.396	I(1)(5%)
K	Capital Stock	0.369	-1.565	-7.90***	I(1)(1%)	0.647	-1.264	-7.74***	I(1)(1%)	1.340+++	0.276+++	0.285	I(1)(1%)
C	Consumption	1.751	-1.075	-8.94***	I(1)(1%)	1.851	1.034	-9.19***	I(1)(1%)	1.324+++	0.349+++	0.576++	I(1)(10%)
A	Productivity	1.239	-2.382	-9.25***	I(1)(1%)	1.307	-0.487	-9.22***	I(1)(1%)	1.324+++	0.339+++	0.679++	I(1)(10%)
P′	Aggregate Penalty	-1.951	-0.552	-4.18***	I(1)(1%)	-2.038	-0.140	-8.68***	I(1)(1%)	1.021+++	0.318+++	0.587++	I(1)(10%)
Tau	Tax rate	-1.120	-2.486	-8.39***	I(1)(1%)	-1.065	-1.903	-9.77***	I(1)(1%)	1.235+++	0.329+++	0.092	TS(10%)
Cre	Credit condition	-0.484	-1.474	-6.77***	I(1)(1%)	-0.056	-1.084	-6.49***	I(1)(1%)	-0.753+++	0.303+++	0.258	I(1)(1%)
Y1	G1′s Income	1.629	-1.204	-8.39***	I(1)(1%)	1.761	-0.486	-9.01***	I(1)(1%)	1.167+++	0.346+++	0.744+++	I(1)(1%)
Y2	G2′s Income	0.164	-2.745	-8.39***	I(1)(1%)	0.206	-2.153	-7.87***	I(1)(1%)	1.420+++	0.280+++	0.122	I(1)(1%)
K1	G1′s Capital	0.682	-1.112	-9.97***	I(1)(1%)	0.682	-1.265	-9.96***	I(1)(1%)	1.331+++	0.243+++	0.198	I(1)(1%)
K2	G2′s Capital	-0.291	-1.117	-10.34***	I(1)(1%)	-0.314	-1.273	-10.02***	I(1)(1%)	1.497+++	0.382+++	0.192	I(1)(1%)
C1	G1′s consumption	0.839	-0.700	-11.03***	I(1)(1%)	-1.337	-2.167	-25.87***	I(1)(1%)	0.936+++	0.322+++	0.435+	I(1)(5%)
C2	G2′s consumption	-0.225	-2.871	-18.62***	I(1)(1%)	-0.337	-3.768	-23.49***	I(1)(1%)	1.397+++	0.412+++	0.231	I(1)(1%)
N1	G1′s labour	-0.652	-2.138	-9.24***	I(1)(1%)	-0.634	-2.035	-9.22***	I(1)(1%)	-0.753+++	0.303+++	0.258	I(1)(1%)
N2	G2′s labour	-1.033	-3.328*	-9.18***	I(1)(1%)	-1.067	-2.877	-9.45***	I(1)(1%)	1.314+++	0.243+++	0.198	I(1)(1%)
P1′	G1′s Penalty	-1.909	-0.543	-6.13***	I(1)(1%)	-2.027	-0.401	-5.94***	I(1)(1%)	1.007+++	0.318+++	0.603++	I(1)(10%)
P2′	G2′s Penalty	-1.851	-0.173	-8.08***	I(1)(1%)	-1.915	-0.316	-8.18***	I(1)(1%)	1.915+++	0.316+++	0.588++	I(1)(10%)
A1′	G1′s productivity	1.485	-1.586	-9.18***	I(1)(1%)	1.616	-1.356	-9.34***	I(1)(1%)	1.223+++	0.306+++	0.523++	I(1)(10%)
A2′	G2′s productivity	0.166	-2.799	-12.05***	I(1)(1%)	0.835	-2.560	-13.46***	I(1)(1%)	1.428+++	0.310+++	0.202	I(1)(1%)
	Critical value (1%)	-3.476	-4.022	-3.476		-3.476	-4.022	-3.476		0.739	0.216	0.739	
	Critical value (5%)	-2.881	-3.441	-2.881		-2.881	-3.441	-2.881		0.463	0.146	0.463	
	Critical value (10%)	-2.577	-3.145	-2.577		-2.577	-3.145	-2.577		0.347	0.119	0.347	

Note: Tests in levels are conducted with model "level (c)" with a constant, and "level (c,t)" with both constant and linear trend. Tests in the first difference are conducted using model "difference (c)" with a constant only. For ADF and PP tests, asterisks denote rejection of the unit root null at 10% (*), 5% (**) and 1% (***) significance levels. The KPSS test evaluates the null of stationarity. Plus signs indicate rejection of the stationarity null for the KPSS test at 10% (+), 5% (++) and 1% (+++) significance levels. In the conclusion column, 'S' stands for stationarity, 'TS' stands for trend stationarity, 'I(1)' stands for integrated of first order with constant, 'I(1) TS' stands for integrated of first order with trend stationarity.

4.3 Calibrated parameters

The parameters employed in our model, namely the capital share in production (α), utility discount factor (β), and capital depreciation rate (δ), are set to typical annual values found in the economic literature: $\alpha = 0.3$, $\beta = 0.97$, and $\delta = 0.045$. These values align with those commonly used in macroeconomic models. For instance, the capital share aligns closely with the empirical estimates presented in seminal works such as Solow (1956) and more recent analyses by Mankiw et al. (1992) which validate its appropriateness across different economic contexts. The utility discount factor is referred from The Green Book (2022) issued by HM Treasury. Also, the capital depreciation rate is referred from the guidelines of the Office of Tax Simplification. For Group 1's population share, we select the top 10%, see justification at section .

Individual income, capital, and consumption shares—$\omega_{y1}, \omega_{y2}, \omega_{k1}, \omega_{k2}, \omega_{c1}, \omega_{c2}$—are derived from the sample averages of the individual data, taking approximate values of $0.40, 0.60, 0.75, 0.25, 0.35, 0.65$. These values represent a distributional parameterization that captures the typical asset and income distributions within the dataset. This range is similar to the one found in data used by Piketty and Saez (2006) in their analysis of income inequality, providing a credible basis for comparison. The consistency of these values with established literature highlights their reliability. For example, Atkinson et al. (2011) also report similar distributions in their comprehensive studies on wealth and income distribution, further validating our approach. Additionally, the use of these shares allows for a more nuanced understanding of economic disparities, as highlighted by Saez and Zucman (2016), who emphasize the importance of detailed distributional analysis in capturing the full extent of economic inequality. By ensuring our parameterization aligns with such reputable sources, we enhance the robustness and empirical grounding of our model.

The steady-state values of the output-consumption ratio (1.4760) and capital-consumption ratio (2.6361) were calculated using the sample means of the relevant variables, providing a realistic snapshot of the economic environment based on observed empirical trends. These calculations follow methodologies akin to those discussed in Smets and Wouters (2007). Additionally, the steady-state variables, including labour inputs ($N_1 = 0.3879$, $N_2 = 0.4415$), penalty rates ($\pi_1 = 0.1747$, $\pi_2 = 0.2011$), real interest rate ($EQR = 0.0202$), and income tax rate ($EQT = 0.2000$), were derived as sample averages from the UK Millennium Dataset. By leveraging sample means from the UK Millennium Dataset, we ensure our model's parameters are not only theoretically sound but also empirically validated. This approach aligns with best practices in macroeconomic modelling, where the integration of real-world data enhances the reliability and applicability of the findings. The use of these steady-state values allows for a nuanced understanding of the economic dynamics at play, providing a solid foundation for further analysis and policy recommendations. For instance, the labour inputs and penalty rates are crucial for understanding the labour market dynamics and the cost of capital, respectively. The real interest rate and income tax rate are fundamental in assessing the broader fiscal and monetary environment, directly influencing investment decisions and consumption patterns.

A summary of these calibrated parameters is provided in Table 4.4, which not only details the values but also references their sources and justifications, ensuring transparency and reliability.

Table 4.4. Calibrated parameters

Series	Description	Source	Value
α	Capital share	Solow (1956)	0.3000
β	Utility Discount Factor	The Green Book (2022)	0.9700
δ	Capital Depreciation Rate	Office of Tax Simplification, UK	0.0453
μ_1	Group1's Population Share	Model calibration	0.1000
ω_{y1}	Top 10% Income Share	WID	0.3978
ω_{y2}	Bottom 90% Income Share	WID	0.6022
ω_{k1}	Top 10% Capital Share	WID	0.7566
ω_{k2}	Bottom 90% Capital Share	WID	0.2434
ω_{c1}	Top 10% Consumption Share	WID	0.3530
ω_{c2}	Bottom 90% Consumption Share	WID	0.6470
Y/C	Steady State Y/C Ratio	Smets and Wouters (2007)	1.4760
K/C	Steady State K/C Ratio	Smets and Wouters (2007)	2.6361
$N1$	Steady State Labour of Group1	UK Millennium Dataset	0.3879
$N2$	Steady State Labour of Group2	UK Millennium Dataset	0.4415
π_1	Steady State Penalty Cost of Group1	UK Millennium Dataset	0.1747
π_2	Steady State Penalty Cost of Group2	UK Millennium Dataset	0.2011
EQR	Steady State Real Interest Rate	UK Millennium Dataset	0.0202
EQT	Steady State Marginal Income Tax Rate	UK Millennium Dataset	0.2000

Chapter 5

Methodology

This chapter explains the methodologies for solving the DSGE model with rational expectations and in estimating the model. Firstly, it introduces the "Type II fix" and the terminal conditions in solving the non-linear model. Secondly, this chapter gives a brief introduction to Indirect Inference in testing and estimating the model.

5.1 Solving the DSGE model with Type II Fix

Suppose the wide class of DSGE models is represented by:

$$E_t \mathbf{y}_{t+1} = g(\mathbf{y}_t, \mathbf{x}_t; \boldsymbol{\theta}) \tag{5.1}$$

where $\mathbf{y}_t$ are endogenous variables with dimension $n_y \times 1$ in period t. and $\mathbf{x}_t$ are exogenous variables with dimension $n_x \times 1$ which we assume are driven by

$$\mathbf{x}_t = h(\mathbf{x}_{t-1}; \gamma) + \boldsymbol{\eta}_t \tag{5.2}$$

The exogenous variables may contain both observable and unobservable variables such as a technology shock. Both $\mathbf{y}_t$ and $\mathbf{x}_t$ are non-stationary. The function $g(\cdot)$ is either linear or non-linear with the structural parameters $\boldsymbol{\theta} \in \Theta$ determined by its equilibrium conditions. Usually, we assume the $h(\cdot)$ as an AR(1) process with AR coefficient γ and structural innovations $\boldsymbol{\eta}_t \sim$ i.i.d. $\mathcal{N}(0, \mathbf{R}_{\eta,t})$.

The perturbation method and the projection method are the most common methods used to solve the DSGE models. The perturbation method approximates solutions around a steady state using a Taylor series, making it efficient for models with small deviations from equilibrium but less accurate for large shocks or nonlinearities; while the projection method uses basis functions to approximate policy or value functions, offering high accuracy and flexibility for nonlinear models but at a higher computational cost (Fernández-Villaverde et al., 2016). The following section explains one of the projection methods, Type II fix, which refers to the definition of "Type II iteration" from Fair and Taylor (1983), see Appendix C.1 for more details. The numerical algorithms originate from the one to solve the Liverpool model, introduced by Minford et al. (1984).

The numerical method described here entails a series of iterations that converge from an arbitrary initial path of values for the expectations to a path of rational expectations, consistent with the forecasts of the model itself. Here is a brief explanation of the steps of the algorithm.

Step 1: solve the model using actual data and structural errors. Starting with initial values of rational expectations (RE) and endogenous variables from period 1 to F, which is defined as the forecasting period, in practice, the initial values were set as the first couple of periods(e.g.$F = 50$ for quarterly data, and $F = 3$ for annual data) of the actual data. Then, input the above initial values into the model and calculate the values of variables in the next period as predictions. This gives F periods of predictions, which are denoted as $y_2^f, \ldots, y_{F+1}^f$,

Step 2: calculate the difference between predictions and the given RE values. All the differences must be less than a tolerance level. If the gap is greater than the tolerance level, then use a certain algorithm (e.g. Gauss-Seidel, Powell-Hybrid, see Appendix C.2 for more details) to renew the given RE from initial values to predictions until predictions are close enough to the newest RE values. Thus, the model is solved in the first period. All the differences between the values of variables when the model is solved and the actual data are saved in all the F periods.

Step 3: iterate over the following periods. The algorithm uses RE values one period ahead to repeat the above steps to obtain predictions $y_3, \cdots, y_{F+2}$, and the model is solved in the second period. The algorithm needs to solve the model in all N periods which requires the data in totally $N + F$ periods. In practice, the dataset is extended at the beginning, the stationary series in the model is kept as the values of the last period's actual data with white noise through the following F period, and the non-stationary series is extended by time trend regression.

Step 4: Set up the terminal conditions. The numerical method could have multiple solutions. To find a unique one, terminal conditions are needed, representing that the model will be in a steady state after a terminal period T. Although T should be infinity in principle, it has to be a finite number in numerical solutions. Both Matthews and Marwaha (1979) and Minford et al. (1979) find that solutions are not sensitive to the choice of T. T is therefore set as the number of the last forecasting period $(N + F)$. If the RE variables depend on the last period of non-stationary variables like productivities, the terminal condition can vary, given different structural parameter values. Generally, the terminal condition is an important component of the mathematical problem that defines the DSGE model and its solution. It ensures that the solution is economically meaningful and consistent with the behaviour of the agents in the model. Over a finite period at the terminal date, we assume that

$$\mathbf{y}_{T-2} = \mathbf{y}_{T-1} = E_{T-1}\mathbf{y}_T = \bar{\mathbf{y}} \tag{5.3}$$

$\bar{\mathbf{y}}$ is a vector of terminal values of all endogenous variables of the model. It verifies at the terminal date, all endogenous variable values can be expressed as a function of values of error processes. In the model, productivity is considered as the unit-root process that drives the non-stationarity property, while other error processes would be zero. The system can be expressed as:

$$\bar{\mathbf{y}} = \bar{g}(\bar{\mathbf{y}}, \mathbf{x}_t; \bar{\boldsymbol{\theta}}) \tag{5.4}$$

Rewrite this as

$$\bar{\mathbf{y}} = f(\mathbf{x}_t; \bar{\boldsymbol{\theta}}) \tag{5.5}$$

This shows at terminal date T, all endogenous variable values can be expressed as a function of values of exogenous processes. In our model, productivity is considered as the unit-root processes drive the non-stationarity property, while other error processes would be zero at T.

5.2 Indirect inference

This section introduces the methodology used in the empirical analysis: the Indirect inference. Indirect Inference provides a classical statistical inferential framework for estimating and testing a model. It is most useful in estimating models for which the likelihood function (or any other criterion function) is analytically intractable or too difficult to evaluate, such as nonlinear dynamic models, models with latent (or unobserved) variables, and models with missing or incomplete data. The following sections present the main framework of indirect inference in the empirical analysis of a DSGE model.

5.2.1 Select the auxiliary model

Like other simulation-based methods, indirect inference requires only that it is possible to simulate data from the economic model for different values of its parameters. Unlike other simulation-based methods, indirect inference uses an approximate, or auxiliary, model to form a criterion function. The auxiliary model serves as a window through which to view both the actual data and the simulated data generated by the economic model. It compares the performance of the auxiliary model estimated on the simulated data derived from the economic model, with the performance of the auxiliary model estimated from the actual data. Indirect inference chooses the parameters of the economic model so that these two estimates of the parameters of the auxiliary model are as close as possible (or so that the actual data and the simulated data look the same from the vantage point of the chosen window (the auxiliary model). The auxiliary model completely independent of the theoretical model is used to compare the performance estimated on the real data and simulated data and it does not need to be an accurate description of the data-generating process. This auxiliary model highlights essential data characteristics, striking a balance between being too broad, which no model could meet, and too narrow, which would fail to filter out poor models effectively. These features can include moments, impulse response functions (IRFs), vector autoregression (VAR) coefficients, or other metrics like scores. This approach is a generalization of the Simulated Method of Moments.

Here is a brief explanation of the steps of the algorithm. The following two steps are the main process during the selection of the auxiliary model.

Step 1: Select the form of the auxiliary model. The auxiliary model can take the form as Vector Autoregression(VAR), Vector Error Correction Model(VECM), Vector Autoregression with Exogenous Variables(VARX), Vector Autoregressive Moving-Average(VARMA), Impulse Response Function(IRF) and moments. If the structural model is correct, then its predictions about the time series properties of the data should match those based on actual data. The state-space representation of the log-linearized DSGE model in general has a restricted VARMA representation for the endogenous variables. It can be approximately rewritten by a finite order reduced from the VAR model. A level VAR can be used if the shocks are stationary. If the shocks or exogenous processes are non-stationary. Non-stationary exogenous processes will drive one or more structural equations to have non-stationary residuals. Since these shock processes are backed by actual data and calibrated parameters, and if we treat these processes as observable variables then the number of cointegrating vectors will be less than the number of endogenous variables. This allows one to represent the solution of the estimated model as a VECM in which the nonstationary residuals appear as observable variables, and to use an unrestricted version of this VECM as the auxiliary model. The VECM model is an approximation of the reduced form of the DSGE model and can be represented as a cointegrated VARX model, see Appendix C.3 for more explanation on the use of the non-stationary data and VARX model.

Step 2: Specify the variables and the order of the auxiliary model. After the selection of the model's form, the variables and order of the model need to be specified. Usually, the order used is one period lag with a limited number of key macro variables. By raising the lag order and increasing the number of variables, the stringency of the overall test of the model is increased significantly. If the structural model is already rejected at order 1, there is no need to conduct a more stringent test based on a higher order. Le et al. (2011) shows the normalized Mahalanobis Distance gets steadily larger, indicating a steadily worsening fit, as the lag order is increased. In fact the general representation of a stationary log-linearised DSGE model is a VARMA, which would imply that the true VAR should be of infinite order, at least if any DSGE model is the true model. However, for the same reason that we have not raised the VECM order above one, we have also not added any MA element. As DSGE models do better in meeting the challenge this could be considered. We could start with a Direct Wald with order 1 in an auxiliary VAR or VARX (from a VECM) and probably increase the order of VAR if the test power is low.

5.2.2 Simulate the data and calculate the Wald statistics

Step 1: Back out residuals and innovations. Residuals in the equations without rational expectations (RE) are directly backed out by LHS(actual data) - RHS(model-fitted data). Residuals in the equations with RE will back out as the following, as suggested by McCallum (1976): Estimate a VAR of variables with expectations. Set the fitted values one period ahead of their expectations. Residuals are calculated as LHS-RHS. Most residuals obtained are non-stationary due to deterministic trends or unit roots. The steady state is driven by non-stationary variables. In practice, after simulating the model from original data in the base run, it computes the differences between the simulation data and original data to get residuals, either stationary or non-stationary. It then adds the BGP on the effects of the shocks, whereas in the version of the model, deterministic components and BGP are fixed. For example, productivity, and its non-stationarity come from both unit root and deterministic trends. Firstly, we take the difference in aggregate productivity and regress it on lagged difference variable $\Delta \ln A_{t-1}$ and a constant β_0^A to get its OLS residuals (ϵ_t^A) which are used for structural innovations for $\ln A_t$.

$$\Delta \ln A_t = \beta_0^A + \beta_1^A \Delta \ln A_{t-1} + \epsilon_t^A \tag{5.6}$$

For residuals of other equations, we regress each on a time trend to get the fitted value, denoted as $\hat{x}_t$, which is our time-detrended equation residual. Time detrending can be skipped if the residuals don't show any time trend. Then regress $\hat{x}_t$ on $\hat{x}_{t-1}$ to obtain the stochastic process and the residuals are structural innovations for the error $\hat{x}_t$

$$\hat{x}_t = \beta_1^x \hat{x}_{t-1} + \epsilon_t^x \tag{5.7}$$

STEP 2: Bootstrap the residuals. Firstly, define a time vector to ensure all the errors in each period will be randomly chosen together due to the interactive volatilities of errors, in which some occasional co-movements occur in certain unobservable environments. The bootstrapped samples in general do not follow the same distributions if these innovations in original samples are not i.i.d. However, we do not release the i.i.d. assumption. We could also use a random generator to randomly draw elements from the time vector to yield a bootstrapped time vector with the same dimension, also resulting in a bootstrapped sample of structural errors. We bootstrap structural innovations instead of randomly drawing innovations from assumed distributions because firstly the interactive volatilities can hardly be realised using single distributions. Secondly, the achieved innovations might imply these values have higher realisation probabilities (same idea as the likelihood), which cannot be reflected by random draws from given distributions. Furthermore, what distributions should be used? Bootstrap is indeed based on the large sample theorem where the original sample of structural innovations has the same distribution with population structural innovations and the sample size in our model is insufficiently large. Le et al. (2011) conduct Monte Carlo experiments to evaluate the accuracy of innovation bootstrap by firstly setting a normal distribution for the original structural errors (equivalent to a distribution of original innovations), randomly drawing innovations, and bootstrapping samples to calculate Wald for each sample. It is found that small sample bootstrap is also accurate, although bootstrap is likely to "under-reject" and is more accurate in large samples. Additionally, they also verify that Id-If Wald based on bootstrapped simulations converges to an χ^2 distribution as the sample size rises using a Monte Carlo experiment.

Step 3: Simulate the data. This simulation is called "Full simulation", different from a simulation for IRFs where innovations are non-zero only in the starting period (regardless of the lagged periods preparing for base run and simulation). The bootstrapped innovations are used to renew structural errors following their stochastic processes achieved in step 2.2. Then the model is solved following the same procedures in 2.3 with two differences that innovations now are non-zero and the "Type II Fix" residuals are added to equations in each period to diminish unobservable uncertainties. The simulated outputs are variable values when the model is solved. The net BGP growth rate generally is quite small for quarterly data, e.g. 0.002-0.004 but significant for annual data. Moreover, BGP has almost no effects on Wald computation. However, it is better to keep this step to compare simulated data with actual data which implicitly contains a time trend.

Step 4: Minimize the Wald statistics. Lastly, Wald and the Transformed MD are calculated after we collect S (1000 in practice) bootstrapped simulations. Define the structural parameter vector and the auxiliary parameter vector as θ and β respectively.

- Estimated auxiliary parameters $\hat{\beta}$ with actual data in the auxiliary model.
- Simulate S samples using the structural model with a given θ.
- Estimate $\tilde{\beta}_{s,\theta}, (s = 1, 2, \cdots, S)$ with simulated data in the auxiliary model.
- The II estimator of θ is the one such that $\hat{\beta}$ and $\bar{\bar{\tilde{\beta}}}_{s,\theta}$ are closest.

Solving the problem to minimize the Wald statistic yields the optimal estimator θ:

$$\min_{\theta} [\hat{\beta} - \bar{\bar{\tilde{\beta}}}_{s,\theta}]' \hat{\Omega}^{-1} [\hat{\beta} - \bar{\bar{\tilde{\beta}}}_{s,\theta}] \tag{5.8}$$

where

$$\bar{\bar{\tilde{\beta}}}_{s,\theta} = \frac{1}{S} \sum \bar{\bar{\tilde{\beta}}}_{s,\theta} \tag{5.9}$$

To make the results more intuitive, As Wald statistics follows

$$\text{Wald}_s = [\hat{\beta} - \bar{\bar{\tilde{\beta}}}_{s,\theta}]' \hat{\Omega}^{-1} [\hat{\beta} - \bar{\bar{\tilde{\beta}}}_{s,\theta}] \sim \chi^2(k) \tag{5.10}$$

$\sqrt{2\chi^2}$ asymptotically follows a normal distribution as $N\left(\sqrt{2k-1}, 1\right)$ and $\{\sqrt{2\text{Wald}} - \sqrt{2k-1}\}$ is close to a t distribution, where k is the number of coefficients in the auxiliary model. Furthermore, a transformed Mahalanobis Distance is defined as below

$$MD = T_c \left[\frac{\sqrt{2 \text{ Wald }_a} - \sqrt{2k-1}}{\sqrt{2 \text{ Wald }_c} - \sqrt{2k-1}} \right] \tag{5.11}$$

where T_c is the critical value of one-tail t distribution on the c% confidence interval, $Wald_a$ is the Wald statistics calculated from the actual data, $Wald_c$ is the c^{th} percentile value of the sorted $\{Wald_s\}$. Then it is compared with the critical value of t distribution with a large degree of freedom on a chosen confidence interval. In practice, the model isn't rejected based on a set of parameters if the MD is less than 1.645 with 95% confidence interval. Based on this, a P value is calculated as

$$\text{P value } = \frac{(100 - \text{ the Wald percentile })}{100} \tag{5.12}$$

5.2.3 Estimate the structural coefficients

Assuming the model is true, we use the II test given continuously adjusted values of parameters until the model cannot be rejected. That is to find the parameter values such as the transformed MD (TMD) or Wald is minimized and this procedure of searching for optimal parameters is called II Estimation. It is an advantage of II that estimation and test are not independent. II estimation tells us the optimal parameters such that the difference between simulated data and the actual data is minimised while II test tells us how good this minimised difference is. Suppose the minimized difference can still be strongly rejected. One could believe that the functional forms of the model are problematic.

The following are basic algorithms of II estimation. Start with a set of calibrated coefficients, set boundaries according to the literature and then randomly generate N sets of coefficients with different algorithms. Then each set of the coefficients will be put into our structural model to simulate the data for endogenous variables. The difference compared to Bayesian estimation is that we use some auxiliary models like VAR/VARX/VECM to fit the simulated data and the actual data. Then we have the difference of the estimated coefficients from the auxiliary model between simulated data and actual data, for each input parameter, we run 1000 simulations. Then we calculate the Wald statistics. The best set of coefficients will have the smallest Wald statistics which means the best coefficients minimised the distance between actual data and the simulated data, which also means our structural model can fit the actual data within the auxiliary framework. The advantage of this methodology is that because we run 1000 simulations for each set of coefficients, we can plot the distribution of distance in the simulations and calculate the P value which could be used to check the fitness of the model statistically. See Figure 5.1 for a quick reference.

Here we briefly introduced some numerical methods we can use to estimate the coefficients. Further details can be found in Appendix D.4.

Random selection method. This methodology is operationalized by delineating boundaries around the calibrated values of the coefficients. These boundaries typically reference established economic principles or conventions found in the scholarly literature. Subsequent to this, a random selection between the upper and lower bounds is undertaken. Another approach is termed grid random selection. After determining the bounds for the coefficients, each decimal is partitioned into ten units. The algorithm identifies coefficients associated with the minimal Wald statistics and then progresses to the subsequent decimal.

Simulated-Annealing method. We start from initial parameter values θ_1 and conduct a full II test to obtain TMD_1 (or $Wald_1$). Then we randomly generate a neighbour parameter vector θ_2 and follow the same steps to obtain TMD_2 (or $Wald_2$). If $TMD_2 < TMD_1, \theta_1$ is chosen as starting parameters and we repeat the two steps above to search for a better one. If $TMD_2 \geq TMD_1$, we do not surely reject to move towards θ_2, but calculate an "acceptance probability" (ACP) which provides a probability to move. Namely, we always surely move to a better new choice, but we still have a certain chance to move towards a worse choice, which is our mechanism to escape from a local optimum. The ACP generally takes the form of $\alpha(TMD_1, TMD_2, T) \in [0, 1]$ where T is called "temperature" which usually starts from unity and decreases after each estimation iteration. ACP could follow Gibbs equation $\alpha = minexp[(TMD_1 - TMD_2)T^{-1}], 1$ in practice where $T_{t+1} = qT_t, 0 < q < 1$. We randomly draw a value u from a uniform distribution $U(0, 1)$ in each estimation iteration. If $\alpha \geq u$, we move to θ_2; if else, we stay at θ_1. Note that the process of T implies that it would be more difficult to move towards a worse choice as more iterations are done. Repeat the steps above until we find an acceptable solution for or reach a maximum number of iterations.

