A model of inequality and growth for the postwar US economy

Thesis Submitted in Fulfilment of the Requirements for the Degree of
Doctor of Philosophy of Cardiff University
Economics Section of Cardiff Business School, Cardiff University

by **Lei Yue**

Primary Supervisor: Professor Patrick Minford
Secondary Supervisor: Dr. David Meenagh
Secondary Supervisor: Dr. Zhirong Ou
Secondary Supervisor: Dr. Zheyi Zhu

May 2024

Abstract:

This thesis puts forward a model of endogenous growth in which entrepreneurial innovation by household-owned firms depends on their wealth and the extent to which their efforts are taxed and regulated. Firms are heterogeneous by wealth and the model implies that wealthier firms will innovate more than others, even if facing the same marginal tax rates and regulative constraints. I have estimated the model on post-war US data by indirect inference and found a parameter set that matches the auxiliary model data behaviour without being rejected at the 2% level. The model when repeatedly hit by the shocks identified in the estimated equations implies that growth and inequality will tend to rise steadily over time. It is to be expected that this will prompt political demands for redistributive transfers. I have examined how a rising transfer system of taxes and tax credits would impact on inequality and growth. I find that it would reduce inequality and raise the welfare of the average household but at a substantial cost in growth, and a rising cost in the rate of long-term growth. The model therefore suggests a strong reason, given the checks and balances in its Constitution designed to reinforce states' powers, why US governments have built only quite limited redistribution into their tax system.

Contents

List of Figures and Tables

Chapter 1: Introduction

Public concerns in the topic of increasing economic inequality have intensified significantly since the Great Recession of 2008-2009. The nonnegligible effect of income and wealth inequality on economic performance has been well documented in the fields of political economy and macroeconomic modelling. Economists design the models not merely to uncover the causal mechanisms in which inequality affects economic growth but also to enable policymakers to find appropriate policies to achieve higher social welfare. For policymakers, there is a large suite of different models to inform policy decisions. Some of these models are large systems of reduced-form equations that reflects complex relationships across different sectors in the economy. Some are Structural Vector Autoregression (SVAR) models in which shock identifications are based on economic theory. Others are structural dynamic stochastic general equilibrium (DSGE) models, usually estimated via Bayesian methods or Indirect Inference[1]. Normally, the models with reduced form equations and SVAR models are widely used for short-term or medium-term forecasting, while DSGE models are designed for understanding the transmission mechanisms of monetary and fiscal policy and their quantitative effects on the economy.

However, the typical quantitative DSGE model, features either a representative consumer or very limited heterogeneity on the household side. The frontier in economic research, however, has recently shifted toward incorporating heterogeneity and distributional considerations on the household side. For example, Heterogeneous Agent models (HAM) are emerging as a leading framework to investigate the impact of monetary and fiscal policy on the macroeconomy. Starting with the work of Bewley (1980), a heterogeneous agent model can

[1] In structural econometrics, the Indirect Inference approach is one method to estimate structural parameters by using an auxiliary model (including auxiliary parameters). The test power of Indirect Inference is substantially greater (Le et al., 2016).

be used to consider various individual behaviours with the introduction of idiosyncratic shocks and incomplete capital asset market, in which optimal decision on aggregate savings and consumption can be made by the conditions across agents and generate heterogeneity. As the development of computational techniques and rich micro data has proceeded, HAMs can feature a representative consumer or heterogeneity in the household side, which helps policymakers to seek appropriate policies that improve social welfare.

The development of HAM has involved several important stages.

Important early examples of HAM are the papers by Hansen (l985), Aiyagari and Gertler (1991), and Aiyagari (1994). A key character is that only idiosyncratic shocks (like individual income uncertainty and employment uncertainty) were employed for heterogeneity and the numerical algorithm.