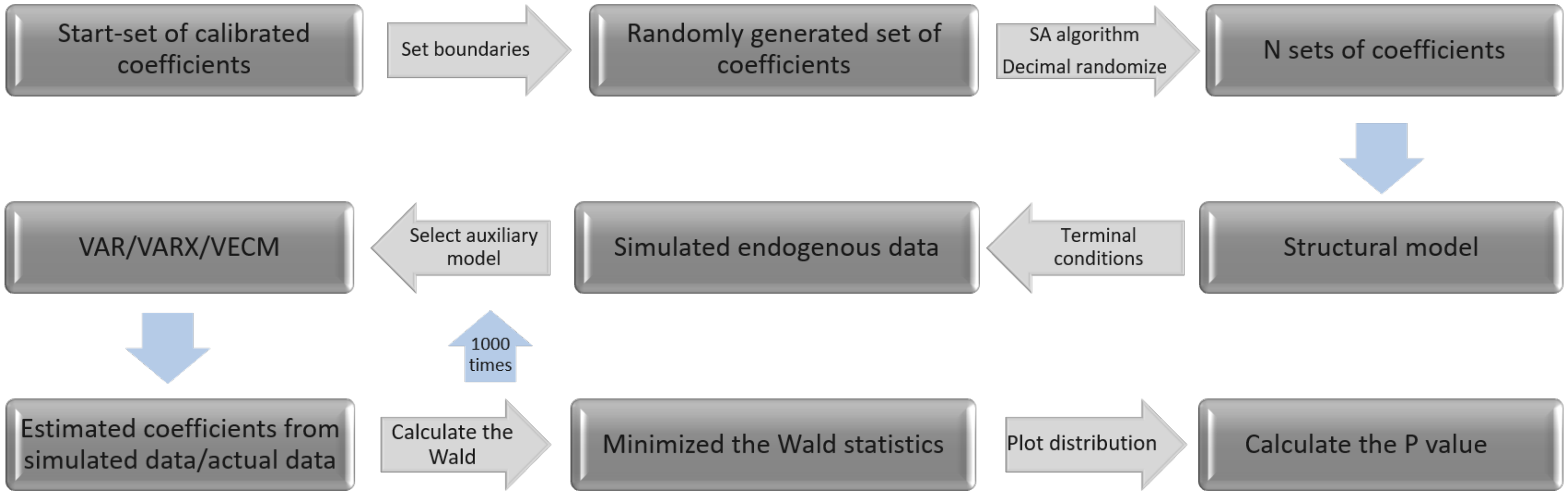


Figure 5.1. Process of Indirect Inference estimation

5.2.4 Test the power of estimation

The size of a test (or significance level, α) is the probability of incorrectly rejecting the null hypothesis when it is true. This is the probability of making a Type I error. The power of a test $(1-\beta)$, is the probability of correctly rejecting the null hypothesis when it is false. A Type II error occurs when one fails to reject a false null hypothesis. For a given hypothesis and test statistic, one constrains the size of the test to be small and attempts to make the power of the test as large as possible. The power of the Wald test can be checked by positing a variety of false models with either parameter mis-estimation or model mis-specification.

Step 1a: Parameter mis-estimation. Generate the falseness by introducing a rising degree of numerical misspecification for the model parameters, moved ±x% away from their true values. The degree of error in the beliefs, whether of parameters or errors, cannot be obviously identified from just examining the data sample; from a given sample with a given θ one can extract different error processes with different methods so that there is no uniquely determined set of ρ's and error moments; similarly, θ can also be variously estimated or calibrated. Thus, for any given data sample the information in the sample itself will be consistent with a wide range of parameters and error moments.

Step 1b: Model misspecification. Generate the falseness by replacing some of the equations with different set-ups. One can consider replacing the equation of the main mechanism in the model with an AR process.

Step 2: Monte Carlo simulations. Monte Carlo simulations can be used to test both parameter mis-estimation and model misspecification. For both mis-specifications,

- Generated 10000-sample data from the estimated model, the True model.
- Calculate the Wald statistics and find the distribution of the Wald for these True samples.
- Generate a set of 10000-sample of the endogenous from the False model.
- Calculate the Wald statistics and find the distribution of the Wald for these False samples.

- Calculate how many of the actual samples from the True model would reject the False Model on this calculated distribution with 95% confidence.

This gives us the rejection rate for a given percentage degree x of misspecification.

5.2.5 Test the robustness of bootstrap

A particular concern with the bootstrap has been its consistency under conditions of near-unit roots. Several authors e.g.,Basawa et al. (1991), Hansen (1999) and Horowitz (2001) have noted that asymptotic distribution theory is unlikely to provide a good guide to the bootstrap distribution of the AR coefficient if the leading root of the process is a unit root or is close to a unit root. This is also likely to apply to the coefficients of a VAR when the leading root is close to unity and may therefore affect indirect inference where a VAR is used as the auxiliary model, as here. The following steps use the Monte Carlo experiment to check whether this was a problem in this model.
Step 1: Drew random samples from the innovations in these error processes, creating 1000 artificial samples of the same length as the original data – 147 observations.
Step 2: Bootstrap each of these samples 1000 times.
Step 3: For each sample they computed the Wald statistic generated by the bootstraps to check whether the model is accepted or rejected at various confidence levels.

With the Likelihood Ratio, we test the bootstrap's accuracy by taking the same model and generating 1000 samples by Monte Carlo. For each one, we generate the bootstrap distribution of the LR (thus we bootstrap that sample's innovations to obtain 1000 data sets; and for each we generate the VAR forecast errors as well as the model's forecast errors, using LIML residuals extracted from the data as in our usual procedure). Then we check whether the true sample LR is rejected on the bootstrap distribution. Over the 1000 true samples, we obtain the bootstrap rejection rate. We do not use the bootstrap for the λ tests (the distribution of λ is generated by the Bayesian stochastic simulation routine) so its accuracy does not arise.

5.2.6 Comparison of Indirect Inference and Bayesian Methods

Indirect Inference and Bayesian Methods are two prominent methods for testing and estimating DSGE models. Indirect Inference is a simulation-based estimation technique. It involves comparing the simulated data generated by the model with actual data using auxiliary models or statistics. Bayesian methods incorporate prior beliefs about the parameters and update these beliefs using the likelihood of observing the data given the model. This results in a posterior distribution of the parameters. See Table 5.1 for a comparative summary.

In general, Indirect Inference follows the following steps, see details in section 5.2.1, 5.2.2, and 5.2.3 .

1. **Model Specification:** Specify the DSGE model.
2. **Simulation:** Simulate data with the model using initial parameter estimates.
3. **Auxiliary Model:** Estimate an auxiliary model (e.g.,VAR) on both actual data and simulated data.
4. **Comparison:** Compare the estimated parameters of the auxiliary model derived from the actual data with those obtained from the simulated data to determine whether the behavior or moments of the auxiliary model, based on the actual data, fall within the 95th percentile of the distribution generated from the simulated data.
5. **Parameter Adjustment:** Adjust the parameters of the DSGE model to minimize the discrepancy(measured by Wald statistics) between these estimates.

Bayesian method mainly follows the following process:

1. **Prior Specification:** Specify prior distributions for the DSGE model parameters.
2. **Likelihood Calculation:** Calculate the likelihood of observing the data given the DSGE model.
3. **Posterior Distribution:** Combine the prior distributions and the likelihood using Bayes' theorem to obtain the posterior distribution of the parameters.

4. **Estimation:** Use numerical methods (e.g., Markov Chain Monte Carlo, MCMC) to sample from the posterior distribution and estimate the parameters.

Table 5.1. Comparison of Indirect Inference and Bayesian Methods

Feature	**Indirect Inference**	**Bayesian Methods**
Approach	Simulation-based	Likelihood and prior-based
Estimation	Comparison through auxiliary models	Posterior distribution via Bayes' theorem
Computational Demand	High due to extensive simulations	High due to MCMC or other sampling techniques
Flexibility	High, robust to model misspecification	Moderate, depends on the correct specification of likelihood
Prior Information	No incorporation of prior beliefs	Incorporates prior information
Uncertainty Measurement	Limited, focuses on point estimates	Comprehensive, provides full posterior distributions
Diagnostic Capability	Formal Wald test	Facilitates model comparison through marginal likelihoods
Sensitivity	Sensitive to choice of auxiliary model	Sensitive to choice of priors

Indirect Inference is a robust and versatile simulation-based estimation technique that stands out for its ability to handle models with intricate structures and multiple sources of economic shocks. By simulating data from the DSGE model and comparing it with actual data through auxiliary models or statistics, this method ensures a comprehensive validation process. One of the main advantages of Indirect Inference is its flexibility; it does not require the exact specification of the likelihood function, making it resilient to model misspecifications and highly adaptable to different types of economic models. Additionally, it serves as an excellent diagnostic tool, highlighting discrepancies between the model and real-world data, which can guide improvements in model specification. However, Indirect Inference has significant computational demands, requiring extensive simulations that can be time-consuming and resource-intensive. The method also has a sensitivity to the choice of the auxiliary model; different auxiliary models may lead to different parameter estimates, potentially affecting the robustness of the results. Furthermore, since the method provides an indirect fit through these auxiliary models, it may not capture all features of the data, potentially leaving some aspects of the model's performance unexplored.

Bayesian Methods provide a thorough and systematic approach to estimating DSGE models by integrating prior information with observed data through Bayes' theorem. This results in a posterior distribution that offers a comprehensive measure of parameter uncertainty, a notable advantage over traditional point estimates. The ability to incorporate prior beliefs is particularly beneficial when dealing with limited or incomplete data, allowing for more informed and accurate estimates. Bayesian Methods are also adept at facilitating model comparison through the computation of marginal likelihoods, enabling researchers to assess the relative performance of different models. Despite these strengths, Bayesian Methods come with their own set of challenges. They are computationally intensive, often requiring sophisticated numerical techniques such as Markov Chain Monte Carlo (MCMC) to sample from the posterior distribution. This can be both time-consuming and computationally expensive. The results are also sensitive to the choice of priors, which can heavily influence the estimates, especially in situations with scarce data. Additionally, Bayesian Methods necessitate the correct specification of the likelihood function, a requirement that can be difficult to fulfil for complex models with numerous parameters and intricate relationships. This need for precise specification adds another layer of complexity to the Bayesian estimation process.

Both Indirect Inference and Bayesian Methods offer valuable tools for estimating and testing DSGE models, each with its strengths and weaknesses. Indirect Inference is particularly useful for its flexibility and robustness in modeling misspecification, while Bayesian Methods provide a comprehensive framework for incorporating prior information and measuring parameter uncertainty. The choice between these methods depends on the specific requirements of the analysis, computational resources, and the availability of prior information. The preference for indirect inference in this work stems from the fact that Bayesian methods would only be preferable if some priors were known to be true, which is not the case here. These conclusions are supported by Monte Carlo experiments, as detailed in Le et al. (2016), Meenagh et al. (2018), and Meenagh et al. (2021).

Chapter 6

Empirical analysis

This chapter offers an empirical examination of the model using United Kingdom data spanning from 1870, grounded in the methodological framework delineated in the preceding chapter. Initially, the model is tested by a series of simulations to check the ability to mimic the long-term inequality-growth dynamics and the business cycle properties found in the UK data. Then the model is estimated through indirect inference techniques. Subsequently, a power test is administered to ascertain the robustness of both the estimations and the bootstrap resamples. The chapter culminates in an exploration of policy implications, specifically focusing on taxation and welfare.

6.1 Simulations

This section first demonstrates the model's capability to replicate the UK's inequality-growth dynamics discussed in the introduction. Following this, it illustrates the model's proficiency in capturing the stylized facts of business cycles.

6.1.1 Tendency simulation

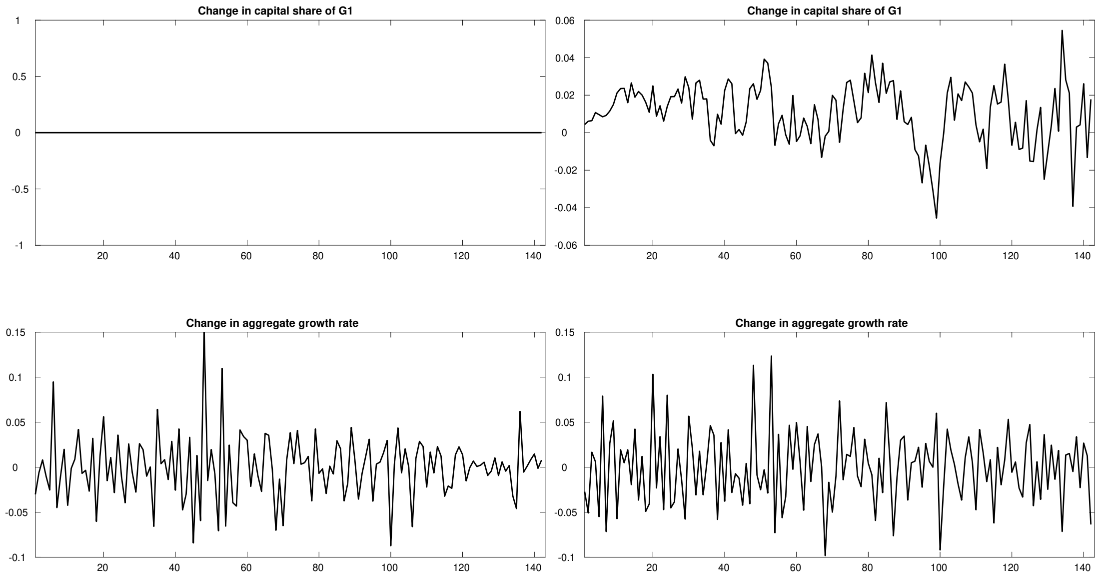
The tendency simulation rigorously assessed the model's capability to replicate the nuanced long-term dynamics between inequality and growth. This economic scenario began with two homogeneous groups, initially devoid of wealth disparity, ensuring a baseline where all individual variable values were consistent across both cohorts. Crucially, the expected values of individual consumption, generated from an AR(1) process and denoted as $E_tC_{1,t+1} = E_tC_{2,t+1}$, serve as a foundational assumption that supports the integrity of the initial conditions. This benchmark model is set as a tax-free system.

As the simulation progresses, the emergence of inequality is attributed solely to innovations in individual labour input. These innovations are not arbitrary but are sourced from independent normal distributions with an identical mean for both groups. This design choice underscores that any resultant inequality is driven by inherent differences in labour input rather than systemic biases or initial conditions, thereby isolating the effects of individual effort on economic outcomes.

Each group is given an equivalent probability of ascending to affluence, emphasizing the model's alignment with meritocratic principles. This aspect of the simulation is particularly vital as it reflects real-world economic ideologies where upward mobility is theoretically based on individual endeavours rather than predetermined conditions.

The results depicted in Figure 6.1, which illustrate the evolution of growth and inequality over a span of 150 years, are not just numerical outputs but a compelling narrative about how small, random differences in productivity can amplify over time to create significant disparities. The positive mean of 0.02 in labour input innovations, as derived from the UK's longitudinal data, distinctly catalyzes dual phenomena—escalating capital inequality and bolstered economic growth, vividly illustrated in Figure 6.1(b). These innovations are not isolated events but accumulate, building on the momentum provided by previous increments. This cumulative nature ensures a sustained elevation in labour input for each group, propelling them towards a mean level that gradually mirrors long-term expectations. Consequently, as these mean levels stabilize, the growth rate moderates.

In contrast, scenarios with a zero-mean shock present a starkly different dynamic: variations in inequality disappeared absolutely and lack a coherent trend with economic growth, as depicted in Figure 6.1(a). This distinction underscores the critical role of the mean shock in shaping economic trajectories. The inherent design of our model reveals that initial random variations fostering inequality are not counterbalanced by subsequent randomness. Instead, the initial disparities in wealth distribution establish a foundation that perpetuates and even intensifies inequality, primarily



(a) Simulation with 0-mean labour innovations (b) Simulation with 0.02-mean labour innovations

Figure 6.1. Tendency simulations with different labour input innovations

benefiting the initially wealthier group. This self-perpetuating cycle of growth and inequality is driven by a mechanism where each new increment in wealth further consolidates the disparity, thereby resisting efforts toward equalization.

This process, once initiated, becomes self-reinforcing. Subsequent innovations that might theoretically reduce inequality serve only to temporarily dampen the established cycle of unbalanced growth. The result is a persistent disparity in capital distribution between Group 1 and Group 2, observed consistently across various simulations. The pivotal factor is the subtle yet deterministic growth rate induced by the mean shock; when paired with random variability, it triggers a disproportionately significant impact among those who, by chance, initially accumulate more wealth.

The nonlinear response of economic growth to initial capital inequality creates a dynamic interplay that is crucial for activating this growth-inequality mechanism. Our findings, robustly supported by simulations rooted in empirical UK data and detailed in Section 6.2, not only lend substantial credibility to our model but also offer profound insights into economic policy and theory. They suggest that seemingly minor adjustments in labour input distribution can have extensive and lasting impacts on the trajectories of economic inequality and growth.

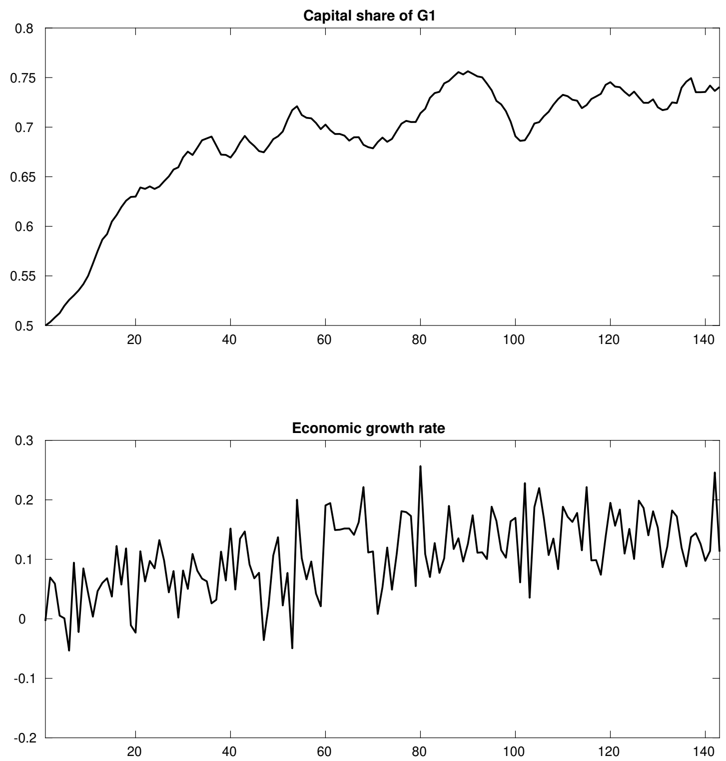


Figure 6.2. Tendency of aggregate growth and capital inequality with all relevant shocks

When all shocks are randomly applied to the base model, either affects the growth rate (equal shocks to π_{it}, A_t) or inequality (different shocks to π_{it} or to A_t), or both (aggregate consumption shocks), the inequality-growth dynamics are illustrated as Figure 6.2. To ascertain the robustness of the findings, 1000 simulations were undertaken. The outcomes indicate that the simulated correlation coefficients(from 0.17-0.42) closely align with those derived from UK data in Table 1.1.

6.1.2 Impulse response

The main concern in this thesis is how a shock to the penalty rate, ϵ_t^{π}, affects both individual behaviours and the aggregate economy. Theoretically, if the entrepreneurship penalty rate, π'_{it}, falls due to a negative shock, both groups (G1 and G2) will be more inclined to engage in entrepreneurship, accepting a slight reduction in consumption. However, due to the differing sensitivities in individual productivity to the penalty rate changes—where G1 is more responsive than G2—the capital distribution will become increasingly unequal. Specifically, the productivity of individuals in G1 increases more significantly than in G2 when the penalty rate decreases, leading to a greater accumulation of capital in G1. Consequently, while aggregate output and overall economic growth rise due to the increased entrepreneurial activity spurred by the reduced penalty rate, inequality also rises because the benefits of the shock are not evenly distributed across the two groups.

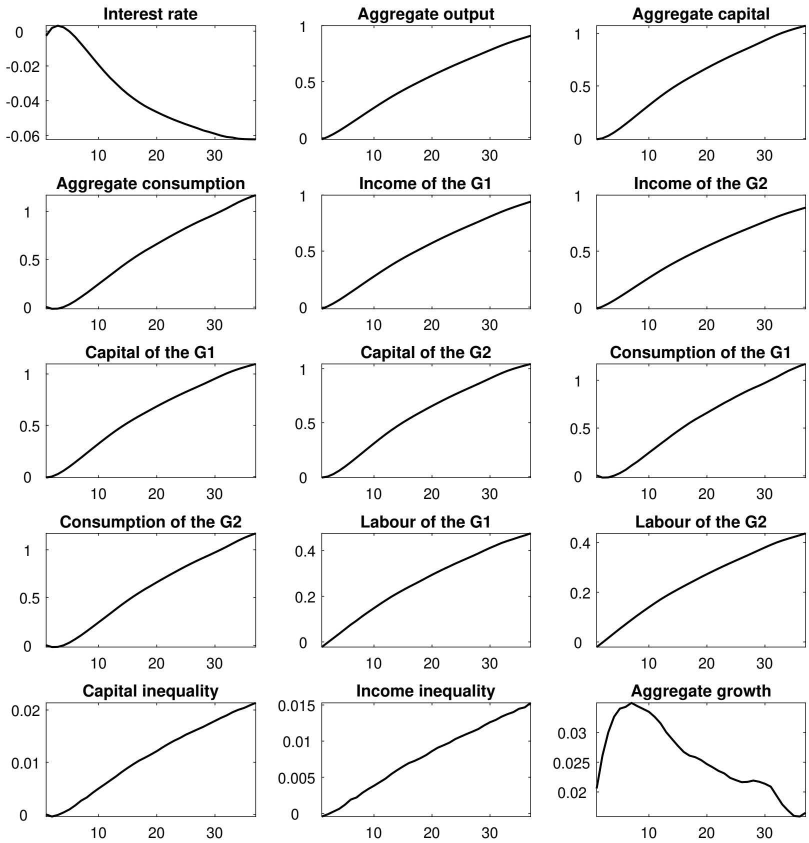


Figure 6.3. IRFs to negative one S.D. of aggregate penalty rate shock

The impulse response functions (IRFs) in Figure 6.3 illustrate the dynamic effects of a negative standard deviation shock to the aggregate penalty rate on the economy. This shock incentivizes all groups to increase their entrepreneurial efforts, thereby boosting overall economic growth. The transmission mechanism operates as follows:

when the penalty rate decreases, the cost of engaging in entrepreneurial activities is reduced, making entrepreneurship more attractive to individuals across all groups. Consequently, both G1 and G2 increase their entrepreneurial activities, leading to higher levels of investment and innovation.

However, the impact of the reduced penalty rate is not uniform across groups. G1 individuals are more sensitive to changes in the penalty rate due to their higher baseline productivity and greater capacity to leverage reduced costs effectively. As a result, G1 benefits disproportionately from the shock. This increased productivity and investment from G1 individuals lead to a significant accumulation of capital within this group, while G2 also benefits but to a lesser extent. Over time, this differential sensitivity results in a more unequal distribution of capital and income, thereby increasing inequality.

In terms of economic variables, the negative shock to the penalty rate leads to several significant changes. The permanent reduction in the penalty rate lowers the cost of borrowing, resulting in a decrease in interest rates. This stimulates further investment as cheaper borrowing costs encourage more capital expenditure. Aggregate output increases as the enhanced entrepreneurial activity boosts production capabilities. Similarly, aggregate capital rises due to the higher levels of investment. Aggregate consumption also increases as the economy grows and individuals have more resources at their disposal. The overall effect is a sustained increase in aggregate growth, driven by the enhanced productivity and investment spurred by the lower penalty rate.

The IRFs clearly show that the negative shock to the penalty rate results in permanent improvements in aggregate economic variables such as output, capital, consumption, and growth. However, the unequal distribution of these benefits, with G1 gaining more significantly due to their higher sensitivity to the penalty rate changes, leads to increased inequality. This detailed explanation of the transmission mechanism clarifies why the reduction in the penalty rate results in both higher growth and higher inequality.

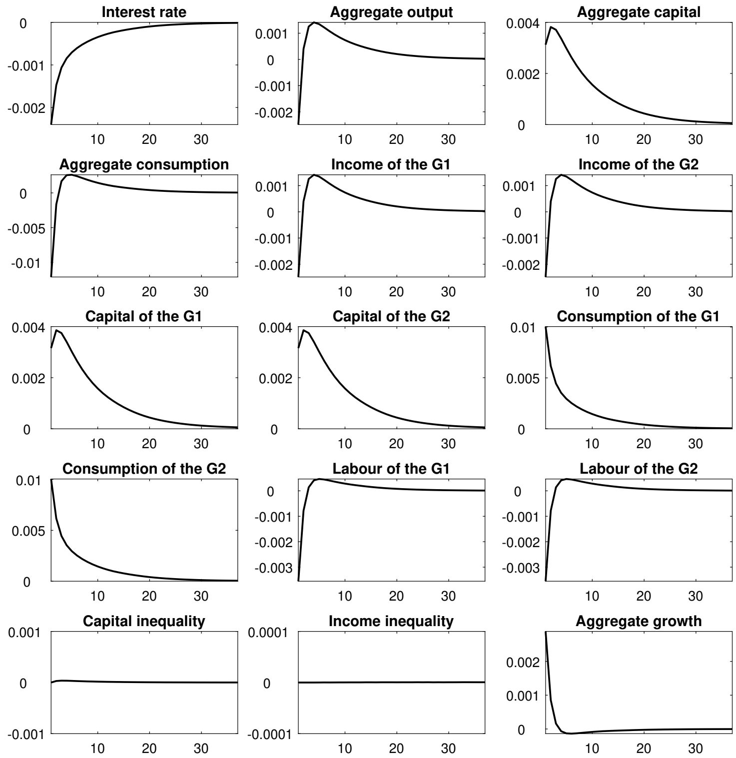


Figure 6.4. IRFs to one S.D. of G2's consumption shock

To evaluate whether the model can effectively respond to a variety of shocks, nine additional shocks were applied to the model. Figure 6.4 specifically illustrates the impact of a temporary shock on G2's consumption, which is representative of a typical demand-side shock. This shock has no lasting effect on inequality and only a temporary impact on growth and other economic variables. The transmission mechanism for this shock can be explained in detail:

When G2 experiences a temporary increase in consumption, it leads to a rise in aggregate demand. This initial surge in demand triggers an immediate response from the central bank, which lowers interest rates to stabilize the economy. The reduction in interest rates is aimed at encouraging investment to meet the increased consumption demand.

However, because the shock to G2's consumption is temporary, the increased consumption is not sustained over time. As G2's consumption returns to its normal level, the temporary boost in aggregate demand fades. This results in a subsequent drop in aggregate output as the economy adjusts back to its equilibrium state. Since the shock is short-lived, it does not have a lasting impact on economic inequality; the temporary changes in consumption do not alter the underlying distribution of wealth or income.

Additionally, the temporary increase in consumption by G2 leads to a short-term increase in aggregate capital. Businesses respond to the initial rise in demand by investing more, which temporarily boosts capital levels. However, as the consumption shock dissipates, aggregate consumption drops back to its original level. This means there is no long-term change in overall consumption patterns, and the increase in capital is not sustained.

In summary, the temporary shock from G2's consumption causes a brief period of increased demand, leading to lower interest rates and a short-term rise in aggregate capital. However, this effect is temporary, and once the shock dissipates, the economy returns to its previous state with no long-term impact on growth or inequality. The IRFs in Figure 6.3 clearly demonstrate these dynamics, showcasing the transient nature of changes in economic variables in response to the consumption shock.

See more IRF figures in Appendix B.1.

6.2 Estimation

For model estimation, this study employs the indirect inference method. As elucidated in the preceding chapter, the fundamental premise of indirect inference revolves around utilizing an auxiliary model. This auxiliary model aids in determining the coefficients, both from the simulated data originating from the structural model and from the actual data. If the divergence between the coefficients derived from these two datasets is deemed statistically significant, one can infer that the structural model aptly replicates the real-world data. In pursuing an understanding of the interplay between capital inequality and overall economic growth, it is imperative for the auxiliary model to encompass both aggregate output and capital inequality. Additionally, individual productivity levels are incorporated as exogenous variables to elucidate their influence on aggregate growth. The subsequent equation delineates this auxiliary VARX model:

$$\mathbb{Y}_t = \beta \mathbb{Y}_{t-1} + \alpha \mathbb{X}_t + e_t \tag{6.1}$$

where $\mathbb{Y}_t$ is $(Y_t, \mathrm{IQK}_t)'$, $\mathbb{X}_t$ is $(A_{1,t}, A_{2,t}, trend)'$ including a time trend, the estimated variance of the error term e_t, is $\hat{\Omega}_e$. Then a 6×1 coefficient vector, including 4 VARX coefficients of the lagged endogenous variables and 2 variances in Ω, is used to calculate the Wald statistics. Table 6.1 presents a juxtaposition of the coefficients derived from the auxiliary model using actual data versus simulated data. The findings indicate that all VAR coefficients reside within the 95% confidence interval generated through simulation, although two error variances deviate outside this range. This overarching pattern of empirical coefficients falling within the bounds of model-based simulation provides a compelling rationale for the test's success, while simultaneously omitting potential simulated covariation, a factor that may be salient in a comprehensive joint test.

Table 6.1. Coefficients of the auxiliary model

Auxiliary coefficients	Actual value $\hat{\beta}$	Simulated value $\tilde{\beta}$		
		Lower 2.5%	Mean $\bar{\beta}$	Upper 2.5%
β_{11}	0.84273	0.43581	0.73774	1.03644
β_{12}	-0.13728	-0.74556	-0.09392	0.65423
β_{21}	-0.02755	-0.05248	0.00564	0.06332
β_{22}	0.95810	0.79738	0.93201	1.04984
Var $(\hat{e}_1)$	0.00076	0.00416	0.01394	0.05306
Var $(\hat{e}_2)$	0.00011	0.00040	0.00170	0.00681

In the structural model, there are 20 coefficients in total. Out of these, 13 have been calibrated as constants, as detailed in Chapter 4.3. The remaining 7 coefficients are estimated through a process wherein 500 sets of these coefficients are randomly generated. These sets are constructed around the pre-calibrated values, bounded by limits determined by fundamental economic principles (refer to Chapter 5.2.3 for an in-depth discussion). Every generated coefficient set is incorporated into the structural model, yielding $S = 1000$ instances of simulated data. Subsequently, the VARX model facilitates the estimation of the auxiliary coefficients $\hat{\beta}$, as delineated in Table 6.1. The mean coefficient value, $\bar{\beta}$ is derived from the S simulated samples. Table 6.2 presents the Wald statistics, the transformed Mahalanobis Distance, and the computed P-value derived from the estimated coefficients. Given the P-value, the estimated model cannot be dismissed at a 5% significance level. Figure 6.5 displays the distribution pattern of the simulated Wald statistics that were used to derive the P-value. The degrees of freedom correspond to the count of coefficients present in the auxiliary model, numbering 6 in total.

Table 6.2. The results of indirect inference estimation

Notation	Description		Value
$-\phi_{2,1}$	Effect of Group 1's penalty rate on their productivity		-0.2121
$-\phi_{2,2}$	Effect of Group 2's penalty rate on their productivity		-0.2043
$-\rho_2^{\pi}$	Effect of capital share on individual penalty rate		-0.0010
ρ_3^{π}	Effect of the credit condition on individual penalty rate		-0.0842
Ψ_1	Elasticity of consumption in the utility		0.2359
Ψ_2	Elasticity of leisure in the utility		0.3759
θ_2	Effect of entrepreneurship time on individual productivity		0.5509
Wald	9.9543	Number of samples	194
TMD	0.8994	P value	0.1082

The effect of the individual penalty rates on their productivity shows that Group 1's productivity is more sensitive to its penalty rate. This is consistent with the theory. Richer individuals typically have more opportunities available to them. If the entrepreneurship penalty rate is high, they might forgo entrepreneurship in favour of other lucrative opportunities, leading to a bigger loss in potential productivity than for someone with fewer alternative opportunities. In addition, richer individuals often have more financial cushioning, allowing them to take on riskier ventures that might have higher potential payoffs. A penalty on entrepreneurship could discourage them from pursuing such high-reward opportunities, leading to a decline in productivity. The credit ratio shows a negative correlation with the penalty rate. One of the primary hurdles entrepreneurs face is access to capital. Better credit conditions make it easier for entrepreneurs to obtain the necessary funds to start or expand their ventures. When credit is readily available, the perceived and actual penalties or barriers to entrepreneurship are lowered. With favourable credit conditions, entrepreneurs might be more willing to take risks, expand their operations, or venture into new markets. The increased activity can lead to a broader acceptance of entrepreneurship in the economy, thereby reducing institutional or regulatory penalties. The estimated marginal effect of capital inequality on the individual penalty rate, at -0.001, should not be misconstrued as indicating that the entrepreneurship penalty rate is unresponsive to shifts in capital distribution. This is attributed to the penalty rate possessing a lower order of magnitude in comparison to capital.

Comparing the elasticity of consumption and leisure in the utility, the higher elasticity for leisure suggests that, in relative terms, changes in leisure have a more significant impact on utility than equivalent percentage changes in consumption. If faced with trade-offs between additional consumption and leisure, these elasticity values suggest that individuals might be more willing to forgo a small amount of consumption for additional leisure. If policymakers wanted to influence behaviour or welfare, interventions or policies that impact leisure time (like working hours, and vacation policies) might have a more pronounced effect than those that influence consumption directly (like subsidies or taxes on goods).

The marginal impact of entrepreneurship time on individual productivity showcases a direct relationship between the duration spent on entrepreneurial ventures and one's efficiency. Specifically, with each additional unit of time (for instance, an hour) committed to entrepreneurial tasks, an individual's productivity is projected to enhance by 0.5509 units. Notably, this estimation closely aligns with the findings presented by Yang et al. (2021) who utilized quarterly data from the year 1978.

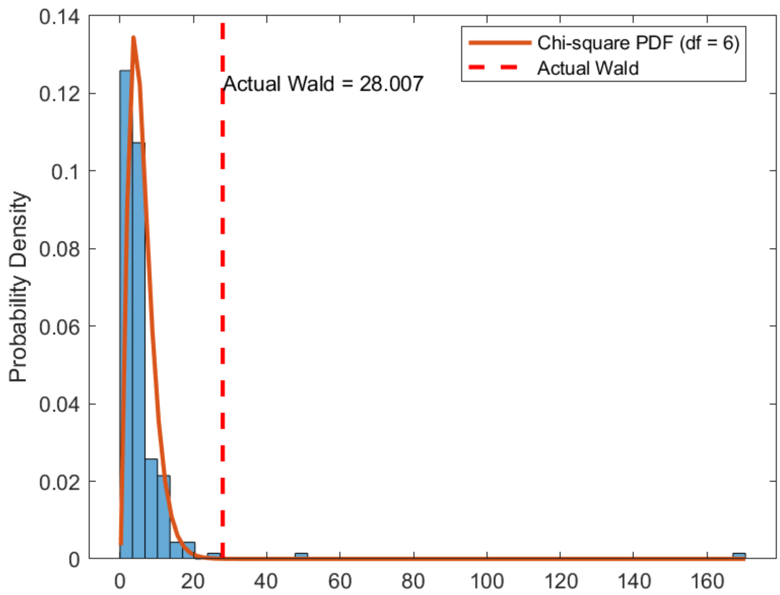


Figure 6.5. Distribution of the Wald statistics

After obtaining the estimators, we substituted the estimated coefficients into the model to calculate the residuals(errors)—the differences between the actual data and the model-implied values. The descriptive statistics and the stationarity test for the seven structural residuals are summarized in Table 6.4. The residuals of the aggregate output and aggregate capital equation are stationary based on the KPSS test. The residuals for the aggregate consumption equation are stationary with 10% significance level and trend stationary with 1% significance level in ADF tests. More specifically, the residuals for G1's consumption are stationary according to both the ADF and KPSS tests; for G2's consumption, the residuals are stationary under the KPSS test at the 1% significance level and also stationary under the ADF test at the 10% significance level. Both residuals from individual labour equations is non-stationary. Regardless of some inconsistency of ADF and KPSS test, we regressed the 7 structural errors on an AR(1) with a deterministic based on Equation 3.23. All the AR coefficients of the seven variables shown in Table 6.3 are less than one, which generally indicates these residuals can be treated as stationary or trend stationary from an empirical perspective, (Box et al., 2015; Brockwell and Davis, 2002). These regressions give us the innovations/shocks of these 7 variables, combined with the other three shocks obtained from 3.20,3.21 and 3.22, Table 6.5 shows all these 10 shocks are stationary.

Table 6.3. AR coefficients of the structural residuals and shocks

$\frac{d\,\epsilon_t^A}{d\,\epsilon_{t-1}^A}$	$\frac{d\,\epsilon_t^\pi}{d\,\epsilon_{t-1}^\pi}$	$\frac{d\,\epsilon_t^Y}{d\,\epsilon_{t-1}^Y}$	$\frac{d\,\epsilon_t^K}{d\,\epsilon_{t-1}^K}$	$\frac{d\,\epsilon_t^C}{d\,\epsilon_{t-1}^C}$	$\frac{d\,\epsilon_t^M}{d\,\epsilon_{t-1}^M}$	$\frac{d\,\epsilon_t^{C1}}{d\,\epsilon_{t-1}^{C1}}$	$\frac{d\,\epsilon_t^{C2}}{d\,\epsilon_{t-1}^{C2}}$	$\frac{d\,\epsilon_t^{N1}}{d\,\epsilon_{t-1}^{N1}}$	$\frac{d\,\epsilon_t^{N2}}{d\,\epsilon_{t-1}^{N2}}$
0.2437	0.4943	0.9892	0.9808	0.7309	0.5176	0.7093	0.3515	0.9714	0.9550

Table 6.5 shows the descriptive statistics and the results of stationarity tests of structural innovations used for bootstrapping. All the innovations are stationary.

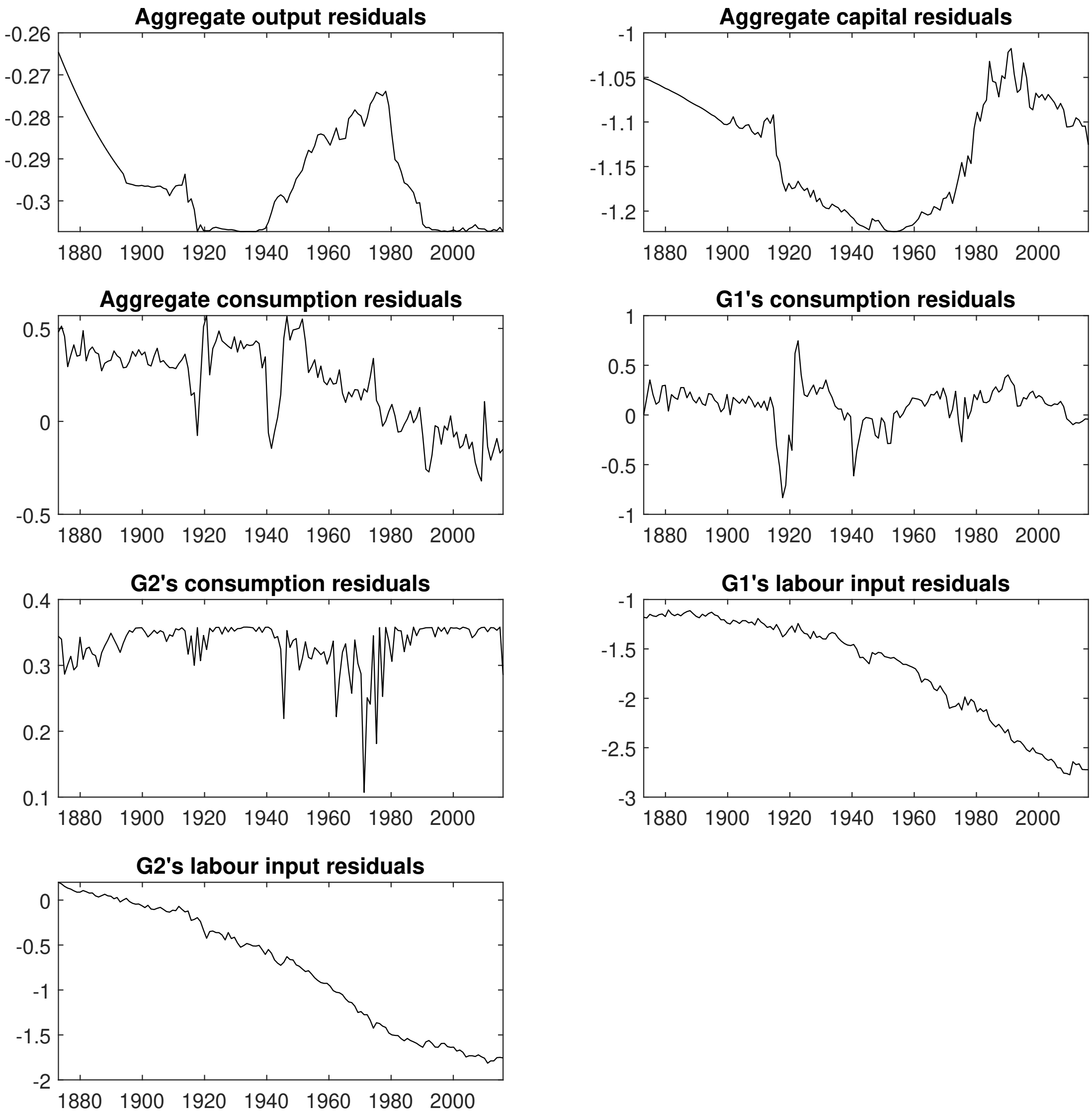


Figure 6.6. Structural residuals

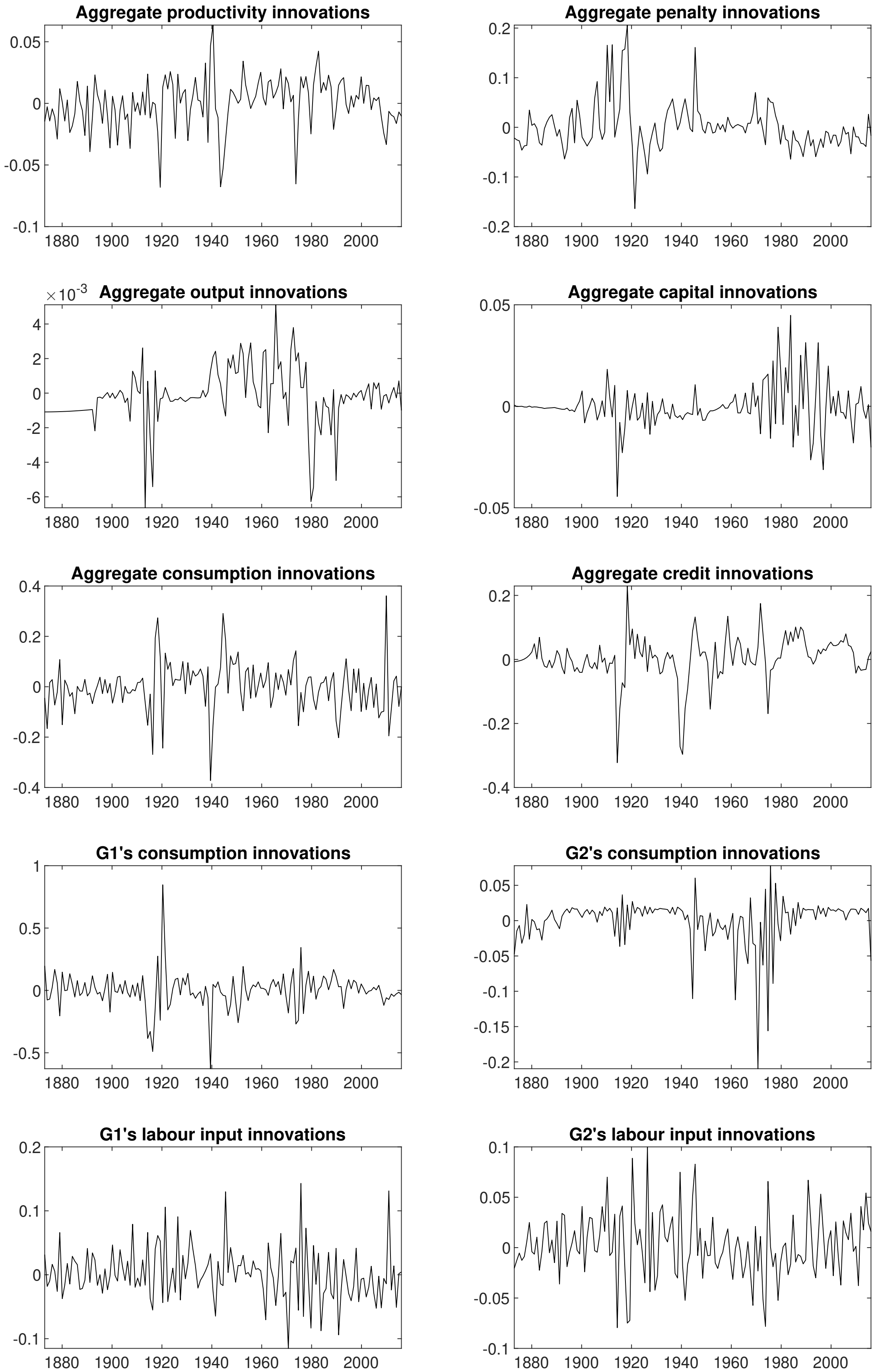


Figure 6.7. Structural innovations

Table 6.4. Descriptive statistics of the structural residuals

Series	Mean	Median	S.D.	Min.	Max.	ADF/KPSS Level(c)	ADF/KPSS Level(c,t)	ADF/KPSS Diff(c)
ε_t^Y	0.05160	0.04594	0.19926	0.025821	0.25821	-2.2503 0.2658	-2.0353 0.1591++	-8.4064*** 0.2590
ε_t^K	0.04979	0.04844	0.17710	0.03439	0.16709	-1.4889 0.3077	-1.4463 0.3065+++	-12.596*** 0.3702+
ε_t^C	0.25712	0.28392	0.14739	0.00210	0.92662	-2.8019* 1.2174+++	-4.6238*** 0.2755+++	-13.009*** 0.0774
ε_t^{c1}	0.00603	0.06735	0.09951	-0.0035	0.07839	-4.8599*** 0.1027	-4.8598*** 0.1023	-12.903*** 0.1059
ε_t^{c2}	0.39785	0.39805	0.27780	0.54990	0.54200	-2.6842* 0.1349	-2.6649 0.1395+	-12.547*** 0.0401
ε_t^{N1}	0.75659	0.82978	0.17804	0.45589	0.94598	1.5177 1.3320+++	-2.1562 0.3497+++	-14.448*** 0.6039++
ε_t^{N2}	0.23097	0.28340	0.11820	0.06640	0.83514	-0.1791 1.4060+++	-2.0179 0.2377+++	-10.524*** 0.1534

Note: For ADF tests, asterisks denote rejection of the unit root null at 10% (*), 5% (**), and 1% (***) significance levels. For KPSS tests. Plug signs denote rejection of the stationarity null at 10% (+), 5% (++), and 1% (+++) significance levels.

Table 6.5. Descriptive statistics of the structural innovations

Series	Mean	Median	S.D.	Min.	Max.	ADF Level(c)
ε_t^A	2.43E-20	0.001722	0.2044	-0.06806	0.063535	-9.2273***
ε_t^π	-0.0001	-0.006442	0.04878	-0.16379	0.206004	-4.3506***
ε_t^M	0.003498	0.004209	0.070859	0.239449	-0.32255	-6.6363***
η_t^Y	-0.00019	-0.00006	0.001653	0.005112	-0.00664	-8.5317***
η_t^K	-0.00047	-0.0006	0.011237	0.04797	-0.0448	-12.4352***
η_t^C	0.00019	0.0009	0.92653	3.36099	-0.3725	-11.4002***
η_t^{c1}	-0.00082	0.005313	0.146793	0.846817	-0.6276	-11.3317***
η_t^{c2}	-0.00024	0.010190	0.034221	0.077810	-0.20961	-2.6680*
η_t^{N1}	0.001590	-0.00229	0.041421	0.149212	-0.1515	-15.2718***
η_t^{N2}	0.000723	-0.00295	0.031819	0.099865	-0.07945	-12.5654***

Note: Unit root with a constant null at 10% (*), 5% (**) and 1% (***) significance levels.

6.3 Power and robustness

Table 6.6 presents the results of the power test against false models, where both structural parameters and the AR coefficients of the errors are systematically misestimated by ± x% increments. The probability of rejecting these false models increases sharply as the degree of falsification in the parameters escalates. This steep rise in rejection probability underscores the robustness and reliability of the estimated parameters obtained through indirect inference.

Table 6.6. Power test against numerical falsity of parameters

Parameter falseness	True	1%	3%	5%	7%	9%
Rejection Rate with 95% Confidence	5%	7%	68%	92%	99%	100%

The findings from this power test are compelling. They demonstrate that the indirect inference method is not only capable of accurately estimating the true parameters but is also highly effective in identifying and rejecting incorrect models. This is particularly important in econometric modelling, where the precision of parameter estimates directly influences the validity and predictive power of the model. By systematically introducing falsification and observing the model's response, we confirm that the indirect inference method possesses a high degree of sensitivity and specificity.

We then conducted a rigorous power test of the indirect inference (II) method against a deliberately misspecified model to further validate its robustness. In this misspecified model, we disabled the fundamental mechanism that links wealth inequality to entrepreneurship by replacing the equations governing the penalty rate (equations 3.15) with an independent AR(1) process without capital inequality and credit conditions. Consequently, in this false model, wealth inequality continues to arise from random processes but no longer influences innovation directly. Despite this critical modification, we maintained all other parameters at their fully estimated values from the benchmark model. As a result, economic growth in the misspecified model is driven solely by shocks to the aggregate penalty rate, effectively severing any linkage between growth and inequality.

Table 6.7. Frequency of rejection of misspecified model

Mis-specified model	Number of Bootstraps	Rejected samples	Rejection rate
Replace eq.3.15 by AR(1)	1000	1000	100%

This scenario presents a stringent test of the II method's power. On the same set of 1000 samples, the rejection rate for this misspecified model at a 95% confidence level was an astonishing 100%. As shown in Table 6.7, the misspecified model was rejected in virtually every instance. This near-universal rejection underscores the exceptional sensitivity and reliability of the II method in identifying and discounting models that fail to accurately represent the underlying economic dynamics.

We also rigorously evaluated the accuracy of the bootstrap method under Indirect Inference through a comprehensive Monte Carlo experiment. We began by configuring the model with error variances matching the estimated values. Subsequently, we drew random samples from the innovations in these error processes, generating 1000 artificial samples, each mirroring the original data's length of 147 observations. Each of these samples was then bootstrapped 1000 times.

For each artificial sample, we computed the Wald statistic from the bootstraps to determine whether the model would be accepted or rejected at various confidence levels. The results, displayed in Table 6.8, indicate that the bootstrap procedure is notably accurate, though it demonstrates a slight tendency toward under-rejection.

Table 6.8. Robustness of the bootstrap

Nominal rejection rate	10%	5%	1%
True rejection rate	8.7%	4.6%	0.9%

The near-alignment of true rejection rates with nominal rates underscores the method's reliability. Specifically, at a 10% nominal rejection rate, the true rate was 8.7%; at a 5% nominal rate, it was 4.6%; and at a 1% nominal rate, it was 0.9%.

The nominal rejection rate represents the expected frequency at which the model would be rejected if it were true, based on pre-defined statistical confidence levels. In essence, a 10% nominal rejection rate means we would expect to reject a true model 10% of the time due to random sampling variability, while the true rejection rates are calculated based on the sorted Wald Statistics.

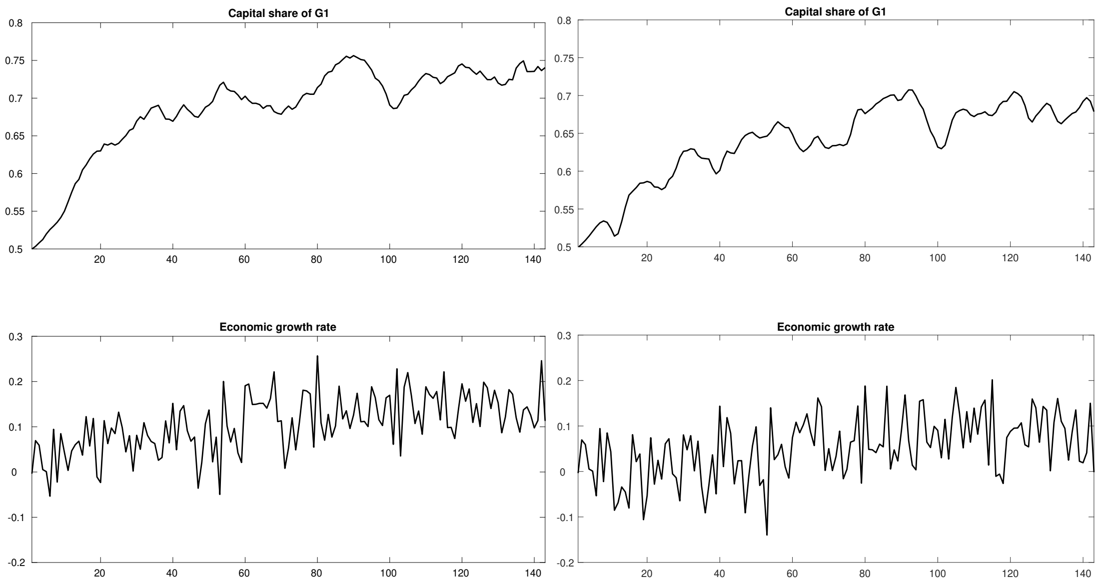
The slight under-rejection observed indicates a conservative bias, where the bootstrap method is slightly more cautious in rejecting the model. This conservatism can be advantageous in econometric analysis, as it minimizes the risk of falsely rejecting a true model, thereby enhancing the robustness of the conclusions drawn from the analysis.

The results highlight the power and robustness of the indirect inference (II) method in DSGE modelling. Tests showed that II accurately estimates true parameters and effectively rejects false models, with rejection probability increasing sharply as parameter falsity increases. A stringent power test against a misspecified model confirmed II's reliability, with an exceptionally high rejection rate. Additionally, a Monte Carlo experiment validated the bootstrap method's accuracy under II, showing true rejection rates closely aligning with nominal rates. Overall, these findings demonstrate that II is a robust and powerful tool for precise parameter estimation and model validation.

6.4 Tax Policy implications

The effect of redistribution stands as a central policy concern within the literature on growth and inequality. Income tax, as a quintessential redistributive tool, can be implemented within various frameworks. The subsequent section commences with simulations of the impact of differing income tax policies on growth and inequality. The constant proportional tax rate will be considered for simplification but with various utilisations of tax revenue: general taxation without subsidy transfer(Policy I), taxation on the rich without transfer (Policy II), and income transfer from the rich to the poor(Policy III). To compare the policy effects, we set the benchmark model as a tax-free system.

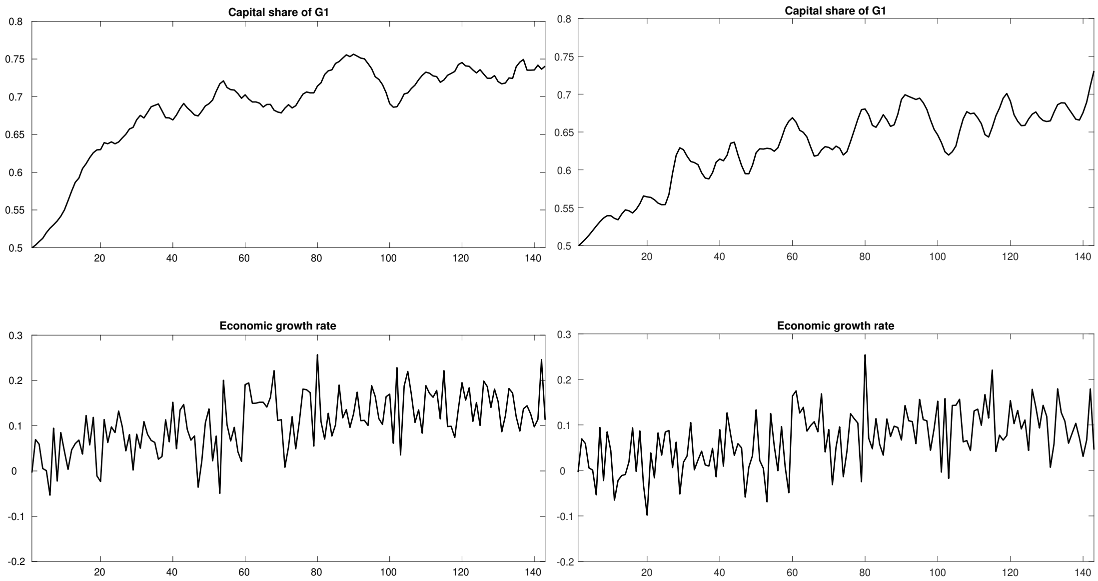
Policy I: Income tax on G1 and G2 without transfer. Consider the tax policy where both the G1 and the G2 are charged a constant income tax rate, $\tau = 0.2$, but with no individual subsidy (tax revenue all funds government spending). Figure 6.8 shows that inequality is reduced to a further lesser extent (top-left vs. top-right) while loss of growth is much higher and quite fluctuated (bottom-left vs. bottom-right). Apparently, taxing without any subsidy is harmful to the economy by reducing aggregate capital accumulation.



(a) Simulation of the benchmark model (b) Simulation with Policy I

Figure 6.8. Effects on growth and inequality with flat income tax rate

Policy II: Income tax only on G1 without transfer to G2. Now consider the policy where a constant income tax rate, $\tau = 0.2$, is solely enforced on the rich without transfer to the poor. Rich has the same φ_{21} in the previous section while the poor have the same φ_{22} as the benchmark model. Figure 6.9 also indicates that the cost of reducing inequality is still a loss of growth.



(a) Simulation of the benchmark model (b) Simulation with Policy II

Figure 6.9. Effects on growth and inequality with tax the rich only

Policy III: Income tax on G1 with transfer to G2. Now consider a more realistic policy, where the government redistributes the tax revenue to improve the living standard of the poor. The model's results for aggregate output, growth, and inequality under different income transfer rates are depicted in Figure 6.10. To mitigate the effects of randomness, the outcomes for each transfer rate are derived from the mean of 500 bootstrap samples. When juxtaposed with the baseline scenario where the transfer rate is zero, it is unsurprising to observe that income transfers lead to decreased inequality. However, this comes with the trade-off of diminished growth and, consequently, reduced long-term output. As the rate of redistribution escalates, inequality continues to decline, yet this is accompanied by an increasingly pronounced adverse impact on growth. Figure 6.10 also provides a temporal perspective on the economic ramifications of these transfers, detailing both the cumulative loss in output over time and the progression of inequality. The decline in inequality is immediate but subsequently undergoes a partial reversal. This is because as output continuously diminishes due to the reduced growth rate, the actual value of the transfer concurrently decreases. This process not only diminishes the redistributive effect over time but also progressively weakens the overall economic health.

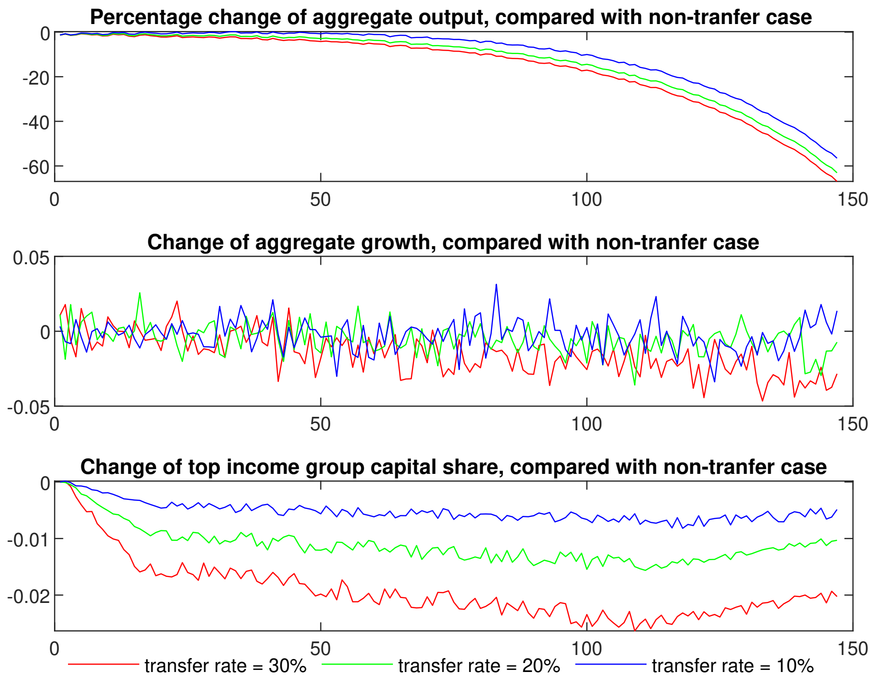


Figure 6.10. Redistributive effects by different income transfer rates over the long period

Table 6.9. Summary of the Inequality-Growth effect with different transfer rates

Transfer rate	0.1	0.2	0.3
Inequality effect	-1.73 %	-3.17 %	-5.2%
Growth effect	-0.61 %	-1.43 %	-2.9 %
Ratio	2.8	2.2	1.8

Table 6.9 shows the main results of the Inequality-Growth effect, mainly the changes in the inequalities and growth rates, with different transfer rates. The base regime was set as non-tax for the comparisons. For simplicity, assume there is no government spending for which income tax revenue needs to be raised so that the government here is purely acting on redistribution. The incremental cost of growth loss in relation to the reduction in inequality becomes steeper with increasing trans-

fer rates. Specifically, an initial move towards redistribution — shifting from a non-existent rate to a rate of 0.1 — results in a near 1.7% decrease in the capital share of the top income bracket, relative to a scenario without transfers. This, however, curtails growth by approximately 0.6%–a ratio of Inequality effect/Growth effect, nearly 2.8. Intensifying the redistributive rate to 0.2 brings about an additional 1.5% drop in the capital share of the top earners when set against the no-transfer baseline, yet depresses growth by another 0.8%, yielding a ratio 2.2. Advancing the transfer rate to 0.3 brings about a further 2% contraction in inequality and a corresponding 1% reduction in growth, a ratio of 1.8. Consequently, the incremental trade-offs between diminishing inequality and suppressing growth become more disadvantageous as redistribution rates ascend.

6.5 Welfare analysis

To elucidate the implications of these policies on both groups, a comprehensive welfare analysis was conducted employing several social welfare functions. One of the primary frameworks utilized was the average welfare concept as postulated by Benabou (2000), which captures pure economic efficiency while disregarding equity considerations. This approach focuses solely on maximizing total welfare without addressing the distribution of resources among different income groups.

In contrast, other methodologies such as utilitarian and Rawlsian welfare functions were also scrutinized. The utilitarian approach aggregates individual utilities to maximize total social welfare, thus providing a balanced view that incorporates both efficiency and some degree of equity by considering the well-being of all individuals. However, it can still be skewed towards those with higher incomes, as their marginal utility of income is lower. The Rawlsian welfare function, built on the foundation of the "veil of ignorance"(Rawls, 2017), places the utmost importance on the economic well-being of the least advantaged members of society. This approach aligns with the principle of justice as fairness proposed by John Rawls, which suggests that social and economic inequalities should be arranged to benefit the least advantaged members of society. Therefore, it emphasizes equity over efficiency, ensuring that policies are evaluated based on their impact on the poorest individuals.

By incorporating these diverse frameworks, the welfare analysis provides a nuanced understanding of the implications of economic policies. The comparison of these methodologies highlights the trade-offs between efficiency and equity, illustrating how different policy choices can lead to varying outcomes for different income groups.

To outline an optimal policy framework, we follow the approach adopted in Boar and Midrigan (2022), which reflects different redistributive priorities. The planner aims to maximize the social welfare function (SWF) defined as:

$$SWF = \left(\int \omega_i(\tau)^{1-\Delta} \, di \right)^{\frac{1}{1-\Delta}} \tag{6.2}$$

Here, ω_i represents the welfare of household i from a transfer rate τ, and $\Delta \geq 0$ is a parameter that indicates the planner's preference for redistribution. For instance, if $\Delta = 0$, the objective is to maximize the average welfare (AW):

$$AW = \int \omega_i(\tau) \, di \tag{6.3}$$

As highlighted by Benabou (2002), this objective captures pure economic efficiency without considering equity, referring to it as risk-adjusted GDP.

By setting $\Delta = \rho$, where ρ is the households' coefficient of relative risk aversion, the planner's objective becomes utilitarian welfare (UW):

$$UW = \left(\int \omega_i(\tau)^{1-\rho} \, di \right)^{\frac{1}{1-\rho}} \tag{6.4}$$

To demonstrate this, consider the utilitarian social welfare function:

$$\int U_i \, di = \frac{1}{(1-\beta)(1-\rho)} \int \omega_i(\tau)^{1-\rho} \, di \tag{6.5}$$

where U_i is the infinite lifetime utility of individual i consuming a constant stream

ω_i. Thus, the individual's welfare ω_i solves:

$$U_i = \sum_{t=0}^{\infty} \beta^t \frac{\omega_i^{1-\rho}}{1-\rho} \tag{6.6}$$

To convert this measure into a consumption equivalent, we determine the constant amount of consumption $\bar{\omega}$ each household needs to achieve the utilitarian level of welfare $\int V_i \, di$:

$$\frac{1}{(1-\beta)(1-\rho)} \bar{\omega}^{1-\rho} = \int V_i \, di \tag{6.7}$$

This implies:

$$\bar{\omega} = \left(\int \omega_i^{1-\rho} \, di \right)^{\frac{1}{1-\rho}} \tag{6.8}$$

Thus, utilitarian welfare is essentially a weighted average of individual household welfare, with weights given by each household's marginal utility, $\omega_i^{-\rho}$.

Generally, a higher Δ indicates a stronger preference for redistribution. As $\Delta \to \infty$, the objective reduces to that of a Rawlsian welfare (RW) planner:

$$RW = \min_i \left(\omega_i(\tau) \right) \tag{6.9}$$

Table 6.10 presents welfare enhancements across various transfer rates, derived from three optimized social welfare functions. Each scenario is averaged over 500 simulations to mitigate the influence of stochastic fluctuations. As transfer rates ascend from 10% to 30%, an upward trajectory is observed in the maximized welfare across all three determinations. Notably, under the Rawlsian paradigm, the welfare of the bottom 90% experiences a significant enhancement, reaching a peak of 26%. Conversely, this comes at a cost to the top 10%, whose welfare deteriorates, spanning a range from 2.6% to 16.6%. A discernible leap in welfare is observed as the transfer rate escalates from 20% to 30%. Such nuances necessitate prudence from policymakers when calibrating transfer rates within this range.

Table 6.10. Welfare gains based on different transfer rates

Transfer rates	0.1	0.2	0.3
A. Maximize Average Welfare			
Average welfare gains	12.7%	13.83%	18.58%
welfare gains, bottom 90%	14.4%	16.0%	22.1%
welfare gains, top 10%	−2.6%	−5.7%	−13.1%
B. Maximize Utilitarian Welfare			
Utilitarian welfare gains	14.14%	18.06%	23.7%
welfare gains, bottom 90%	16.2%	20.9%	27.9%
welfare gains, top 10%	−4.4%	−7.5%	−14.1%
C. Maximize Rawlsian Welfare			
Rawlsian welfare gains	16.67%	19.65%	26.42%
welfare gains, bottom 90%	19.4%	23.3%	31.2%
welfare gains, top 10%	−7.9%	−13.2%	−16.6%

Table 6.10 presents welfare enhancements across various transfer rates, derived from three optimized social welfare functions: Average Welfare, Utilitarian Welfare, and Rawlsian Welfare. Each scenario is averaged over 500 simulations to mitigate the influence of stochastic fluctuations, providing a robust understanding of the potential impacts of different transfer rates.

As transfer rates ascend from 10% to 30%, an upward trajectory is observed in the maximized welfare across all three determinations. Notably, under the Rawlsian paradigm, the welfare of the bottom 90% experiences a significant enhancement, reaching a peak of 31.2% at a 30% transfer rate. This substantial increase highlights the effectiveness of Rawlsian principles in reducing inequality and enhancing the welfare of the majority. However, this comes at a cost to the top 10%, whose welfare deteriorates significantly, spanning a range from -2.6% at a 10% transfer rate to -16.6% at a 30% transfer rate.

In detail, the average welfare gains increase from 12.7% at a 10% transfer rate to 18.58% at a 30% transfer rate. The bottom 90% see their welfare gains rise from 14.4%

to 22.1% as transfer rates increase. Conversely, the top 10% experience a decline in welfare, from -2.6% at 10% to -13.1% at 30%. Utilitarian welfare gains are slightly higher compared to average welfare gains, starting at 14.14% for a 10% transfer rate and reaching 23.7% at 30%. The welfare gains for the bottom 90% follow a similar pattern, increasing from 16.2% to 27.9%. The top 10% experience a more pronounced decline, from -4.4% at 10% to -14.1% at 30%. Rawlsian welfare gains are the highest among the three measures, starting at 16.67% at a 10% transfer rate and reaching 26.42% at 30%. The bottom 90% benefit the most under this paradigm, with welfare gains increasing from 19.4% to 31.2%. The top 10%, however, see the most significant welfare losses, from -7.9% at 10

The results suggest that increasing transfer rates can significantly enhance overall welfare, particularly for the bottom 90% of the population. However, policymakers must consider the adverse effects on the top 10%, who face substantial welfare losses. The pronounced leap in welfare as the transfer rate escalates from 20% to 30% underscores the sensitivity of welfare outcomes to transfer rate adjustments.

When juxtaposed with the tax revenue over GDP trends shown in Figure 4.4, we observe an upward trend leading up to the 1980s. The political motivation for these continuous tax reforms largely stemmed from catering to the median voter, whose welfare considerably influenced various social welfare measures. This strategic approach aligns with the observed sustained decrease in inequality throughout UK history, largely attributed to redistribution policies.

Governments, driven by political competition, tend to implement redistribution policies to appeal to the average voter. This political economy perspective explains the historical tendency of the UK to engage in more extensive redistribution. As political parties vie for votes, they adopt policies that promise to reduce inequality and enhance social welfare, thereby appealing to the median voter who typically benefits from such measures. This dynamic has resulted in sustained redistribution efforts aimed at mitigating inequality and boosting overall welfare.

However, from the 1980s onward, inequality began to rise, spurred by the Thatcher reforms. These reforms were aimed at addressing the growing demands of skilled workers for enhanced growth and elevated wages. These skilled workers emerged as a pivotal voting demographic with interests closely aligned with the affluent. This shift highlights the complex interplay between economic policies and political dynamics, suggesting that the interests of different voter demographics can significantly influence policy directions and outcomes. Thus, while historical trends towards redistribution have been driven by political incentives to please the median voter, changes in the economic landscape and voter demographics have led to shifts in these policies over time.

This analysis underscores the importance of understanding the political economy behind welfare calculations and redistribution policies. The observed welfare enhancements across different transfer rates provide valuable insights for policymakers aiming to balance growth and inequality while navigating the political landscape.

Chapter 7

Conclusion

This thesis constructs a structural model to investigate the relationship between capital inequality and economic growth in the UK. The model was built on the idea that the marginal utility cost of entrepreneurial activities decreases with rising wealth, so wealth inequality enhances the entrepreneurship incentives of the rich to stimulate growth, and the growth, in turn, aggravates inequality. The model incorporates heterogeneity by classifying the population into two groups for simplicity: the rich (the top 10%) who own higher capital holdings and the rest. This structure generates a stable tendency of the relationship between capital inequality and economic growth, which is almost independent of parameter values and population shares of individual groups, making it applicable to different countries and different population group-segmentations.

When considering the 10%–90% income segmentation, the benchmark model cannot be rejected by the Indirect Inference test and can fit the main characteristics of the UK data from 1870 to 2016. This study is notable as it is the first to apply the powerful method of indirect inference to UK long-term history, joining recent studies of UK and US postwar history. This method does not check the model's ability to predict events but rather its ability to simulate behaviour similar to that found in the data.

Compared to Yang et al. (2021), this version of the model incorporates an exogenous credit ratio, which measures the credit conditions that entrepreneurs face, modelling the changing credit constraints over two centuries of data. By analyzing the influence of external shocks on deviations of endogenous variables, we find that some shocks, such as those on the entrepreneurship penalty rate, have a substantial impact on both aggregate growth and inequality. In contrast, some shocks have almost no effect on either aspect, like aggregate consumption shock and aggregate output shock, while others moderately affect both, like a capital aggregate shock. In practice, policymakers must face a trade-off between wealth equalization and economic growth when implementing redistribution policies. The model suggests that tax policies transferring income from the top-income group to the low-income group are more efficient in balancing inequality and growth. However, as the transfer rate rises, the trade-off worsens, implying that an appropriately low tax rate is preferable for policymakers. The welfare analysis shows that under different social welfare functions, the welfare gains of the two groups vary. Policymakers prioritizing lower-income groups can improve their welfare by increasing transfer rates, while those focusing on economic growth can decrease transfer rates by a relatively small loss on average welfare.

The present study also has some limitations and shortcomings. Firstly, while the model primarily addresses the mechanism through which inequality influences growth, it's challenging to overlook the reciprocal impact: growth may also positively affect inequality. This is evident as growth and inequality consistently demonstrate a positive relationship in the benchmark model, tendency experiment, and redistribution experiments. Regarding model configuration, notable heterogeneity exists within the top 10% group. Recent data from the 2021 Census by ONS indicates that the wealthiest 1% of households possess over 20% of the wealth within the top 10% decile, potentially highlighting variations in entrepreneurial acumen and old money within this bracket.

Moreover, individual bonds are excluded from the linearized model equations due to the unavailability of individual credit data. Although macro credit conditions are incorporated, this may result in approximations when considering tax transfer. Concerning the model's adequacy, the simulated outcomes for inequality and growth display greater volatility compared to the actual data. Finally, the study relies heavily on proxy variables, which could be enhanced with improved datasets in the future. Given ample micro-data in subsequent research, incorporating individual credit equations into the model could lead to more nuanced behaviours of the real interest rate and better differentiation between individual capital and wealth.

From a policy perspective, beyond income tax, introducing other instruments like asset return tax might be worth exploring. Despite the concerns and areas for future refinement, the model demonstrates a significant long-term trade-off between growth and redistribution aimed at reducing inequality throughout the UK's extensive history.

This thesis presents a new model of growth and inequality, which has been rigorously tested by the method of indirect inference—marking the first application of this technique to the long-term economic history of the UK. The findings highlight critical insights into the trade-offs between economic growth and inequality, offering valuable guidance for future policy formulation.

Bibliography

Acemoglu, D. and Robinson, J. A. (2000), 'Why did the west extend the franchise? democracy, inequality, and growth in historical perspective', *The quarterly journal of economics* **115**(4), 1167–1199. (cited on page 14)

Acemoglu, D. and Robinson, J. A. (2002), 'The political economy of the kuznets curve', *Review of development economics* **6**(2), 183–203. (cited on pages 14 and 16)

Adams Jr, R. H. and Page, J. (2003), 'Poverty, inequality and growth in selected middle east and north africa countries, 1980–2000', *World development* **31**(12), 2027–2048.

Aghion, P., Akcigit, U., Bergeaud, A., Blundell, R. and Hémous, D. (2019), 'Innovation and top income inequality', *The Review of Economic Studies* **86**(1), 1–45. (cited on page 13)

Aghion, P. and Bolton, P. (1997), 'A theory of trickle-down growth and development', *The review of economic studies* **64**(2), 151–172. (cited on page 15)

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. (cited on page 11)

Aiyagari, S. R. (1995), 'Optimal capital income taxation with incomplete markets, borrowing constraints, and constant discounting', *Journal of political Economy* **103**(6), 1158–1175. (cited on page 16)

Alesina, A. and Perotti, R. (1996), 'Income distribution, political instability, and investment', *European economic review* **40**(6), 1203–1228. (cited on page 15)

Alesina, A. and Rodrik, D. (1994), 'Distributive politics and economic growth', *The quarterly journal of economics* **109**(2), 465–490. (cited on pages 12 and 15)

Allen, R. C. (2009), *The British industrial revolution in global perspective*, Cambridge University Press. (cited on page 36)

Andrle, M. (2008), The role of trends and detrending in dsge models, Technical report, University Library of Munich, Germany. (cited on page 136)

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. (cited on page 13)

Atkeson, A. and Irie, M. (2020), Understanding 100 years of the evolution of top wealth shares in the us: What is the role of family firms?, Technical report, National Bureau of Economic Research. (cited on page 13)

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. (cited on page 44)

Bagchi, S. and Svejnar, J. (2015), 'Does wealth inequality matter for growth? the effect of billionaire wealth, income distribution, and poverty', *Journal of Comparative Economics* **43**(3), 505–530. (cited on page 12)

Bandyopadhyay, D. and Basu, P. (2005), 'What drives the cross-country growth and inequality correlation?', *Canadian Journal of Economics/Revue canadienne d'économique* **38**(4), 1272–1297. (cited on page 14)

Bandyopadhyay, D. and Tang, X. (2011), 'Parental nurturing and adverse effects of redistribution', *Journal of Economic Growth* **16**, 71–98. (cited on page 16)

Banerjee, A. V. and Duflo, E. (2003), 'Inequality and growth: What can the data say?', *Journal of economic growth* **8**, 267–299. (cited on pages 18 and 28)

Banerjee, A. V. and Newman, A. F. (1993), 'Occupational choice and the process of development', *Journal of political economy* **101**(2), 274–298. (cited on page 15)

Barro, R. J. (2000), 'Inequality and growth in a panel of countries', *Journal of economic growth* **5**, 5–32. (cited on page 12)

Barro, R. J. (2008), Inequality and growth revisited, Technical report, ADB Working paper series on regional economic integration. (cited on page 12)

Basawa, I. V., Mallik, A. K., McCormick, W. P., Reeves, J. H. and Taylor, R. L. (1991), 'Bootstrapping unstable first-order autoregressive processes', *The annals of Statistics* **19**(2), 1098–1101. (cited on page 60)

Basu, P. and Getachew, Y. (2020), 'Redistributive innovation policy, inequality, and efficiency', *Journal of Public Economic Theory* **22**(3), 532–554. (cited on page 15)

Basu, P. and Guariglia, A. (2007), 'Foreign direct investment, inequality, and growth', *Journal of Macroeconomics* **29**(4), 824–839. (cited on pages 11 and 14)

Baxter, M. and King, R. G. (1999), 'Measuring business cycles: approximate band-pass filters for economic time series', *Review of economics and statistics* **81**(4), 575–593. (cited on page 136)

Benabou, R. (1993), 'Workings of a city: location, education, and production', *The Quarterly Journal of Economics* **108**(3), 619–652. (cited on page 14)

Benabou, R. (1994), 'Human capital, inequality, and growth: A local perspective', *European Economic Review* **38**(3-4), 817–826. (cited on page 14)

Benabou, R. (2000), 'Unequal societies: Income distribution and the social contract', *American Economic Review* **91**(1), 96–129. (cited on page 88)

Benabou, R. (2002), 'Tax and education policy in a heterogeneous-agent economy: What levels of redistribution maximize growth and efficiency?', *Econometrica* **70**(2), 481–517. (cited on page 89)

Berg, A., Ostry, J. D., Tsangarides, C. G. and Yakhshilikov, Y. (2018), 'Redistribution, inequality, and growth: new evidence', *Journal of Economic Growth* **23**, 259–305. (cited on page 12)

Boadway, R., Marchand, M. and Pestieau, P. (1991), 'Optimal linear income taxation in models with occupational choice', *Journal of Public Economics* **46**(2), 133–162. (cited on page 15)

Boar, C. and Midrigan, V. (2022), 'Efficient redistribution', *Journal of Monetary Economics* **131**, 78–91. (cited on pages 16 and 89)

Bourguignon, F. (1981), 'Pareto superiority of unegalitarian equilibria in stiglitz'model of wealth distribution with convex saving function', *Econometrica: Journal of the Econometric Society* pp. 1469–1475. (cited on page 13)

Box, G. E., Jenkins, G. M., Reinsel, G. C. and Ljung, G. M. (2015), *Time series analysis: forecasting and control*, John Wiley & Sons. (cited on page 76)

Brockwell, P. J. and Davis, R. A. (2002), *Introduction to time series and forecasting*, Springer. (cited on page 76)

Cagetti, M. and De Nardi, M. (2006), 'Entrepreneurship, frictions, and wealth', *Journal of political Economy* **114**(5), 835–870. (cited on pages 14 and 18)

Cagetti, M. and De Nardi, M. (2009), 'Estate taxation, entrepreneurship, and wealth', *American Economic Review* **99**(1), 85–111. (cited on pages 15 and 18)

Canova, F. (1998), 'Detrending and business cycle facts', *Journal of monetary economics* **41**(3), 475–512. (cited on page 136)

Canova, F. and Ferroni, F. (2011), 'Multiple filtering devices for the estimation of cyclical dsge models', *Quantitative Economics* **2**(1), 73–98. (cited on page 136)

Carroll, C. D. (1998), 'Why do the rich save so much?'. (cited on page 14)

Castelló, A. and Doménech, R. (2002), 'Human capital inequality and economic growth: some new evidence', *The economic journal* **112**(478), C187–C200. (cited on page 12)

Castelló-Climent, A. (2010), 'Inequality and growth in advanced economies: an empirical investigation', *The Journal of Economic Inequality* **8**, 293–321. (cited on page 12)

Chadha, J. S., Janssen, N. and Nolan, C. (2000), *An examination of UK business cycle fluctuations: 1871-1997*, University of Reading, Department of Economics. (cited on page 34)

Chambers, D. and Krause, A. (2010), 'Is the relationship between inequality and growth affected by physical and human capital accumulation?', *The Journal of Economic Inequality* **8**, 153–172. (cited on page 12)

Chamley, C. (1986), 'Optimal taxation of capital income in general equilibrium with infinite lives', *Econometrica: Journal of the Econometric Society* pp. 607–622. (cited on page 16)

Chatterjee, S. and Turnovsky, S. J. (2012), 'Infrastructure and inequality', *European Economic Review* **56**(8), 1730–1745. (cited on page 15)

De La Croix, D. and Doepke, M. (2003), 'Inequality and growth: why differential fertility matters', *American Economic Review* **93**(4), 1091–1113. (cited on page 14)

De Nardi, M. and Yang, F. (2016), 'Wealth inequality, family background, and estate taxation', *Journal of Monetary Economics* **77**, 130–145. (cited on page 15)

Deaton, A. (2003), 'Health, inequality, and economic development', *Journal of economic literature* **41**(1), 113–158. (cited on page 17)

Deininger, K. and Squire, L. (1998), 'New ways of looking at old issues: inequality and growth', *Journal of development economics* **57**(2), 259–287. (cited on page 16)

Doorn, D. (2006), 'Consequences of hodrick–prescott filtering for parameter estimation in a structural model of inventory behaviour', *Applied Economics* **38**(16), 1863–1875. (cited on page 136)

Ehrhart, C. et al. (2009), 'The effects of inequality on growth: a survey of the theoretical and empirical literature', *ECINEQ WP* **107**. (cited on page 11)

Eicher, T. S. and Garcıa-Penalosa, C. (2001), 'Inequality and growth: the dual role of human capital in development', *Journal of Development Economics* **66**(1), 173–197. (cited on page 14)

Fair, R. C. and Taylor, J. B. (1983), 'Solution and maximum likelihood estimation of dynamic nonlinear rational expectations models', *Econometrica: Journal of the Econometric Society* pp. 1169–1185. (cited on pages 48 and 130)

Fernández-Villaverde, J., Rubio-Ramírez, J. F. and Schorfheide, F. (2016), Solution and estimation methods for dsge models, *in* 'Handbook of macroeconomics', Vol. 2, Elsevier, pp. 527–724. (cited on page 48)

Ferroni, F. (2011), 'Trend agnostic one-step estimation of dsge models', *The BE Journal of Macroeconomics* **11**(1). (cited on page 137)

Foellmi, R. and Baselgia, E. (2022), 'Inequality and growth: A review on a great open debate in economics', *CEPR Discussion Papers* . (cited on page 11)

Foellmi, R. and Oechslin, M. (2010), 'Market imperfections, wealth inequality, and the distribution of trade gains', *Journal of International Economics* **81**(1), 15–25. (cited on page 14)

Foellmi, R. and Oechslin, M. (2020), 'Harmful procompetitive effects of trade in presence of credit market frictions', *Journal of Money, Credit and Banking* **52**(6), 1493–1525. (cited on page 14)

Foellmi, R. and Zweimüller, J. (2006), 'Income distribution and demand-induced innovations', *The Review of Economic Studies* **73**(4), 941–960. (cited on page 17)

Forbes, K. J. (2000), 'A reassessment of the relationship between inequality and growth', *American economic review* **90**(4), 869–887. (cited on page 12)

Galor, O. and Tsiddon, D. (1997), 'Technological progress, mobility, and economic growth', *The American Economic Review* pp. 363–382. (cited on page 17)

Galor, O. and Zang, H. (1997), 'Fertility, income distribution, and economic growth: theory and cross-country evidence', *Japan and the world economy* **9**(2), 197–229. (cited on page 14)

Galor, O. and Zeira, J. (1993), 'Income distribution and macroeconomics', *The review of economic studies* **60**(1), 35–52. (cited on pages 14 and 18)

García-Peñalosa, C. and Turnovsky, S. J. (2007), 'Growth, income inequality, and fiscal policy: What are the relevant trade-offs?', *Journal of Money, Credit and Banking* **39**(2-3), 369–394. (cited on page 15)

García-Peñalosa, C. and Wen, J.-F. (2008), 'Redistribution and entrepreneurship with schumpeterian growth', *Journal of Economic Growth* **13**, 57–80. (cited on page 15)

Getachew, Y. Y. (2010), 'Public capital and distributional dynamics in a two-sector growth model', *Journal of Macroeconomics* **32**(2), 606–616. (cited on page 15)

Getachew, Y. Y. and Turnovsky, S. J. (2015), 'Productive government spending and its consequences for the growth–inequality tradeoff', *Research in Economics* **69**(4), 621–640. (cited on page 16)

Gomez, M. (2023), 'Decomposing the growth of top wealth shares', *Econometrica* **91**(3), 979–1024. (cited on page 13)

Gorodnichenko, Y. and Ng, S. (2010), 'Estimation of dsge models when the data are persistent', *Journal of Monetary Economics* **57**(3), 325–340. (cited on page 137)

Halter, D., Oechslin, M. and Zweimüller, J. (2014), 'Inequality and growth: the neglected time dimension', *Journal of economic growth* **19**, 81–104. (cited on page 17)

Hansen, B. E. (1999), 'The grid bootstrap and the autoregressive model', *Review of Economics and Statistics* **81**(4), 594–607. (cited on page 60)

Hansen, G. D. (1985), 'Indivisible labor and the business cycle', *Journal of monetary Economics* **16**(3), 309–327.

Hopkins, E. and Kornienko, T. (2006), 'Inequality and growth in the presence of competition for status', *Economics Letters* **93**(2), 291–296. (cited on page 16)

Horowitz, J. L. (2001), The bootstrap, *in* 'Handbook of econometrics', Vol. 5, Elsevier, pp. 3159–3228. (cited on page 60)

Judd, K. L. (1985), 'Redistributive taxation in a simple perfect foresight model', *Journal of public Economics* **28**(1), 59–83. (cited on page 16)

Kaldor, N. (1955), 'Alternative theories of distribution', *The review of economic studies* **23**(2), 83–100. (cited on page 13)

King, R. G. and Rebelo, S. T. (1993), 'Low frequency filtering and real business cycles', *Journal of Economic dynamics and Control* **17**(1-2), 207–231. (cited on page 136)

Kolev, G. V. and Niehues, J. (2016), The inequality-growth relationship: An empirical reassessment, Technical report, IW-Report. (cited on page 28)

Le, V. P. M., Meenagh, D., Minford, P. and Wickens, M. (2011), 'How much nominal rigidity is there in the us economy? testing a new keynesian dsge model using indirect inference', *Journal of Economic Dynamics and Control* **35**(12), 2078–2104. (cited on pages 51 and 53)

Le, V. P. M., Meenagh, D., Minford, P. and Wickens, M. (2012), 'Testing dsge models by indirect inference and other methods: some monte carlo experiments'. (cited on page 139)

Le, V. P. M., Meenagh, D., Minford, P., Wickens, M. and Xu, Y. (2016), 'Testing macro models by indirect inference: a survey for users', *Open Economies Review* **27**, 1–38. (cited on pages 63 and 139)

Levine, R. and Rubinstein, Y. (2017), 'Smart and illicit: who becomes an entrepreneur and do they earn more?', *The Quarterly Journal of Economics* **132**(2), 963–1018. (cited on page 38)

Lin, Y.-C., Yeh, C.-C. et al. (2009), 'Joint determinations of inequality and growth', *Economics Letters* **103**(3), 163–166. (cited on page 12)

Lindert, P. H. (2000), 'Three centuries of inequality in britain and america', *Handbook of income distribution* **1**, 167–216. (cited on page 36)

Loury, G. C. (1981), 'Intergenerational transfers and the distribution of earnings', *Econometrica: Journal of the Econometric Society* pp. 843–867. (cited on page 14)

Lucas Jr, R. E. (1988), 'On the mechanics of economic development', *Journal of monetary economics* **22**(1), 3–42. (cited on page 116)

Lundberg, M. and Squire, L. (2003), 'The simultaneous evolution of growth and inequality', *The economic journal* **113**(487), 326–344. (cited on page 12)

Mankiw, N. G., Romer, D. and Weil, D. N. (1992), 'A contribution to the empirics of economic growth', *The quarterly journal of economics* **107**(2), 407–437. (cited on page 44)

Matthews, K. and Marwaha, S. (1979), *Numerical properties of the LITP model*, The University. (cited on page 49)

McCallum, B. T. (1976), 'Rational expectations and the natural rate hypothesis: some consistent estimates', *Econometrica: Journal of the Econometric Society* pp. 43–52. (cited on page 52)

Mdingi, K. and Ho, S.-Y. (2021), 'Literature review on income inequality and economic growth', *MethodsX* **8**, 101402. (cited on page 11)

Meenagh, D., Minford, P. and Wang, J. (2007), 'Growth and relative living standards-testing barriers to riches on post-war panel data'. (cited on page 25)

Meenagh, D., Minford, P. and Wickens, M. (2012), 'Testing macroeconomic models by indirect inference on unfiltered data'. (cited on pages 137 and 138)

Meenagh, D., Minford, P. and Wickens, M. (2021), 'Estimating macro models and the potentially misleading nature of bayesian estimation', *CEPR Discussion Paper No. DP15684* . (cited on page 63)

Meenagh, D., Minford, P., Wickens, M. R. and Xu, Y. (2018), The small sample properties of indirect inference in testing and estimating dsge models, Technical report, Cardiff Economics Working Papers. (cited on page 63)

Minford, L. and Meenagh, D. (2019), 'Testing a model of uk growth: A role for r&d subsidies', *Economic Modelling* **82**, 152–167. (cited on pages 15, 18, and 28)

Minford, L. and Meenagh, D. (2020), 'Supply-side policy and economic growth: A case study of the uk', *Open Economies Review* **31**, 159–193. (cited on pages 18 and 28)

Minford, P., Marwaha, S., Matthews, K. and Sprague, A. (1984), 'The liverpool macroeconomic model of the united kingdom', *Economic Modelling* **1**(1), 24–62. (cited on page 48)

Minford, P., Matthews, K. and Marwaha, S. (1979), 'Terminal conditions as a means of ensuring unique solutions for rational expectations models with forward expectations', *Economics Letters* **4**(2), 117–120. (cited on page 49)

Mirrlees, J. A. (1971), 'An exploration in the theory of optimum income taxation', *The review of economic studies* **38**(2), 175–208. (cited on page 15)

Neves, P. C., Afonso, Ó. and Silva, S. T. (2016), 'A meta-analytic reassessment of the effects of inequality on growth', *World Development* **78**, 386–400. (cited on page 12)

Noh, Y.-H. and Yoo, K. (2008), 'Internet, inequality and growth', *Journal of Policy Modeling* **30**(6), 1005–1016. (cited on page 11)

Okun, A. M. (1975), *Equality and efficiency: The big tradeoff*, Brookings Institution Press. (cited on page 16)

Parente, S. L. and Prescott, E. C. (1994), 'Barriers to technology adoption and development', *Journal of political Economy* **102**(2), 298–321. (cited on pages 14 and 18)

Parente, S. L. and Prescott, E. C. (2002), *Barriers to riches*, MIT press. (cited on pages 14 and 18)

Perotti, R. (1996), 'Growth, income distribution, and democracy: What the data say', *Journal of Economic growth* **1**, 149–187. (cited on page 12)

Persson, T. and Tabellini, G. (1994), 'Does centralization increase the size of government?', *European Economic Review* **38**(3-4), 765–773. (cited on pages 12 and 15)

Piketty, T. (2014), *Capital in the twenty-first century*, Harvard University Press. (cited on page 11)

Piketty, T. and Saez, E. (2006), 'The evolution of top incomes: a historical and international perspective', *American economic review* **96**(2), 200–205. (cited on page 44)

Rawls, J. (2017), A theory of justice, *in* 'Applied ethics', Routledge, pp. 21–29. (cited on page 88)

Rebelo, S. (1991), 'Long-run policy analysis and long-run growth', *Journal of political Economy* **99**(3), 500–521. (cited on page 15)

Restuccia, D. and Rogerson, R. (2008), 'Policy distortions and aggregate productivity with heterogeneous establishments', *Review of Economic dynamics* **11**(4), 707–720. (cited on page 14)

Romer, P. M. (1986), 'Increasing returns and long-run growth', *Journal of political economy* **94**(5), 1002–1037. (cited on page 116)

Saez, E. (2001), 'Using elasticities to derive optimal income tax rates', *The review of economic studies* **68**(1), 205–229. (cited on page 16)

Saez, E. and Zucman, G. (2016), 'Wealth inequality in the united states since 1913: Evidence from capitalized income tax data', *The Quarterly Journal of Economics* **131**(2), 519–578. (cited on pages 13 and 44)

Scholl, N. and Klasen, S. (2019), 'Re-estimating the relationship between inequality and growth', *Oxford Economic Papers* **71**(4), 824–847. (cited on page 13)

Smets, F. and Wouters, R. (2007), 'Shocks and frictions in us business cycles: A bayesian dsge approach', *American economic review* **97**(3), 586–606. (cited on pages 44 and 46)

Solow, R. M. (1956), 'A contribution to the theory of economic growth', *The quarterly journal of economics* **70**(1), 65–94. (cited on pages 44 and 46)

Townsend, R. M. and Ueda, K. (2006), 'Financial deepening, inequality, and growth: a model-based quantitative evaluation', *The Review of Economic Studies* **73**(1), 251–293. (cited on page 11)

Trefethen, L. N. and Bau, D. (2022), *Numerical linear algebra*, SIAM. (cited on page 132)

Wickens, M. R. (1982), 'The efficient estimation of econometric models with rational expectations', *The Review of Economic Studies* **49**(1), 55–67. (cited on page 136)

Yang, X., Minford, P. and Meenagh, D. (2021), 'Inequality and economic growth in the uk', *Open Economies Review* **32**, 37–69. (cited on pages 18, 28, 75, and 95)

Yang, X. and Zhou, P. (2022), 'Wealth inequality and social mobility: A simulation-based modelling approach', *Journal of Economic Behavior & Organization* **196**, 307–329. (cited on page 16)

Appendices

A Modelling details

A.1 Substitution of $1 + r_t$

To prove the equation $1 + r_t = \alpha(1-\tau)\frac{E_t Y_{i,t+1}}{K_{i,t}} + 1 - \delta$, we start from the basics of the Cobb-Douglas production function and derive the expression step-by-step.

1. Cobb-Douglas Production Function:

$$Y_{i,t} = A_t K_{i,t}^{\alpha} L_{i,t}^{1-\alpha} \tag{A1}$$

2. Marginal Product of Capital (MPK):

$$MPK = \frac{\partial Y_{i,t}}{\partial K_{i,t}} = \alpha A_t K_{i,t}^{\alpha-1} L_{i,t}^{1-\alpha} \tag{A2}$$

3. Equate MPK to the Rental Rate of Capital:

$$r_t = \alpha A_t K_{i,t}^{\alpha-1} L_{i,t}^{1-\alpha} - \delta \tag{A3}$$

4. Re-express A_t in Terms of Output $Y_{i,t}$:

$$A_t = \frac{Y_{i,t}}{K_{i,t}^{\alpha} L_{i,t}^{1-\alpha}} \tag{A4}$$

5. Substitute A_t into the MPK Equation:

$$r_t = \alpha \left(\frac{Y_{i,t}}{K_{i,t}^{\alpha} L_{i,t}^{1-\alpha}} \right) K_{i,t}^{\alpha-1} L_{i,t}^{1-\alpha} - \delta \tag{A5}$$

Simplify:

$$r_t = \alpha \frac{Y_{i,t}}{K_{i,t}} - \delta \tag{A6}$$

6. In a dynamic model, firms decide on capital stock $K_{i,t}$ based on the expected future output $E_t Y_{i,t+1}$

$$r_t = \alpha \frac{E_t Y_{i,t+1}}{K_{i,t}} - \delta \tag{A7}$$

7. Rearrange the Terms:

$$1 + r_t = \alpha \frac{E_t Y_{i,t+1}}{K_{i,t}} + 1 - \delta \tag{A8}$$

8. consider the tax the equation can be rewritten as

$$1 + r_t = \alpha(1 - \tau) \frac{E_t Y_{i,t+1}}{K_{i,t}} + 1 - \delta \tag{A9}$$

A.2 Approximating C-Y ratio by a random walk

Step 1: Redefining Income

We start by redefining an individual's net income $\tilde{Y}_{i,t}$:

$$\tilde{Y}_{i,t} = Y_{i,t} + (1 - \delta)K_{i,t-1} - K_{i,t} - \pi_t Z_{i,t}$$

Here, $Y_{i,t}$ is the individual's income, $(1 - \delta)K_{i,t-1}$ represents the depreciated capital from the previous period, $K_{i,t}$ is the capital stock, and $\pi_t Z_{i,t}$ is the cost of investment.

Step 2: Individual Budget Constraint

The individual's budget constraint can be rewritten as:

$$(1 + r_{t-1})b_{i,t} = C_{i,t} + b_{i,t+1} - \tilde{Y}_{i,t}$$

Where:

- $b_{i,t}$ is the amount of bonds held at time t,
- r_{t-1} is the return on these bonds,
- $C_{i,t}$ is the consumption,

- $\tilde{Y}_{i,t}$ is the redefined net income.

Step 3: No-Ponzi Condition The no-Ponzi condition ensures that an individual cannot perpetually finance consumption by issuing new debt. Mathematically, it states that the present value of all future bond holdings must be equal to the current value of the bonds:

$$b_{i,t+1} = E_t \left[\sum_{j=1}^{\infty} \frac{C_{i,t+j} - \tilde{Y}_{i,t+j}}{\prod_{s=1}^{j}(1 + r_{t+s-1})} \right] \tag{A10}$$

$$(1 + r_{t-1}) b_{i,t} = C_{i,t} - \tilde{Y}_{i,t} + E_t \left[\sum_{j=1}^{\infty} \frac{C_{i,t+j} - \tilde{Y}_{i,t+j}}{\prod_{s=1}^{j}(1 + r_{t+s-1})} \right]$$

The no-Ponzi condition prevents infinite borrowing by ensuring that the present value of future borrowing (or bond issuance) does not exceed the current value of the bonds.

Step 4: Rewrite the Budget Constraint

The Euler equation 3.7 describes the intertemporal consumption choice:

$$C_{i,t}^{\Psi_1} = \frac{1}{\beta} E_t \left[\frac{C_{i,t+1}^{\Psi_1}}{1 + r_t} \right] \approx \frac{1}{\beta^j} E_t \left[\frac{C_{i,t+j}^{\Psi_1}}{\prod_{s=1}^{j}(1 + r_{t+s-1})} \right]$$

For simplicity, assume $\Psi_1 = 1$:

$$C_{i,t} = E_t \left[\frac{C_{i,t+j}}{\beta^j \prod_{s=1}^{j}(1 + r_{t+s-1})} \right]$$

Substitute the Euler equation into the budget constraint:

$$C_{i,t} = (1 + r_{t-1})\, b_{i,t} + \widetilde{Y}_{i,t} - E_t \sum_{j=1}^{\infty} \frac{C_{i,t+j} - \tilde{Y}_{i,t+j}}{\Pi_{s=1}^{j}\,(1 + r_{t+s-1})} \tag{A11}$$

$$C_{i,t} = (1 + r_{t-1})\, b_{i,t} + \widetilde{Y}_{i,t} + E_t \sum_{j=1}^{\infty} \frac{\tilde{Y}_{i,t+j}}{\Pi_{s=1}^{j}\,(1 + r_{t+s-1})} - E_t \sum_{j=1}^{\infty} \frac{C_{i,t+j}}{\Pi_{s=1}^{j}\,(1 + r_{t+s-1})} \tag{A12}$$

$$C_{i,t} + E_t \sum_{j=1}^{\infty} \frac{C_{i,t+j}}{\Pi_{s=1}^{j} (1 + r_{t+s-1})} = (1 + r_{t-1})\, b_{i,t} + \widetilde{Y}_{i,t} + E_t \sum_{j=1}^{\infty} \frac{\tilde{Y}_{i,t+j}}{\Pi_{s=1}^{j} (1 + r_{t+s-1})} \tag{A13}$$

Consider the present value of future consumption:

$$\frac{C_{i,t}}{1-\beta} = (1 + r_{t-1})\, b_{i,t} + \widetilde{Y}_{i,t} + E_t \sum_{j=1}^{\infty} \frac{\tilde{Y}_{i,t+j}}{\Pi_{s=1}^{j} (1 + r_{t+s-1})} \tag{A14}$$

$$C_{i,t} = (1-\beta) \left[(1 + r_{t-1})\, b_{i,t} + \widetilde{Y}_{i,t} + E_t \sum_{j=1}^{\infty} \frac{\tilde{Y}_{i,t+j}}{\Pi_{s=1}^{j} (1 + r_{t+s-1})} \right] \tag{A15}$$

The term inside the braces is the household's spendable wealth hence the whole RHS expression is permanent net income or

$$C_{i,t} = (1-\beta)\,(1 + r_{t-1})\, b_{i,t} + \overbrace{Y_{i,t}} \tag{A16}$$

$$\overbrace{Y_{i,t}} = \widetilde{Y}_{i,t} + E_t \sum_{j=1}^{\infty} \frac{\tilde{Y}_{i,t+j}}{\Pi_{s=1}^{j} (1 + r_{t+s-1})} \tag{A17}$$

In steady state (at T) we have

$$C_{i,T} = (1-\beta)\,(1 + r^{*})\, b_{i,T} + \overbrace{Y_{i,T}} \tag{A18}$$

Divide through by $Y_{i,t}$

$$\frac{C_{i,t}}{Y_{i,t}} = (1-\beta)\,(1 + r_{t-1})\, \frac{b_{i,t}}{Y_{i,t}} + \frac{\overbrace{Y_{i,t}}}{Y_{i,t}} \tag{A19}$$

Step 5: Approximating $b_{i,t}/Y_{i,t}$

In the steady state, we assume bond holdings evolve according to:

$$(1 + r_{T-1}) b_{i,T} = b_{i,T+1} \tag{A20}$$

Before reaching the steady state, bonds typically follow an AR process:

$$b_{i,t+1} = (1 + r_{t-1})b_{i,t} + x_{i,t} \tag{A21}$$

The ratio $b_{i,t}/Y_{i,t}$ evolves as:

$$\frac{b_{i,t+1}}{Y_{i,t+1}} \frac{Y_{i,t+1}}{Y_{i,t}} = (1 + r_{t-1})\frac{b_{i,t}}{Y_{i,t}} + \frac{x_{i,t}}{Y_{i,t}}$$

Given the relationship in equation A19, we see that $\frac{C_{i,t}}{Y_{i,t}}$ depends on $(1 + r_{t-1})\frac{b_{i,t}}{Y_{i,t}}$ and the expected present value of future income streams. Since $\frac{b_{i,t}}{Y_{i,t}}$ behaves like a random walk because the random growth rate $Y_{i,t+1}/Y_{i,t}$ is generally close to $1 + r_{t-1}$,

$$\frac{b_{i,t+1}}{Y_{i,t+1}} = \frac{b_{i,t}}{Y_{i,t}} + \frac{x_{i,t}}{Y_{i,t}}$$

Therefore, we can conclude that $C_{i,t}/Y_{i,t}$ can be approximated by a random walk.

A.3 The balanced growth path

Balanced growth process (BGP) refers to a situation where all key economic variables (e.g., output, capital, and technology) grow at constant rates over time, leading to a stable long-run equilibrium. Except for some stationary variables, such as the real interest rate and individual labour inputs, most series in our model are non-stationary.

In our model, we use the estimated coefficient of the deterministic trend of aggregate productivity as the basis for the BGP. This coefficient represents a steady-state time trend, not the true BGP growth rate. To calculate the true BGP growth rate for actual productivity, we start from the steady-state process for aggregate productivity which is described by the equation:

$$\Delta A = \beta_0^A + \beta_1^A \Delta A$$

Thus, the BGP growth rate for aggregate productivity is:

$$\Delta A = \frac{\beta_0^A}{1 - \beta_1^A}$$

The BGP growth rate of individual productivity is therefore $\beta_0^A/(1 - \beta_1^A)$, which is the same as the aggregate productivity growth rate.

To calculate the BGP growth rates for output, capital, and consumption, we start from the implicit aggregate production function in the steady state:

$$Y = AK^{\alpha}N^{(1-\alpha)}$$

where Y (output) and K (capital) grow at the same BGP growth rate, and N (labour) has a zero growth rate.

Hence, the BGP growth rate for both output and capital is:

$$\frac{\beta_0^A}{(1 - \beta_1^A)(1 - \alpha)}$$

This characterization of the balanced growth rates ensures that all key economic variables are growing at a consistent rate, maintaining the equilibrium necessary for a stable long-term growth trajectory. This approach is consistent with the endogenous growth model's requirement that growth is driven by factors within the model, such as productivity improvements and capital accumulation, as described in Romer (1986) and Lucas Jr (1988). By specifying these relationships and growth rates, we ensure that the model accurately reflects the endogenous nature of long-term economic growth and adheres to the theoretical foundations of balanced growth processes.

A.4 Derivations of the log linearized model

Linearization of the individual Euler euqation 3.7,

$$(v_{i,t}C_{i,t})^{-\Psi_1} = (1+r_t)\,\beta E_t\left[(v_{i,t+1}C_{i,t+1})^{-\Psi_1}\right]$$

1. Take the natural logarithm of both sides:

$$\ln\left((v_{i,t}C_{i,t})^{-\Psi_1}\right) = \ln\left((1+r_t)\,\beta E_t\left[(v_{i,t+1}C_{i,t+1})^{-\Psi_1}\right]\right) \tag{A22}$$

2. Simplify using logarithm properties:

$$-\Psi_1 \ln(v_{i,t}C_{i,t}) = \ln(1+r_t) + \ln\beta + \ln E_t\left[e^{-\Psi_1(\ln v_{i,t+1}+\ln C_{i,t+1})}\right] \tag{A23}$$

3. Distribute the logarithm inside the expectation:

$$-\Psi_1(\ln v_{i,t} + \ln C_{i,t}) = \ln(1+r_t) + \ln\beta + E_t\left[-\Psi_1(\ln v_{i,t+1} + \ln C_{i,t+1})\right] \tag{A24}$$

4. Simplify and isolate the terms:

$$-\Psi_1 \ln v_{i,t} - \Psi_1 \ln C_{i,t} = \ln(1+r_t) + \ln\beta - \Psi_1 E_t[\ln v_{i,t+1}] - \Psi_1 E_t[\ln C_{i,t+1}] \tag{A25}$$

5. Approximate $\ln(1+r_t) \approx r_t$ since r_t is typically small.

$$-\Psi_1 \ln v_{i,t} - \Psi_1 \ln C_{i,t} = r_t + \ln\beta - \Psi_1 E_t[\ln v_{i,t+1}] - \Psi_1 E_t[\ln C_{i,t+1}] \tag{A26}$$

Thus, the log-linearized Euler equation without the idiosyncratic shocks to consumption is:

$$r_t = \Psi_1(E_t[\ln C_{i,t+1}]) - \ln C_{i,t}) - \ln\beta \tag{A27}$$

Linearization of individual capital equation 3.8,

Rewrite euqation A9 as

$$K_{i,t} = \alpha(1-\tau)\frac{E_t Y_{i,t+1}}{[\delta + r_t]} \tag{A28}$$

1. Take the natural logarithm of both sides:

$$\ln K_{i,t} = \ln\left(\alpha(1-\tau)\frac{E_t Y_{i,t+1}}{\delta + r_t}\right) \tag{A29}$$

2. Log-linearize the expected output $E_t Y_{i,t+1}$ using

$$E_t\left[\frac{Y_{i,t+1}}{C_{i,t+1}}\right] \approx \frac{Y_{i,t}}{C_{i,t}} \tag{A30}$$

$$E_t \ln Y_{i,t+1} \approx \ln Y_{i,t} + E_t \ln C_{i,t+1} - \ln C_{i,t} \tag{A31}$$

3. Substitute the expected output into the equation and linearized it as :

$$\ln K_{i,t} \approx [\alpha(1-\tau)(Y_i/K_i)\left(\ln Y_{i,t} + E_t \ln C_{i,t+1} - \ln C_{i,t}\right) - r_t] \, / (\delta + r) \tag{A32}$$

4. Substitute the Euler equation A27 into the equation :

$$\ln K_{i,t} \approx [\alpha(1-\tau)(Y_i/K_i)\left(\ln Y_{i,t} + (r_t + \ln \beta)/\Psi_1\right) - r_t] \, / (\delta + r) \tag{A33}$$

5. Regardless of some constant terms

$$\ln K_{i,t} = \ln Y_{i,t} - \left[\frac{K}{(1-\tau)\alpha Y} - \frac{1}{\Psi_1}\right] r_t \tag{A34}$$

which gives individual capital euqation A47 and euqation A53.

Linearization individual labour equation 3.9,

$$(1-\Phi)\left(1 - u_{i,t}N_{i,t} - Z_{i,t}\right)^{-\Psi_2} = \Phi\left(v_{i,t}C_{i,t}\right)^{-\Psi_1}(1-\tau)(1-\alpha)\frac{Y_{i,t}}{N_{i,t}}$$

1. Take the natural logarithm of both sides:

$$\ln\left[(1-\Phi)\left(1 - u_{i,t}N_{i,t} - Z_{i,t}\right)^{-\Psi_2}\right] = \ln\left[\Phi\left(v_{i,t}C_{i,t}\right)^{-\Psi_1}(1-\tau)(1-\alpha)\frac{Y_{i,t}}{N_{i,t}}\right] \tag{A35}$$

2. Use the properties of logarithms to separate the terms

$$\ln(1-\Phi) - \Psi_2 \ln\left(1 - u_{i,t}N_{i,t} - Z_{i,t}\right) = \ln(\Phi) - \Psi_1 \ln(v_{i,t}) - \Psi_1 \ln(C_{i,t}) + \ln(1-\tau) + \ln(1-\alpha) + \ln(Y_{i,t}) - \ln(N_{i,t}) \tag{A36}$$

3. Log-linearize around the steady state:

$$\ln(1 - (u_{i,t}N_{i,t} + Z_{i,t})) \approx -(u_{i,t}N_{i,t} + Z_{i,t}) = -(\bar{u}\bar{N} + \bar{Z} + \tilde{u}_{i,t}\bar{N} + \bar{u}\tilde{N}_{i,t} + \tilde{Z}_{i,t}) \tag{A37}$$

4. Combining and simplifying, we obtain:

$$\Psi_2(\tilde{N}_{i,t} + \bar{N}\tilde{u}_{i,t} + \tilde{Z}_{i,t}) = -\Psi_1(\tilde{v}_{i,t} + \tilde{C}_{i,t}) + \tilde{Y}_{i,t} - \tilde{N}_{i,t} \tag{A38}$$

5. Rearrange it as

$$\tilde{N}_{i,t} = \frac{1}{\Psi_2 + 1}\left[\tilde{Y}_{i,t} - \Psi_1\tilde{C}_{i,t} - \Psi_2\tilde{Z}_{i,t} - \Psi_2\bar{N}\tilde{u}_{i,t} - \Psi_1\tilde{v}_{i,t}\right] \tag{A39}$$

6. Substitute $Z_{i,t}$ out using equation 3.6 and equation A41.

$$\ln N_{i,t} = \frac{1}{1+\Psi_2}\left(\ln Y_{i,t} - \Psi_1 \ln C_{i,t} + \frac{2\Psi_2\phi_{2,i}}{\theta_2}\pi'_{i,t}\right) \tag{A40}$$

which gives individual capital euqation A49 and euqation A55. Individual bonds are removed from the equation list because they take a small share of individual capital resources which we are not interested in.

Linearization of the individual entrepreneurship time equation 3.14,

$$\frac{A_{i,t+1}}{A_{i,t}} = \frac{\beta\theta_2(1-\tau)Y_{i,t}/w_{i,t}}{(1-\beta)(1-\tau+\pi'_{i,t})}$$

Taking a first-order Taylor expansion of the right-hand side around a point where $Y_{i,t}/w_{i,t} = Y_i/w_i$ and $\pi_{i,t} = \pi_i$ gives

$$\frac{A_{i,t+1}}{A_{i,t}} = \frac{\beta\theta_2(1-\tau)Y_i/w_i}{(1-\beta)(1-\tau+\pi'_i)} + \frac{\beta\theta_2(1-\tau)}{(1-\beta)(1-\tau+\pi'_{i,t})}d\frac{Y_{i,t}}{w_{i,t}} - \frac{\beta\theta_2(1-\tau)Y_i/w_i}{(1-\beta)(1-\tau+\pi'_i)^2}d\pi'_{i,t}$$

Treating the ratio Y/w as roughly time invariant - on the basis that $w_{i,t}/Y_{i,t} = \alpha N_{i,t}$ and labour is a long-run stationary variable and modelling the penalty rate as stationary, a linear relationship exists between $\frac{A_{i,t+1}}{A_{i,t}}$ and $\pi'_{i,t}$ of the form

$$\ln A_{i,t+1} - \ln A_{i,t} = \phi_{1,i} - \phi_{2,i} \ln \pi'_{i,t} \tag{A41}$$

where $\phi_{1,i} = \frac{\beta\theta_2(1-\tau)Y_i/w_i}{(1-\beta)(1-\tau+\pi'_i)}$, $\phi_{2,i} = \frac{\beta\theta_2(1-\tau)Y_i/w_i}{(1-\beta)(1-\tau+\pi'_i)^2}$.

The full log-lineaized model is:

$$r_t = \Psi_1(E_t[\ln C_{2,t+1}]) - \ln C_{2,t}) - \ln \beta \tag{A42}$$

$$\ln Y_t = \ln [\mu_1 \exp(\ln Y_{1,t}) + \mu_2 \exp(\ln Y_{2,t})] + \epsilon_t^Y \tag{A43}$$

$$\ln K_t = \ln [\mu_1 \exp(\ln K_{1,t}) + \mu_2 \exp(\ln K_{2,t})] + \epsilon_t^K \tag{A44}$$

$$\ln C_t = (1-\tau)\frac{Y}{C} \ln Y_t - \frac{K}{C} [\ln K_t - (1-\delta) \ln K_{t-1}] + \epsilon_t^M \tag{A45}$$

$$\ln Y_{1,t} = \alpha \ln K_{1,t-1} + (1-\alpha) \ln N_{1,t} + \ln A_{1,t} \tag{A46}$$

$$\ln K_{1,t} = \ln Y_{1,t} - \left[\frac{K}{(1-\tau)\alpha Y} - \frac{1}{\Psi_1}\right] r_t \tag{A47}$$

$$\ln C_{1,t} = E_t[\ln C_{1,t+1}] - \frac{r_t + \ln \beta}{\Psi_1} + \epsilon_t^{C1} \tag{A48}$$

$$\ln N_{1,t} = \frac{1}{(1+\Psi_2)} \left(\ln Y_{1,t} - \Psi_1 \ln C_{1,t} + 2\frac{\Psi_2 \phi_{2,1}}{\theta_2} \ln \pi'_{1,t}\right) + \epsilon_t^{N1} \tag{A49}$$

$$\ln A_{1,t+1} = \ln A_{1,t} + \phi_{1,1} - \phi_{2,1} \ln \pi'_{1,t} + \epsilon_t^A \tag{A50}$$

$$\ln \pi'_{1,t} = \rho_0^\pi + \rho_1^\pi \ln \pi'_{1,t-1} - \rho_2^\pi \frac{\mu_1}{\omega_{Y1}} \left(\frac{K_{1,t-2}}{K_{t-2}}\right)^2 + \rho_3^\pi \text{Cre}_t + \epsilon_t^\pi \tag{A51}$$

$$\ln Y_{2,t} = \alpha \ln K_{2,t-1} + (1-\alpha) \ln N_{2,t} + \ln A_{2,t} \tag{A52}$$

$$\ln K_{2,t} = \ln Y_{2,t} - \left[\frac{K}{(1-\tau)\alpha Y} - \frac{1}{\Psi_1}\right] r_t \tag{A53}$$

$$\ln C_{2,t} = \frac{1}{\omega_{C2}} (\ln C_t - \omega_{C1} \ln C_{1,t}) + \epsilon_t^{C2} \tag{A54}$$

$$\ln N_{2,t} = \frac{1}{(1+\Psi_2)} \left(\ln Y_{2,t} - \Psi_1 \ln C_{2,t} + 2\frac{\Psi_2 \phi_{2,1}}{\theta_2} \ln \pi'_{2,t}\right) + \epsilon_t^{N2} \tag{A55}$$

$$\ln A_{2,t+1} = \ln A_{2,t} + \phi_{1,2} - \phi_{2,2} \ln \pi'_{2,t} + \epsilon_t^A \tag{A56}$$

$$\ln \pi'_{2,t} = \rho_0^\pi + \rho_1^\pi \ln \pi'_{2,t-1} - \rho_2^\pi \frac{\mu_2}{\omega_{Y2}} \left(\frac{K_{2,t-2}}{K_{t-2}}\right)^2 + \ln \rho_3^\pi \text{Cre}_t + \epsilon_t^\pi \tag{A57}$$

B Supplemental results

B.1 Impulse response functions

Table B1. Summary of Shocks, Standard Deviations, and Corresponding IRFs

Shocks	SD	Figures
Aggregate productivity shock	0.020445628852	Figure B1
(Negative) Aggregate penalty rate shock	0.048787239982	Figure 6.3
Aggregate output shock	0.001653150266	Figure B2
Aggregate capital shock	0.011236556192	Figure B3
Aggregate consumption shock	0.092490529822	Figure B4
Consumption shock of G1	0.146793070014	Figure B5
Consumption shock of G2	0.034221186146	Figure 6.4
Labour supply shock of G1	0.041421012817	Figure B6
Labour supply shock of G2	0.031818610053	Figure B7
Credit shock	0.070889134136	Figure B8

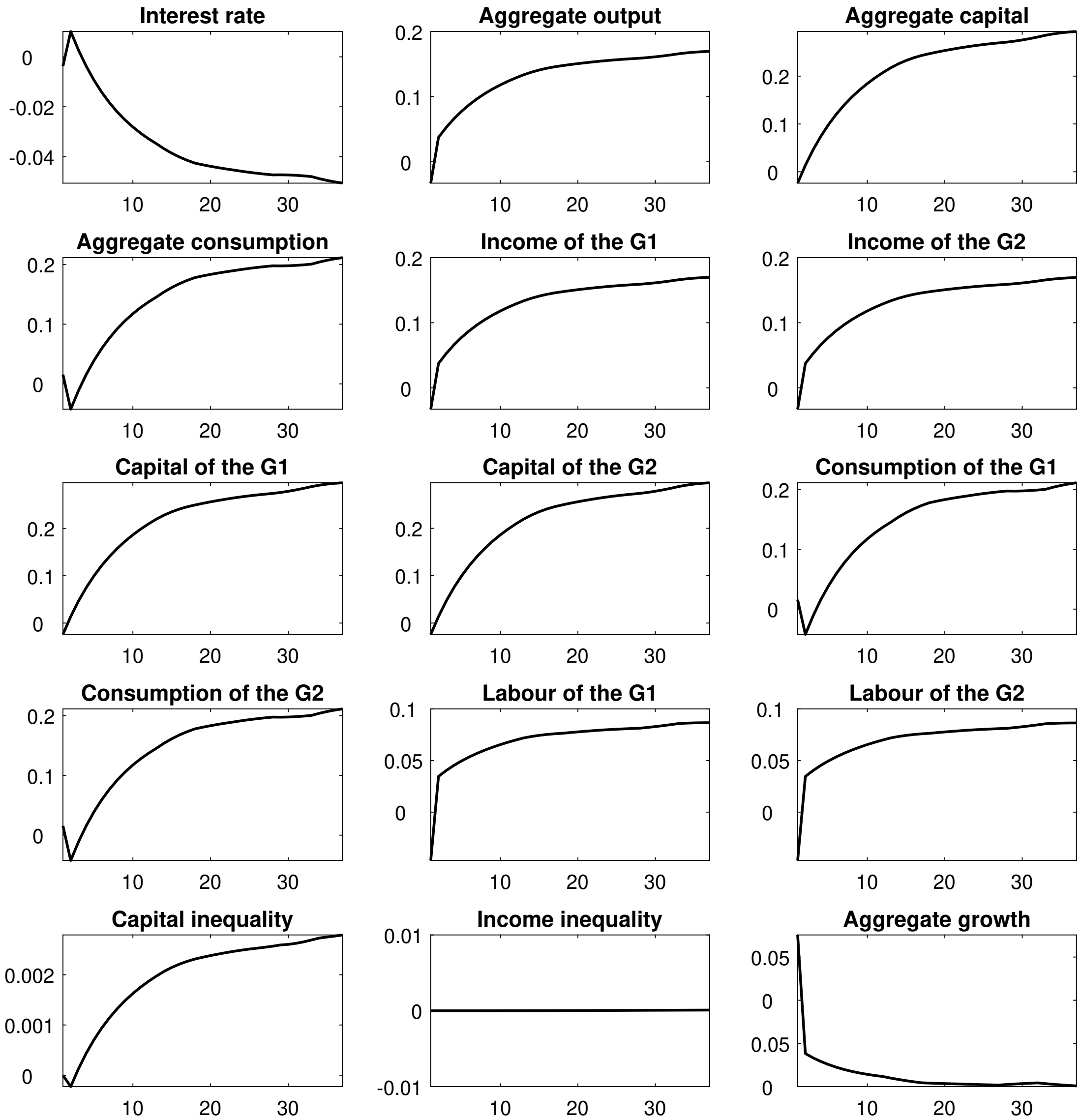


Figure B1. Impulse responses to negative one S.D. of aggregate productivity shock

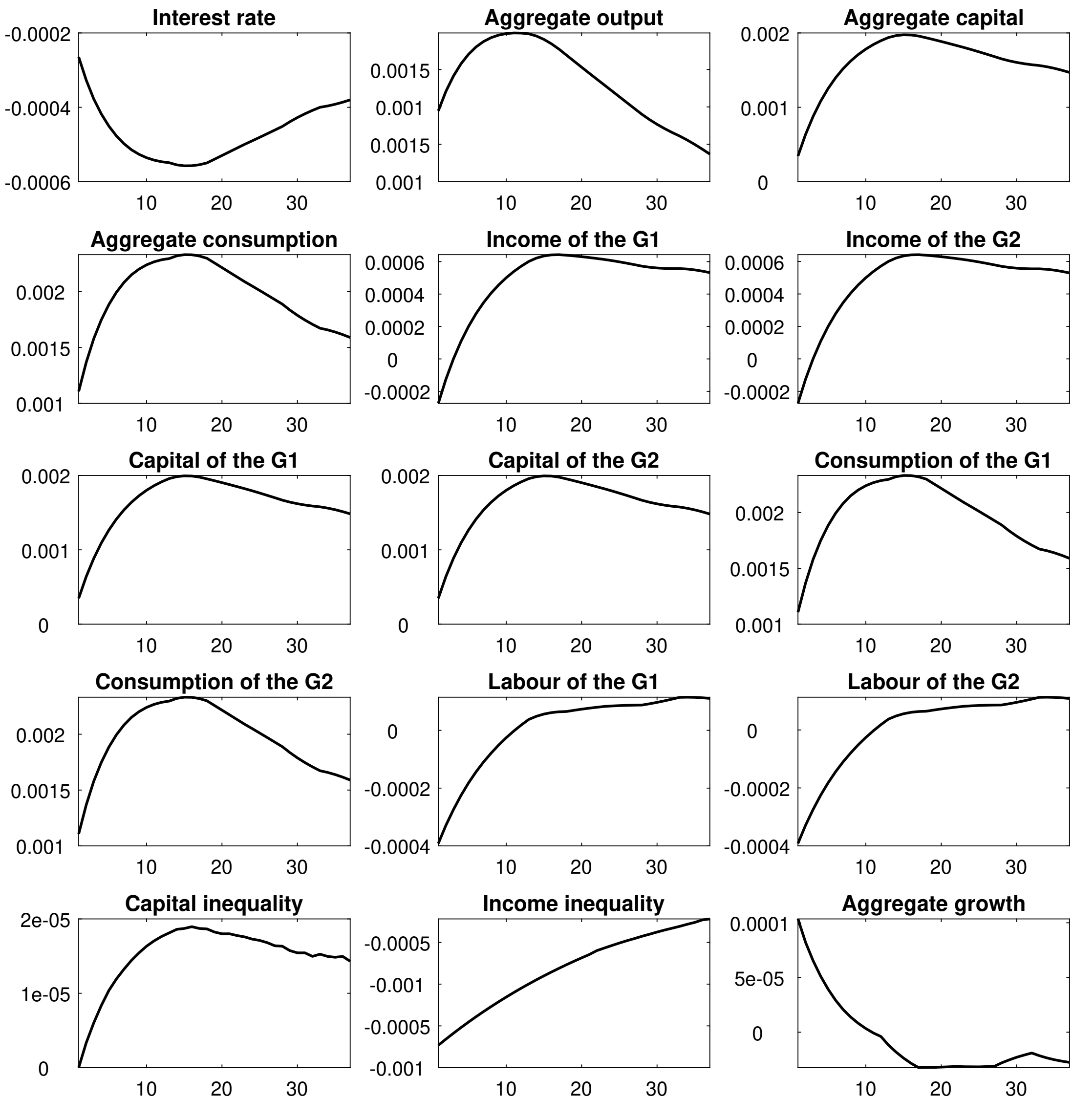


Figure B2. Impulse responses to negative one S.D. of aggregate output shock

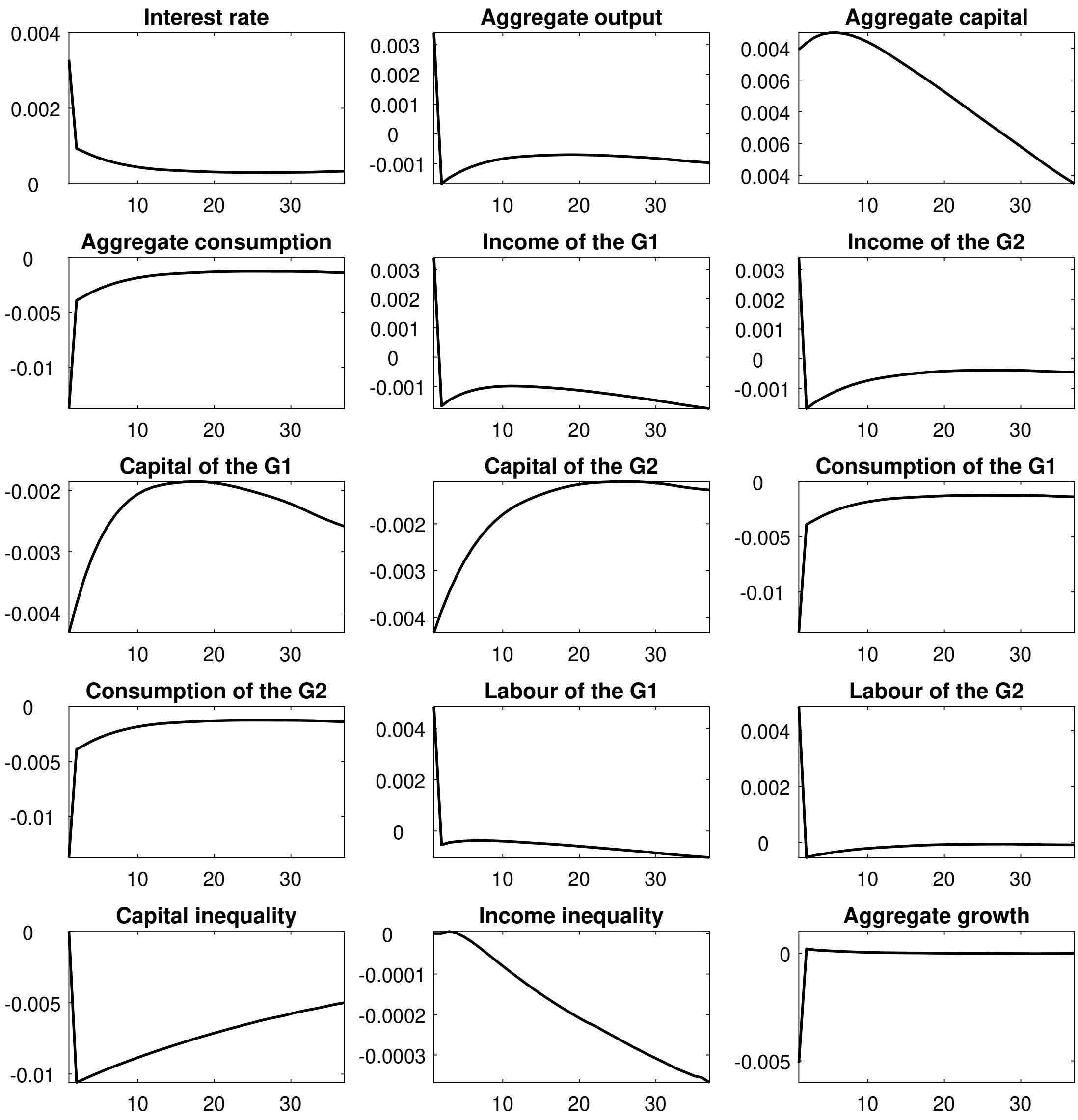


Figure B3. Impulse responses to negative one S.D. of aggregate capital shock

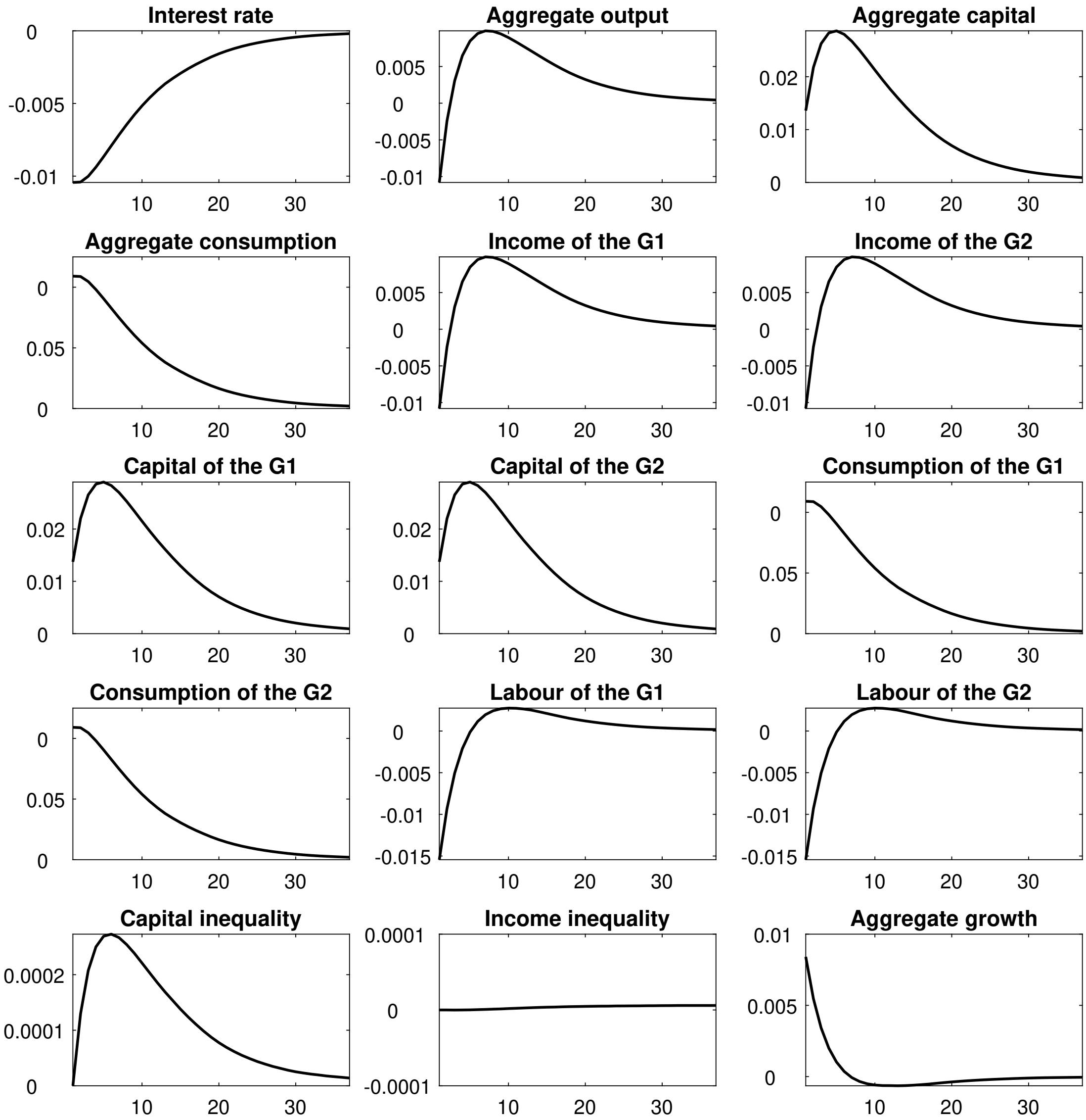


Figure B4. Impulse responses to negative one S.D. of aggregate consumption shock

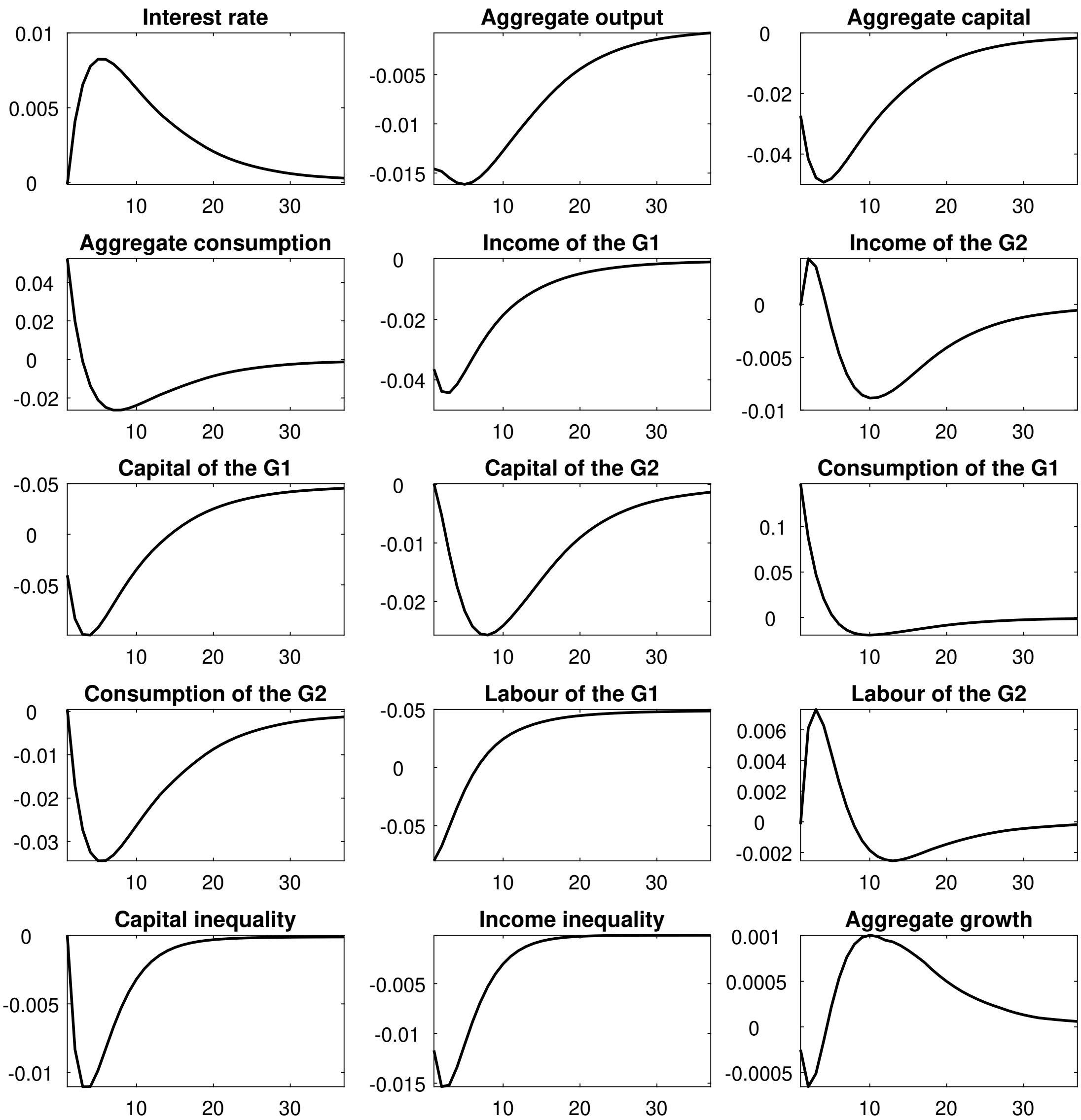


Figure B5. Impulse responses to negative one S.D. of G1's consumption shock

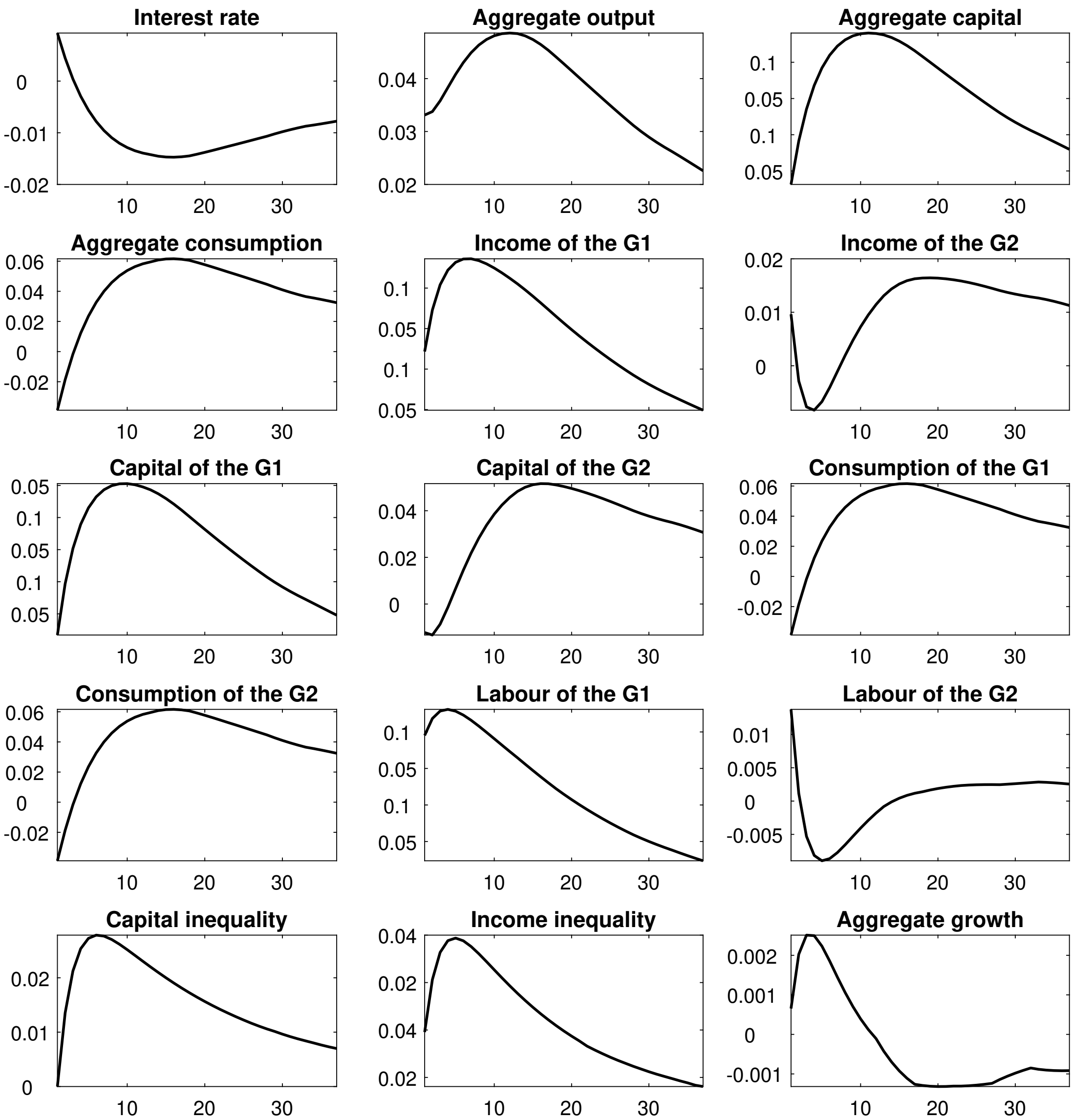


Figure B6. Impulse responses to negative one S.D. of G1's labour shock

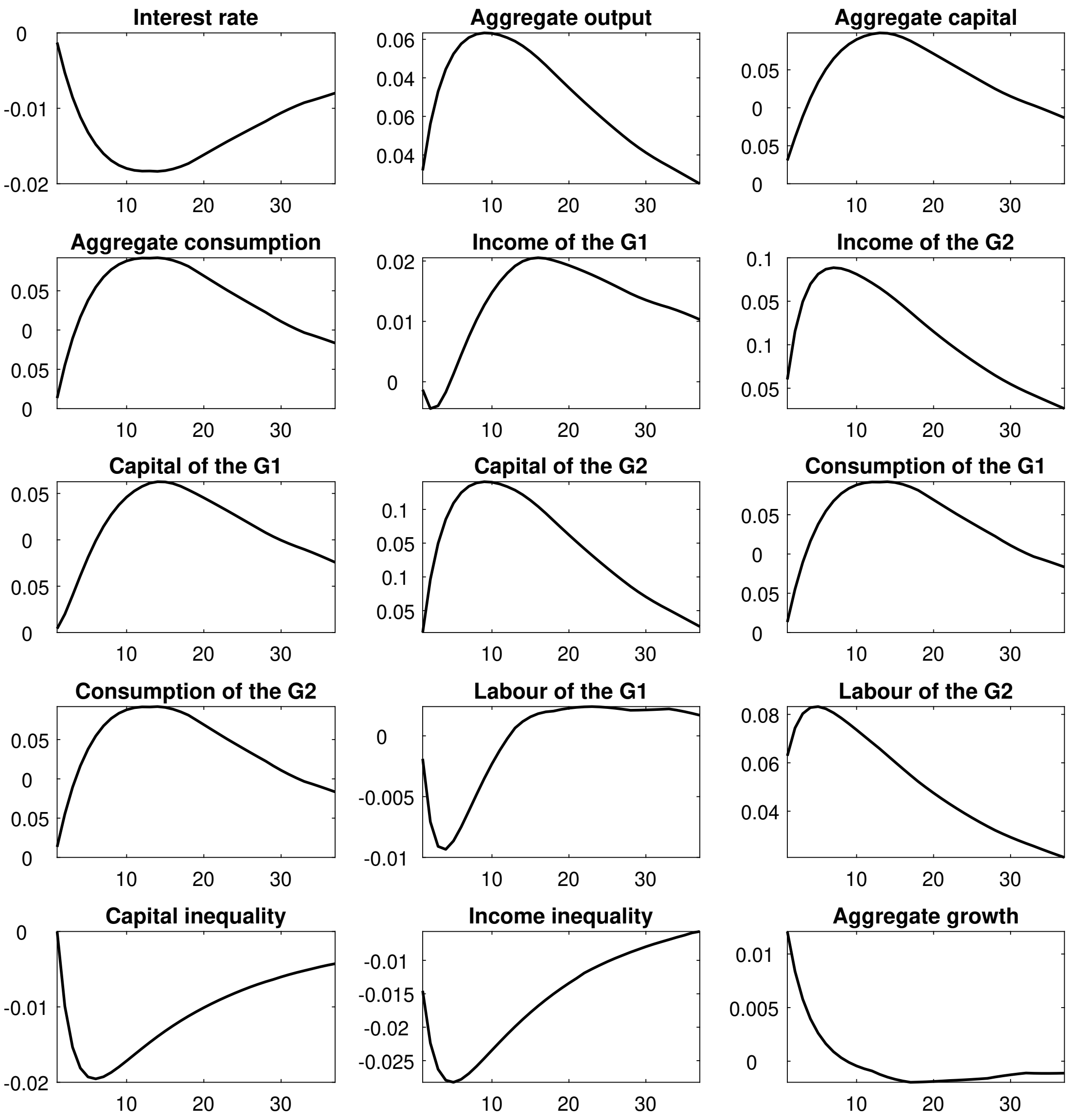


Figure B7. Impulse responses to negative one S.D. of G2's labour shock

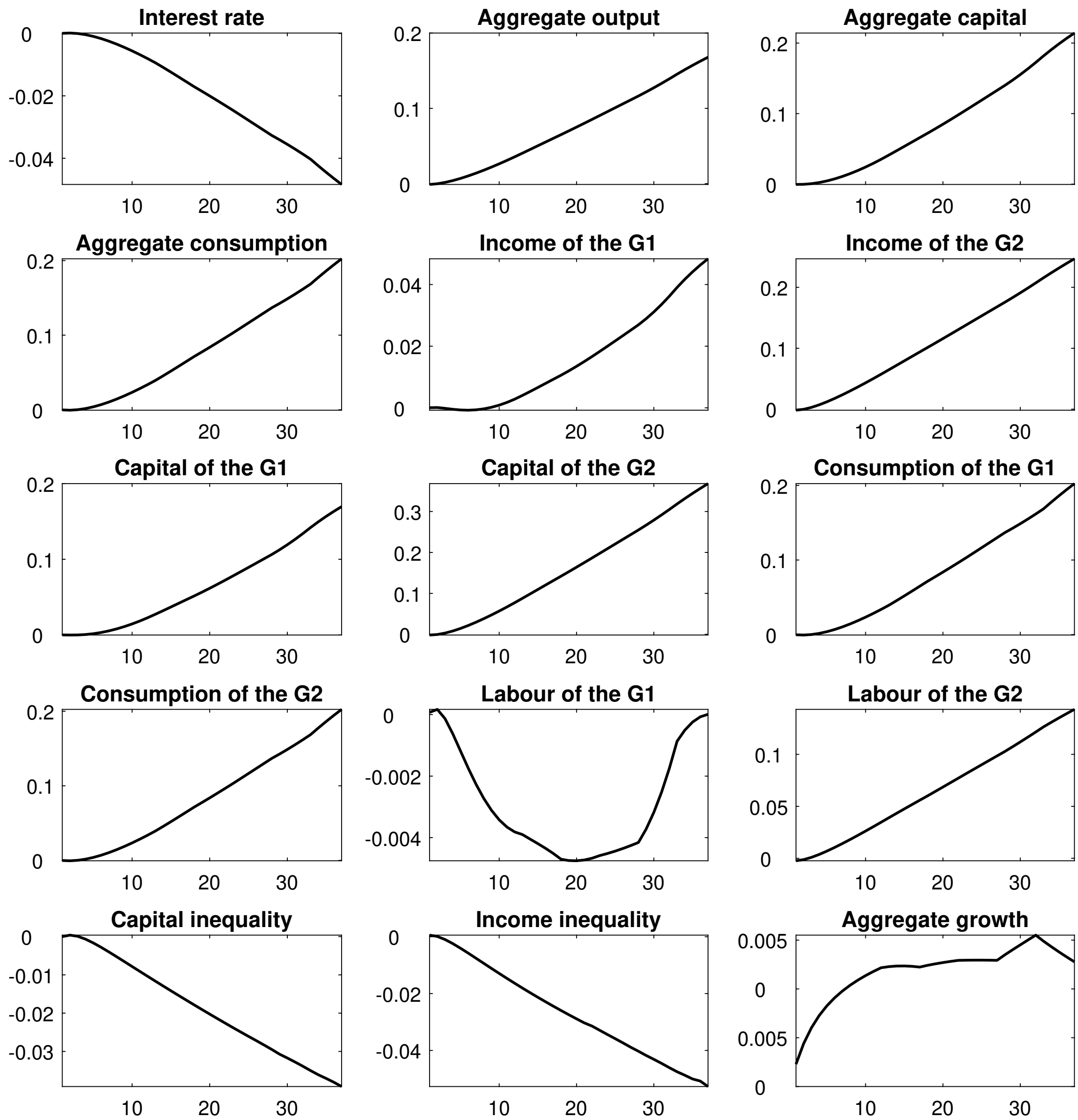


Figure B8. Impulse responses to negative one S.D. of credit shock

C Methodology details

C.1 The Type II iteration

Fair and Taylor (1983) proposed a basic projection method for solving the DSGE model. Considering the dynamic rational expectations model given by

$$f_i\left(y_t, y_{t-1}, \ldots, y_{t-p}, E_{t-1}y_t, E_{t-1}y_{t+1}, \ldots, E_{t-1}y_{t+h}, x_t, \alpha_i\right) = u_{it} \quad (i = 1, \ldots, n) \quad \text{(C1)}$$

where y_t is an n-dimensional vector of endogenous variables at time t, x_t is a vector of exogenous variables at time t, E_{t-1} is the conditional expectations operator based on the model and on information through period $t-1$, α_i is a vector of parameters, and u_{it} is a stationary scalar random variable which has mean zero and which may be correlated across equations ($E[u_{it}u_{jt}] \neq 0$ for $i \neq j$) and over time ($E[u_{it}u_{is}] \neq 0$ for $t \neq s$). The model is *nonlinear* in that the function f_i may be nonlinear in the variables, parameters, and expectations.

If one were given numerical values for the expected endogenous variables in the model (C1) for all periods from s on, then it would be straightforward to solve the model for period s using the Gauss-Seidel iterative technique. The numerical method described here entails a series of iterations that converge from an arbitrary initial path of values for these expectations to a path of *rational expectations*, consistent with the forecasts of the model itself. Let the initial set of values for the expected endogenous variables, $E_{s-1}y_{s+r}$, be represented as g_r, $r = 0, 1, \ldots$. Since in general the model will have no natural termination date, an infinite number of these values need to be specified in principle. In practice, however, only a finite number of these will be used in obtaining a solution with a given finite tolerance range. We require that the initial values be bounded: $|g_r| < M$ for every r, where M is not a function of r.

The solution method can be described in terms of 5 steps:

1. Choose an integer k, which is an initial guess at the number of periods beyond the horizon h for which expectations need to be computed in order to obtain a solution within a prescribed tolerance level δ. Set $E_{s-1}y_{s+r}$ equal to g_r, $r = 0, 1, \ldots, k + 2h$. For the purpose of describing the iterations, call these initial

values $e_r(1, k), r = 0, 1, \ldots, k + 2h$; the values at later iterations will then be called $e_r(i, k), i > 1$.

2. Obtain a new set of values for $E_{s-1}y_{s+r}, r = 0, 1, \ldots, k + h$, by solving the model dynamically for $y_{s+r}, r = 0, 1, \ldots, k + h$. This is done by setting the disturbances to their expected values (usually zero), using the values $E_{s-1}x_s, \ldots, E_{s-1}x_{s+h+k}$ in place of the actual x's, and using the values $e_r(i, k)$ in place of $E_{s-1}y_{s+r}$. Call these new guesses $e_r(i + 1, k), r = 0, 1, \ldots, k + h$. If the model is nonlinear, then the solution for each period requires a series of Gauss-Seidel iterations. Call each of these a Type I iteration.

3. Compute for each expectation variable and each period the absolute value of the difference between the new guess and the previous guess, i.e., compute the absolute value of the difference between each element of the $e_r(i + 1, k)$ vector and the corresponding element of the $e_r(i, k)$ vector for $r = 0, 1, \ldots, h + k$. If any of these differences are not less than a prescribed tolerance level (i.e., if convergence has not been achieved), increase i by 1 and return to step (ii). If convergence has been achieved, go to step (iv). Call this iteration (performing steps (ii) and (iii)) a Type II iteration. Let $e_r(k)$ be the vector of the convergent values of a series of Type II iterations ($r = 0, 1, \ldots, k + h$).

4. Repeat steps (i) through (iii), replacing k by $k + 1$. Compute the absolute value of the difference between each element of the $e_r(k + 1)$ vector and the corresponding element of the $e_r(k)$ vector, $r = 0, 1, \ldots, h$. If any of these differences are not less than δ, increase k by 1 and repeat steps (i) through (iv). If convergence has been achieved, go to step (v). Call this iteration (performing steps (i) through (iv)) a Type III iteration. Let e_r be the vector of the convergent values of a series of Type III iterations ($r = 0, 1, \ldots, h$).

5. Use e_r for $E_{s-1}y_{s+r}, r = 0, 1, \ldots, h$, and the actual values for x_t to solve the model for period s. This gives the desired solution, say $\hat{y}_s$, and concludes the solution method.

C.2 Update rational expectations

The following is a general explanation and comparison of the Jacobb method, Gauss-Seidel, Powell-Hybrid and Levenberg-Marquardt algorithms. See Trefethen and Bau (2022) for more details. Given a system of linear equations represented in matrix form as

$$Ax = b \tag{1}$$

where $\boldsymbol{A}$ is a square matrix of size $n \times n, \boldsymbol{x}$ is the vector of unknowns, $\boldsymbol{b}$ is the known vector. The system (matrix $\boldsymbol{A}$)must be diagonally dominant or the method may not converge.

Jacobb method. The Jacobi method is an iterative algorithm used to solve a system of linear equations. It is named after the German mathematician Carl Gustav Jacob Jacobi. The idea behind the Jacobi method is to solve for each unknown x_i on the left-hand side, using the other unknowns from the previous iteration on the righthand side. The algorithm mainly consists of 2 steps:

Step 1: Initialize. Choose an initial $x^{(0)}$ and set a tolerance level and a maximum number of iterations.

Step 2: Iterate. For $k = 1, 2, \ldots$, until convergence or maximum iterations reached:

a. For $i = 1$ to n : compute $\sum_{j \neq i} a_{ij} x_j^{(k-1)}$ and update the unknown: $x_i^{(k)} = \frac{b_i - \sum_{j \neq i} a_{ij} x_j^{(k-1)}}{a_{ii}}$

b. Check convergence: If $\left\| x^{(k)} - x^{(k-1)} \right\| <$ tolerance, then stop.

The vector $x^{(k)}$ is the approximate solution to the system. The method is simple to implement. Each update is independent of the others within the same iteration, making the method easily parallelizable. The method might not converge if the matrix is not diagonally dominant. Convergence can be slow, especially if the matrix $\boldsymbol{A}$ is illconditioned.

Gauss-Seidel Algorithm. It is named after German mathematicians Carl Friedrich Gauss and Philipp Ludwig von Seidel, and is a modification of the Jacobi method. The idea behind the Gauss-Seidel method is to decompose the matrix $\boldsymbol{A}$ into a lower triangular part $\boldsymbol{L}$ and an upper triangular part $\boldsymbol{U}$, so that $\boldsymbol{A} = \boldsymbol{L} + \boldsymbol{U}$. Then the original system can be written as:

$$(L + U)x = b \tag{2}$$

$$Lx = b - Ux \tag{3}$$

This leads to an iterative method where you solve the left side for x given an initial guess, and then use that result as the guess for the next iteration. The Gauss-Seidel method is similar to the Jacobi method but differs in how the updates are performed. In the Jacobi method, all the updates in a given iteration are based on the values from the previous iteration. In the Gauss-Seidel method, updates within the same iteration immediately use the newly computed values, potentially accelerating convergence. However, the method might not converge if the matrix is not diagonally dominant or if it's not positive definite in the symmetric case. Unlike the Jacobi method, it's harder to parallelize as each update depends on previous ones within the same iteration. In practice, the Gauss-Seidel method is useful for solving large sparse systems of linear equations where direct methods might be computationally expensive. It is often used as a smoother in multigrid methods.

Levenberg-Marquardt (LM) Algorithm. The LM algorithm is a popular iterative optimization method used to solve nonlinear least squares problems. It's a combination of the steepest descent method and the Gauss-Newton method, aiming to inherit the stability of the former and the rapid convergence of the latter. A nonlinear least squares problem aims to find the parameters that minimize the sum of squared residuals between observed data and a nonlinear model. Mathematically, it is formulated as:

$$\min_{x} \|F(\boldsymbol{x})\|^2$$

where $F(\boldsymbol{x})$ is a vector of nonlinear functions, and $\boldsymbol{x}$ is the vector of parameters to be estimated. The LM algorithm works by iteratively updating the parameter estimates. Here's an outline of the algorithm:

Step 1: Initialize. Choose an initial guess $x^{(0)}$, set a tolerance level, maximum number of iterations, and initialize λ, a damping factor.

Step 2: Iterate. For $k = 1, 2, \ldots$, until convergence or maximum iterations reached:

a. Evaluate the Jacobian matrix J of F at $x^{(k-1)}$.

b. Form the augmented normal equations: $\left(J^T J + \lambda I\right) \Delta x = -J^T F \left(x^{(k-1)}\right) \backslash.$

c. Solve for Δx.

d. Update $x^{(k)} = x^{(k-1)} + \Delta x$.

e. Check for convergence: If $\|\Delta x\| <$ tolerance, then stop.

f. Adjust λ : If the reduction in error is significant, decrease λ; otherwise, increase λ.

$x^{(k)}$ is the approximate solution. Combines the best properties of steepest descent and Gauss-Newton, making it suitable for a wide range of problems. Convergence is usually faster than using steepest descent alone. The damping factor helps in handling ill-conditioned problems. However, it requires the calculation of the Jacobian matrix, which can be expensive and the damping factor may require careful tuning for optimal performance. The Levenberg-Marquardt algorithm provides a powerful tool for solving nonlinear least squares problems, combining the robustness of gradient descent with the efficiency of second-order methods.

Powell's Hybrid Algorithm. Powell's Hybrid Algorithm is a nonlinear equation-solving method used to find the roots of a system of nonlinear equations. It combines Powell's dogleg method, which itself is a combination of the steepest descent and Newton's methods, with the Levenberg-Marquardt method. The algorithm is often used when you have a system of nonlinear equations and you want to find the values of the variables that make all the equations equal to zero. The method is iterative, meaning that it improves an initial guess through a series of steps until it reaches a solution that is accurate enough. Here's an overview of how the algorithm works:

1. Initialization: Choose an initial guess for the vector of unknowns, x, and set tolerances for convergence.

2. Evaluate the Function and Jacobian: Calculate the value of the functions and the Jacobian matrix at the current guess.

3. Compute the Dogleg Step: Determine the direction in which to move x by forming a dogleg path that combines the steepest descent and Newton's methods.

4. Levenberg-Marquardt Adjustment: If the dogleg step does not lead to a reduction in the value of the functions, modify the step using the Levenberg-Marquardt technique.

5. Update: Update the guess for x using the computed step.

6. Check for Convergence: If the change in x or the value of the functions is less than the specified tolerances, or if the maximum number of iterations has been reached, stop the iterations.

7. Repeat: Return to step 2 and continue the iterations.

Powell's Hybrid Algorithm can be used on problems that are not well-scaled or have poorly conditioned Jacobians. It is often effective in practice for solving systems of nonlinear equations. The algorithm is more complex to understand and implement compared to some simpler methods. Requires the calculation of both the function values and the Jacobian matrix, which may be computationally intensive. There is no guarantee of convergence, particularly if the initial guess is not close to the solution or if the functions are not sufficiently smooth. Powell's Hybrid method can be an excellent choice when you are dealing with a system of nonlinear equations, particularly if other methods have failed to converge. Its combination of multiple techniques can provide a powerful tool for finding solutions. It is implemented in various software packages, making it accessible to practitioners who want to solve nonlinear problems.

While both LM Algorithm and Powell's Hybrid Algorithm are used for solving nonlinear least squares problems, they approach the problem in different ways. Levenberg-Marquardt switches between Gauss-Newton and Gradient Descent based on a damping parameter and tends to be faster near the solution. Powell's Hybrid method avoids second derivative computation, uses a trust region approach, and can be more suitable when the initial guess is farther from the solution. Powell's Hybrid method was used to solve our model in this thesis.

C.3 Why use the non-stationary data and the VARX auxiliary model

According to Wickens (1982), non-stationary data helps differentiate between temporary and permanent shocks. For stationary or trend-stationary processes, variables have short memories, so shocks only temporarily affect them before they return to their steady trend. In contrast, for unit root processes, variables have long memories, meaning shocks have permanent effects, and variables do not return to their previous path after a disturbance. Additionally, permanent shocks can shift endogenous variables sharing the same balanced growth path (BGP) permanently, where the former describes business cycle effects ('cyclical component'), and the latter affects the long-run growth path.

Traditionally, we make data stationary by applying linear (or higher-order polynomial) detrending for deterministic trends or first differencing for stochastic trends. However, these transformations don't effectively isolate fluctuations with the desired periodicity (Canova, 1998). Linear detrending is unsuitable for data with stochastic trends, while first differencing amplifies high-frequency noise. Both methods risk leaving significant influences of permanent shocks in the detrended data.

To address this, researchers often use the Hodrick-Prescott (HP) filter or similar band-pass (BP) filters to decompose economic time series into trend and cyclical components (Baxter and King, 1999; King and Rebelo, 1993). However, the HP filter has two main drawbacks: it can create artificial cycles that don't exist and distort key business cycle characteristics. Its two-sided moving average filter can misalign the timing of data, affecting forward-looking properties and potentially biasing dynamic parameter estimates in DSGE models (Doorn, 2006). Andrle (2008) criticizes that the detrending data in the DSGE model may be unable to explain co-movements of filtered time series because permanent shocks inducing dynamics usually have a large influence on the business cycle and models using detrended data are less likely to capture the true business cycle dynamics. Canova and Ferroni (2011) compares several univariate filtering devices and finds that different approaches yield significantly different estimates of parameters. Approaches that can potentially extract the cyclical component rely on assumptions about trend processes that can cause mismeasurement of cyclical components and bias the estimation of deep parameters. Other criticisms can also be found in Ferroni (2011); Gorodnichenko and Ng (2010).

When the data is non-stationary, the choice of auxiliary model typically should be a Vector Error Correction (VECM) model, also it can be a Vector Auto Regression with an Exogenous variable model (VARX) as the auxiliary model. Here is a brief proof provided by Meenagh et al. (2012). The structural DSGE model after log-linearisation usually can be written as a function:

$$A(L)y_t = B(L)E_t y_{t+1} + C(L)x_t + D(L)e_t \tag{C58}$$

where y_t is a vector of endogenous variables, $E_t y_{t+1}$ is a vector of expected future endogenous variables, x_t is an exogenous variable which is assumed to be driven by

$$\Delta x_t = \alpha(L)\Delta x_{t-1} + d + b(L)z_{t-1} + c(L)\varepsilon_t \tag{C59}$$

The exogenous variables x_t include stationary and non-stationary shocks like productivity shocks. e_t and ε_t are both i.i.d and the means are zero. x_t is nonstationary, y_t is linearly dependent on x_t. Therefore, y_t is also non-stationary. L is the lag operator $Y_{t-s} = L^s Y_t$ and $A(L), B(L)$ etc is a matrix polynomial functions in the lag operator of order h that have roots of the determinantal polynomial lies outside the complex unit circle. The general solution of y_t can be written as

$$y_t = G(L)y_{t-1} + H(L)x_t + f + M(L)e_t + N(L)\varepsilon_t \tag{C60}$$

where f is a vector of constants and polynomial functions in the lag operator have roots outside of the unit circle. Since y_t and x_t are both non-stationary, the solution of the model has p cointegrated relations given by:

$$y_t = [I - G(1)]^{-1}[H(1)x_t + f] = \Pi x_t + g \tag{C61}$$

The matrix Π is a $p * p$ matrix, which has rank $0 \le r < p$, where r is the number of linearly independent cointegrating vectors. $y_t - [\Pi \bar{x}_t + g] = \eta_t$, where η_t is the error correction term. y_t is a function of deviation from the equilibrium in the short run. In the long run, the solution to the model is given by:

$$\bar{y}_t = \Pi\bar{x}_t + g\bar{x}_t = [1-\alpha(1)]^{-1}[dt + c(1)\xi] \tag{C62}$$

$$\xi_t = \sum_{s}^{t-1} \varepsilon_{t-s} \tag{C63}$$

where $\bar{y}_t$ and $\bar{x}_t$ are the long run solution to y_t and x_t respectively. It can be seen that the long run solution of $\bar{x}_t$ can be decomposed into two components: a deterministic trend $\bar{x}_t^d = [1-\alpha(1)]^{-1}dt$ and a stochastic trend $\bar{x}_t^s = [1-\alpha(1)]^{-1}c(1)\xi_t$. There are two components in the endogenous variables: this trend and a VARMA in deviations from it. Meenagh et al. (2012) formulate this as a cointegrated VECM with a mixed moving average error term, w_t.

$$\begin{aligned}\Delta y_t &= -[I-G(1)]\,(y_{t-1} - \Pi x_{t-1}) + P(L)\Delta y_{t-1} + Q(L)\Delta x_t + f + M(L)e_t + N(L)\varepsilon_t \\ &= -[I-G(1)]\,(y_{t-1} - \Pi x_{t-1}) + P(L)\Delta y_{t-1} + Q(L)\Delta x_t + f + w_t \end{aligned} \tag{C64}$$

where

$$w_t = M(L)e_t + N(L)\varepsilon_t \tag{C65}$$

This suggests that the VECM can be approximated by the VARX:

$$\Delta y_t = -K\,(y_{t-1} - \Pi x_{t-1}) + R(L)\Delta y_{t-1} + S(L)\Delta x_t + g + \zeta_t \tag{C66}$$

where ζ_t is an i.i.d with zero mean, since

$$\bar{x}_t = \bar{x}_{t-1} + [1-\alpha(1)]^{-1}\,[d+\varepsilon_t]\,\bar{y}_t = \Pi\bar{x}_t + g \tag{C67}$$

The VECM can also be rewritten as:

$$\Delta y_t = K\,[(y_{t-1} - \bar{y}_{t-1}) - \Pi\,(x_{t-1} - \bar{x}_{t-1})] + R(L)\Delta y_{t-1} + S(L)\Delta x_t + h + \zeta_t \tag{C68}$$

which distinguishes between the effect of the trend component and the temporary deviation of x_t from the trend. The advantage is that it is possible to estimate the parameters of equation (6) using classical OLS methods. It was also proved by Le et al. (2012) that this procedure is extremely accurate using Monte Carlo experiments.

According to Le et al. (2016), either equations C66 or equations C68 can be used as the auxiliary model. The equations C66 can be rewritten as following:

$$y_t = [I - K]y_{t-1} + K\Pi x_{t-1} + n + t + q_t \tag{C69}$$

where the errors q_t now consist of the lagged difference regressors and the deterministic time trend in $\bar{x}_t$ which affect both endogenous and exogenous variables. Hence we use VARX(1) as the auxiliary model and estimate coefficients by OLS in this research.

D Indirect Reference of DSGE models with supercomputer

The following mind map shows the whole technical process of the project in simulating and estimating the model. For an overview of the files, refer to the notation.xlsx[1].

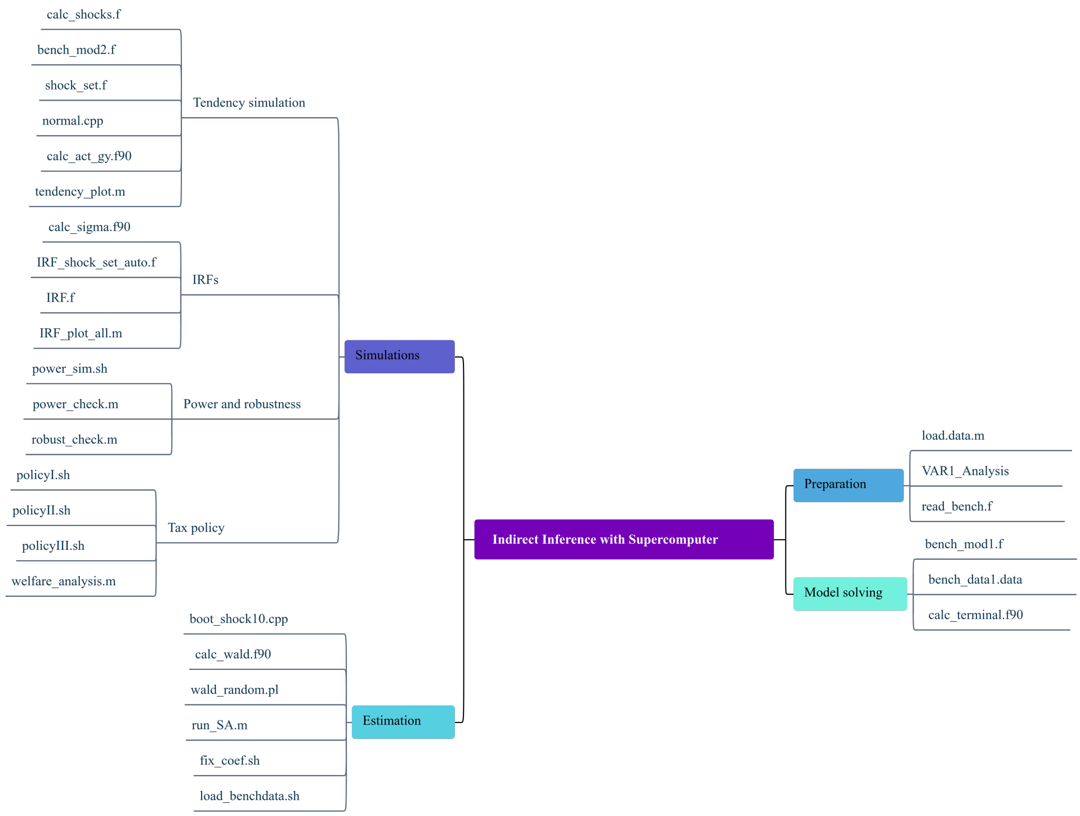


Figure D9. Technical structure of Indirect Inference with Supercomputer

D.1 Preparation

STEP 1: Start the HPC project. Create an account with MySCW and apply for the MySCW project. The current Welsh HPC is called HAWK. Then, apply for the Quality of Service. To run the simulation and estimation, it needs to apply for quality and service at the maximum job 1500, which allows the user to run 1500 jobs on the

[1] Available at Gitlab of Cardiff University, contact wangh53@cardiff.ac.uk for details. We acknowledge the support of the Supercomputing Wales project, which is part-funded by the European Regional Development Fund (ERDF) via the Welsh Government.

HAWK simultaneously (the parallel calculation now available in HAWK). Finally, use Putty and FileZilla in the Windows system or Linux/Mac terminal to interact with the HAWK.

STEP 2: Set up the work environment. To take full advantage of different programming languages and software, This project used Fortran, C++, Matlab, Perl and Microsoft Visual Studio as the programming tools.

STEP 3: Process the raw data. Save the actual data of variables into "raw_data.xslx".

STEP 4: Toolboxes. There are mainly two toolboxes in this project: toolbox77.f and toolbox90.f90.

The following sections explain the main programmes used for the preparation steps.

D.1.1 load_data.m

Step 1: Load all the actual data from Excel.

Step 2: Use VAR1_Analysis.m to generate the initial data of expectation variables.

Step 3: Write the actual data with expectation data into act_data.data.

Step 4: Save the starting endo and exo data into plain data files.

INPUT: raw_data.xlsx, VAR1_Analysis.m

OUTPUT: actual_data.data, start_lags_endog, start_lags_exog

D.1.2 VAR1_Analysis.m

Step 1: Use OLS to fit the actual data with VAR(1).

Step 2: Use the estimated VAR(1) to generate new data.

D.1.3 read_bench.f

Step 1: Load the act_data.data.

Step 2: Extending the data to forecasting periods.

Step 3: Write the bench data files.

INPUT: act_data.data, VAR1_Analysis.m

OUTPUT: bench_end.data, bench_exo.data.

D.2 Model solving

The main algorithm used is type II fix by using **bench_mod1.f** and **bench_data1.data.** Here are the main steps,

STEP 1: Load the input parameters and actual data from bench_data1.data.

STEP 2: Call the **Get_TypeII** in **bench_mod1.f** to calculate Type II fix.

- Specify the dimensions of the innovations of shocks.
- Relate the coefficients in listing to the parameter orders.
- Load the innovations from err(k,j)
- Put all the exogenous variables and shocks in block X.
- Get the type II fix ER(n,i) and the endogenous variables V(N).
- Read and write the type II fix.

D.2.1 bench_mod1.f

The **bench_mod1.f** is the main file used to solve the non-linear DSGE model, including 44 subroutines and functions in total. The **Main** sets up the structure of the project with 6 sections,

- Section 1: Reading in errors and terminal coefficients from external files.
- Section 2: Reading control cards that define parameters and settings for the program.
- Section 3: Reading in initial values of endogenous and exogenous variables, residuals and coefficient by calling "**DATSIM**".
- Section 4: Solving a system of equations by calling "**DEBCAL**".
- Section 5: Performing a rolling forecast by updating the values of the endogenous and exogenous variables and residuals, and then calling "**DEBCAL**" again.

- Section 6: Output is generated and saved to an external file.

Use the **Get_TypeII** to get the type II fix error term and the endogenous variables, which will be used in the model-solving subroutines. The "er" array contains the initial values of the residuals while the "x" array contains the initial values of the exogenous variables. The number of rows in each array is equal to the number of endogenous and exogenous variables, respectively, while the number of columns is equal to the maximum number of time periods over which the the simulation will be run.

Use **CONT(IDY)** to select the required solution routine. For example, **IDY=3** is the Powell-Hybrid algorithm (Originated from MINPAC) and the codes will call **FUNCTH**, which controls the Powell-Hybrid procedure. One important section of **FUNCTH** is it calls **HYBRD1**, which is a modification of the Powell hybrid routine **HYBRD**. **HYBRD1** requires **FCNH** and calls the **HYBRD** to solve the system.

Use **FCNH** to calculate the LHS values of the functions in the model with a new set of values for the endogenous variables, and to calculate the residuals(stored in **FVEC**). **FCNH** calls **EQN** to get new values for each endogenous variable, which **EQN** achieved this by call **typeII**.

HYBRD calls many subroutines to finish the algorithm, including

- calls **FDJAC1** to calculate the Jacobian matrix.
- calls **QRFAC** to compute the QR factorization of the Jacobian matrix.
- calls **QFORM** to accumulate the M by M orthogonal matrix Q from its factored form.
- calls **DOGLEG** to determine the direction P.
- calls **R1UPDT** to determine an Q such that $(S + U * V') * Q$ is lower trapezoidal.
- calls **R1MPYQ** to computes $A * Q$ in the QR factorization.
- uses function **ENORM** to calculate the Euclidean norm of the residuals.
- use function **SPMPAR** to set the machine precision

If **IDY=4**, the model will be solved based on the Levenberg-Marquardt algorithm. The code will choose **FNCTLM**. Similar to **FUNCTH, FNCTLM** controls the Levenberg-Marquardt procedure. One important section of **FNCTLM** is it calls **LMDIF1**, which is a modification of the Levenberg-Marquardt method **LMDIF**. **LMDIF1** calls the original **LMDIF** to solve the system.

LMDIF calls many subroutines to finish the algorithm, including

- calls **FDJAC2** to calculate the Jacobian matrix.
- calls **FCN** to calculate the LHS values of the functions.
- calls **QRFAC** to compute the QR factorization of the Jacobian matrix.
- calls **LMPAR** to determine the Levenberg-Marquardt parameter.
- **LMPAR** calls **QRSOLV** to compute the QR factorization.
- uses function **ENORM** to calculate the Euclidean norm of the residuals.

D.2.2 bench_data1.data

The structure of bench_data1.data was determined by codes in bench_mod1.f. The following table shows the contents of the bench_data1.data

D.2.3 calc_terminal.f90

STEP 1 Set up terminal conditions for the simulations with **subroutine CALCFX(FX)**.
STEP 2 Use **boot_shocks.cpp** to randomise the shocks' order using a random seed.
STEP 3 Use **bench_mod2.f** and **bench_data2.data** to do the simulation.

D.3 Simulations

Simulations of a DSGE model can established from different standpoints which depends on the objectives of the research, the following parts shows the tendency simulation, impulse response and the policy simulation.

Table D2. Model Parameters and Settings

Notation	Value	Description
NEGP	1	Number of endogenous variable groups
NXGP	1	Number of exogenous variable groups
NPER	377	Number of periods for simulation/length of forecast period
LAG	3	Maximum number of lags in the system
NSTART	4	Starting period of forecast
ITMX	99	Maximum number of iterations
IB	0	Data input instruction
NCG	1	Number of coefficient groups
IRAND	0	Maximum number of random errors
NSK	0	Joint or single shock option
IDYN1	3	Parameters to control the solving algorithms
NDOG	16	Number of endogenous variables
LEXOG	32	Number of exogenous variables
IPAR	30	Number of calibrated coefficients in coef.data
IRHO	10	Number of coefficients of shock processes in 'ar_coeff.data'
PX	-1.0	Print option for exogenous terms
P	0	Print option for iteration steps
TOL	0.01	Tolerance level (0.0 < tol < 0.25)
B	0.4	Damping factor (0.0 < b < 1.0)
BR	0	Baserun comparison option
PR	1.0	Baserun print option
NARG	100	Number of expectation variables multiplied by the number of periods in each forecast run (50)
MAXITR	9999	Maximum number of iterations the model uses to solve for the rational expectations
BSTEP	0.15	Step size for updating the coefficients
TOLR	0.09	Tolerance for the rational expectations solution
LDOFIX	1	Indicates whether the model is performing a Type II fix

D.3.1 Tendency simulation

This directory aims to check the relation between inequality and growth in the long-run, start from two identical groups in using the TENDENCY1 folder. Typing **bash** runme.sh *to run the following:*

STEP 1: ***Generate new data files for identical groups.*** *Run* load_data_UK.m to obtain act_data.data with VAR1_Analysis.m to get the values of the expected variables for both identical groups.

STEP 2: Format the data files for Fortran. Use read_bench.f90 to get the ex-

Table D3. Structure of bench_data1.data

Row	Colum	Contents
1	1-19	Parameters control the model solving algorithms
2	1-2	Number of endogenous/exogenous variables
3	1-2	Number of calibrated coefficients and the AR coefficients of shock processes
4	1-7	Parameters control the model solving algorithms
5	1-8	Parameters control the model solving algorithms
6-21	1	Names of the endogenous variables
22-37	1	Names of the solved endogenous variables
38-53	1	Names of the exogenous variables
54-?	1-4	Actual data of the endogenous variables
?-?	1-4	Extended data of the endogenous variables
?-?	1-4	Actual data of the exogenous variables
?-?	1-4	Extended data of the exogenous variables
4622-4624	1-5	Names of the calibrated coefficients
4625-4626	1-5	Names of AR coefficients of shock processes

tended data with forecasting value and format the data with bench_data1.data , bench_data2.data, start_lags_endog start_lags_exog.

STEP 3: ***Compile all relevant programmes.*** calc_terminal.f90,calc_shocks.f90, bench_mod1.f,bench_mod2.f,normal.cpp,shock_set.f,calc_act_gy.f90.

STEP 4: ***Run the main programme.*** *Obtain the* simulation *data by*

- use calc_terminal.f90, calc_shocks_bench.f90, bench_mod1.f and bench_data1.data to get the base data.
- use normal.cpp, shock_set.f, base, bench_mod2.f and bench_data2.data to get the simulation data.

STEP 5: ***Plot relevant tendency data.***

- *use* calc_act_gy.f90 *to* gets the act_growth data in logarithm and algorithm.
- matlab < tendency_plot.m

STEP 6: Clear the intermediate files for next simulation.

- rm -rf bench_data1.data . . .

The following section gives more details about the programme used in the tendency simulation.

calc_shocks.f

Specify the persistence and innovations of the shock process. Errors = actual data - fitted value.

STEP 1: Preparations

- Specify the names for any parameters and steady state of the variables.
- State the dimension of model variables and data.
- Loads the endogenous variables from **'act_data.data'** ;
- Generate some missing data (value=0) with functions: e.g AR_param;
- Replace the missing data (value=0) by generated data.
- Represent the full data inputs.

STEP 2: Calculate the fitted value of variables which has shocks in its equation.

STEP 3: Calculate the residuals by using the $actual\ data - fitting\ data$.

STEP 4: Save the residual data as **'resids.data'**.

STEP 5: Get the detrend residuals by the function **OLS.**

STEP 6: Save the AR coefficients and the innovations.

INPUT: coef.data, act_data

OUTPUT: resid.data, shock.data, ar_coeff.data, bgp_const.data

bench_mod2.f The difference between bench_mod1.f and bench_mod2.f is just the read or write the residuals. bench_mod1.f was used to solve the model we write the residuals, bench_mod2.f was used to simulate the data, we read in the residuals.

shock_set.f Use shock_set.f to set the chosen shocks with the random.data generated from normal.cpp.

There are 10 shocks could be used in the model, for TENDENCY1 directory, we used the labour input shocks only, for TENDENCY2 directory, we used 5 shocks related to the inequality and growth.

normal.cpp Use normal.cpp to randomly draw a number from a normal distribution. We can also use either the building-in random number generators or the Box-Muller transformation to generate normally distributed random numbers from uniformly distributed ones. This is an older method but is still valid.

calc_act_gy.f90 In this file, we calculated the logarithmic/ arithmetic growth rates to the extended periods using OLS. This work can also be done in read_bench.f by using bench_exo and bench_endo data.

INPUT: act_data.data

OUTPUT: act_growth.txt

tendency_plot.m plot the relevant plot in twp steps

STEP 1: Read in data from simulation/act_growth data files

STEP 2: generate capital/income share of G1 and logarithmic/ arithmetic growth

Capital share of G1: $SIM\_OUT(i, 17) = 0.5 * e^{K2}/e^{K}$

Income share of G1: $SIM\_OUT(i, 18) = 0.5 * e^{Y1}/e^{Y}$

Logarithmic growth rate: $SIM\_OUT(i, 19) = Y_{i+1} - Y_i$

Arithmetic growth rate: $SIM\_OUT(i, 20) = e^{Y_{i+1}}/e^{Y_i} - 1$

STEP 3: Calculate the relevant series to be plotted

STEP 4: Plot the relevant figures.

D.3.2 IRFs

This part aims to conduct the IRF analysis by giving one temporary shock equal to each exogenous shock's standard deviation. It could be executed either before or after the estimation.

STEP 1: *Compile all the programmes.*

STEP 2: *Run the base.*

- *use the base model to get the base data.*

STEP 3: *Do the IRF simulation.*

- *use **calc_sigma.f90**, **IRF_shock_set_auto.f** to get the irfsim data*
- *use **IRF.f** to calculate the difference between base and irfsim.*

STEP 4: *Plot figures*

- use ***IRF_plot_all.m*** *to plot the 10 IRF figures.*

calc_sigma.f90 *Calculates the standard deviations of the structural shocks. Input the* shocks.data *with number of periods and number of shocks, it outputs the SD of the shocks to* IRF_sigma.data by using the *VARIANCE* function

INPUT: shocks.data

OUTPUT: IRF_sigma.data

IRF_shock_set_auto.f *Sets the SD to be used by reading* IRF_sigma.data *and setting it as a shock for IRF analysis one by one according to the input* IRF_number, *the IRF shocks data are saved in* irfshocks.data.

INPUT: IRF_sigma.data, IRF_number

OUTPUT: irfshokcs.data

IRF.f *Generates the forecasted impulse response which is the difference between the* base *and* irfsim.

INPUT: base,irfsim

OUTPUT: IRFs*.txt

IRF_plot_all.m *INPUT:* IRFs*.txt *OUTPUT:* IRFs*.jpg

D.3.3 Power and robustness

power_sim.sh This file has two parts, first one resets the coef.data by 1%, 3%, 5%, 7%, 9% randomly fault, and then simulates the data. Second part reset the equations of penalty rate in bench_mod1.f/ bench_mod2.f and then simulate the data.

power_check.m This file mainly used to calculate the Wald statistics and P value generated from above simulations.

robust_check.m This file is used to drew random samples from the innovations in error processes, creating 1000 artificial samples of the same length as the original data – 147 observations. Then bootstrap each of these samples 1000 times. For each sample compute the Wald statistic generated by the bootstraps to check whether the model is accepted or rejected at various confidence levels.

D.3.4 Tax policies

We have three policies experiments in the POLICY directory. We need compare the difference of output and inequality between the results of non-tax regime and each policy scenario. Since we need to change the model equations a little bit, the following table summarized the changes in each scenario:

Table D4. Model Parameters and Settings

Types	Tax policies	Changes of the model codes
Bench	*zero tax rate*	*bench_mod1/bench_mod2 eq. 16: delate the CEQT;*
Policy I	*20% income tax on both groups*	*coef.data 25th: set the CEQT as 0.2 and phi2 a little bit different.*
Policy II	*20% income tax on the rich but no transfer to the poor*	*coef.data 25th: set the CEQT as 0.2 and phi2 a little bit different; Delate the CEQT in G2's equations.*
Policy III	*20% tax on the rich with transfer to the poor*	*coef.data 25th: set the CEQT as 0.2 and phi2 a little bit different; Delate the CEQT in G2's equations. Add* $CEQT * Y_{2t}$ *in G2's income equation.*

For welfare analysis, substract the simulated consumption data and then calculate the relevant welfare with welfare_analysis.m.

D.4 Estimation

The basic logics of the estimation is generating N groups of coefficients, then use one group of coefficients to do the simulation, then calculate the Wald statistics, the target is to find the corresponding group of coefficients with the Wald statistics < 1.65. The following show the detailed steps of the estimation. [2]

***STEP 1**: Preparation for the data and starting values, coefficients*

- *Sub*step 1: Run "load_data.m" to obtain data file "act_data.data"[3], "start_lags_endog" and "start_lags_exog" by

module load matlab/R2019a
matlab < load_data.m

- Substep 2: Set the values like number of variables and parameters, etc. in the upper part "bench_data1a.data" and "bench_data2a.data" as well as in the lower part "bench_datab.data".

- Substep 3: Obtain the model data file "bench_data1.data" and "bench_data2.data" by

bash load_benchdata.sh[4]

- *Sub*step 4: *Run the run_SA.m and fix_coef.sh to get "all_coef.data".*

matlab < *run_SA.m*
bash *fix_coef.sh*

***STEP 2**: Compile all the scripts to the format could be run on the HPC by typing*

[2]Step 1-3 are compiled in the runall.sh, type bash runall.sh to run the project.
[3]Use VAR1_Analysis.m to get the values of the expected variables.
[4]Use read_bench.f90 to get the extend the data with forecasting value and format the data.

***bash** compile.sh*

***STEP 3**: Run Indirect Inference estimation N times by typing*

***bash** sub_search*

The following part shows some details of the relevant programmes used here

D.4.1 compile.sh

```
module load compiler/intel/2018

ifort bench_mod1.f -o bench_mod1
ifort bench_mod2.f -o bench_mod2
ifort calc_terminal.f90 -o calc_terminal
ifort calc_shocks_bench.f90 -o calc_shocks
#gfortran calc_wald_bench.f90 -o calc_wald
ifort calc_wald_bench.f90 -o calc_wald
icc boot_shocks10.cpp -o boot_shocks
```

Load the intel compilers and compile all the codes to the executable programmes.

D.4.2 sub_search.sh

```
#!/bin/bash
rm Wald Trans nsims
rm MD.data
touch MD.data

 sbatch --account=scw1969 -p htc -t 3-00:00 -J UK1870_INEQ_search -o search_out.out -e search_out.err ./wald_perl_script
```

Eliminates some files and create "MD.data" for saving all newly simulated Walds later.

Runs the file "wald_random.pl" for continuing reading new coefficients, simulating and calculating the Wald by using the file "submit_jobs"

D.4.3 submit_jobs.sh

```
sbatch --account=scw1969 -p htc -n 1 -t 0-3:00 -J UK1870_INEQ -o base_out/base.%J.out -e base_out/base.%J.err ./1_run_base
sbatch --account=scw1969 -p htc --qos=maxjobs1500 -n 1 -t 0-3:00 -J UK1870_INEQ -o sims_out/sims.%J.out -e sims_out/sims.%J.err --array=1-1000 -d singleton ./2_run_sims
sbatch --account=scw1969 -p htc --qos=maxjobs1500 -n 1 -t 0-3:00 -J UK1870_INEQ -o wald_out/wald.out -e wald_out/wald.err -d singleton ./3_run_wald
sbatch --account=scw1969 -p htc --qos=maxjobs1500 -n 1 -t 0-3:00 -J UK1870_INEQ -o check_out/check.out -e check_out/check.err -d singleton ./4_run_check
```

This file runs the following files one after another: "1_run_base", "2_run_sims", "3_run_wald", "4_run_check". All the simulated Walds and relevant statistics, combined with the coefficients used will be saved in the data file "log.data".

1_run_base

./calc_shocks *#Calculate the shocks.*

./calc_terminal *#Calculate the terminal condition.*

./bench_mod1.f90 *#Runs the base to write the residuals(type II fix).*

This creates the residual files and exogenous data files in the tmp directory. when reading in the residuals the model should converge in 1 iteration. AR coefficients of the shocks can be found in the ar_coeff.data

2_run_sims ***cp base_run/tmp...*** ***$WDPATH*** *# Copy base_run/tmp to the sims directory so that*

#the simulation model files can read in the

#residuals and exogenous variable files.

./boot_shocks $randseed *#Randomise the shocks' order using a random seed*

./bench_mod2 *#Read in the residuals and do the simulation*

For i {4..246} controls the periods of simulation(the number of files in tmp)

3_run_wald ***./calc_wald*** *# Calculate the Wald statistics*

Find the VAR coefficients in the SIMS directory, actVARcoef.txt and simVARcoef.txt

4_run_check *Check if the Calculated Wald statistics is normal, otherwise, it will be 9999 to show something wrong in above process.*

calc_wald Call the **ADD_BGP_TO_SIMS(nsims)** to add back BGP to simulated data and get the **all_diff_bgp.data**.

Call the **CALC_WALD(nboot)** to calculate the Wald and get **Wald/Trans** data.

Calculate the model specified data: err_y_trend[5], inequalities and aggregate Growth.

Estimate the auxiliary model, VARX , with the N simulated samples.

Estimate the auxiliary model, VARX, with the actual data.

Calculate the Wald statistics.

Write the Wald/Trans data

[5]There is no need of this if we use a VARX in calculating the Wald statistics

Calculate the P value[6]

D.5 Checklist for a new model

Variations could be done by the following changes:

Table D5. Checklist for a new model

Variations	**Code need to be changed**
Change of the periods of simulation	1.bench_mod1.f & bench_mod2.f: %s/248/147/gc 2.main.sh: for i {4..248} → for i {4..147} 3.Gather_sim.f: NSIM=245,NPER=244 → NSIM=144,NPER=143
Change of the shocks	1.shock_set.f :Comment the relevant shocks don't want to count in.

The following table summarises the factors that may impact the final results

Table D6. Factors that may impact the final results

Files	**Components**
normal.cpp	mean=0/0.01/0.02
shock_set.f	Labour input shock/All relevant shocks
normal.cpp	methods to generate the random number

[6] *Find the Wald.txt and MD.data, compare the value in Wald.txt and MD.data. locate the number of the values which is larger than the one in Wald.txt, then divided by nsim.*