The second generation is represented by Díaz-Giménez et. Al. (1992, 2003), and Krusell and Smith (1998, 2006); they developed a new solution method that searches for an equilibrium law of motion for the wealth distribution, generally described by finite order moments for simplicity, with individual decisions assumed to be made based on the distribution of moments. The new generation of HAMs attempts to remove the dependence on aggregate laws of motion when solving for individual behaviour by searching for equilibrium cross-agent distributions in each period described by density functions (Algan et al. 2008; Reiter 2009; Young 2010; Bohacek and Kejak 2018).

All three generations' models can be used to explain some properties of business cycles, however, the effects of wealth and income distributions on economic growth in the long run are not among them. Although there is no consensus on exact intermediate theoretical mechanisms in which inequality affects economic growth, some relevant theories characterised the macroeconomic relations between inequality and economic growth can be summarised as four basic categories according to main features: financial development, political economy, 'human capital accumulation' theory, saving rate and innovation incentive (See Barro, 2000; Cingano, 2014). On the other hand, there is an abundant theoretical literature attempting to establish the links between economic growth and inequality. From the perspective of modelling, there are several primary models of growth and inequality, such as the extensions of endogenous growth Model (See Pagano, 1993; Cingano, 2014), DSGE Monetary Framework (Ferrara and Tirelli, 2015; Colciago, Samarina and Haan, 2019), and Panel VAR models (Atems and Jones, 2014). And the lack of sufficient variations of some inequality indicators and complex interactions between inequality and growth make regression analysis less efficient.

This paper aims to establish a heterogeneous-agent growth model to fit the distributional characteristics of the US economy in recent decades and examine the relationship between inequality and growth. Here we aim to test a DSGE model which creates a bidirectional causal link between them; growth causes inequality and inequality in turn feeds back positively into growth. A DSGE model of this sort with a single household type only, was first estimated by Minford and Meenagh (2020) on the UK postwar economy, with particular focus on the role of supply-side reforms of the 1980s in boosting growth, as these changed the returns to entrepreneurship. They tested it by Indirect Inference on the facts of the episode. The facts of UK inequality also showed substantial movement during this period. This led Yang et al (2021) to explore the links between growth and inequality in the UK, differentiating household types by wealth as described above. Building on the model of Minford and Meenagh and supplementing it with this heterogeneous agent set-up in which random shocks have distributional effects and higher wealth increases the incentive to innovate as an entrepreneur as just explained, testing the model by Indirect Inference they found that this model of growth and inequality was not rejected as a match to the facts of data behaviour. The model implies that there are trade-offs between growth and inequality for policymakers to explore via redistribution in a developed economy like the UK; they showed that for the UK these trade-offs faced diminishing returns, with a rising sacrifice of growth required for a rising reduction of inequality. This occurs because in the model richer households are more prone to undertake entrepreneurial activity; they do so because they have greater risk-tolerance in that the possible costs of failure, which cannot be insured owing to the insurer being unable to observe the firm's idiosyncratic shocks and inputs of effort, cost less in marginal utility to them, the rich, than to the poor. Hence when the economy creates profit opportunities for entrepreneurial innovation, these are grasped more vigorously by the rich than by the poor households. The resulting growth therefore creates greater inequality as its fruits go predominantly to rich households. In turn as the rich group gets richer, it responds more strongly to the economy's opportunities, generating more growth.

The central mechanism in the model is Meenagh et al. (2007)'s endogenous growth mechanism. In this, individuals have entrepreneurship incentives which drive individual productivity growth and further aggregate growth. In addition, we relate individual entrepreneurship incentives to the wealth distribution so that the rich have larger entrepreneurship incentives than the poor, their wealth reducing the costs of entrepreneurial entry. Levine and Rubinstein (2017) also provide some supporting evidence: using NLSY79 data in the US, they find that more entrepreneurs come from well-educated and high-income families. This mechanism reinforces wealth inequality since productivity growth tends to originate mainly with the rich, who in turn reap larger rewards. The mechanism causes wealth to be gradually concentrated on the rich while also gradually raising the growth rate. Nevertheless, this process can be interrupted and even temporarily reversed by aggregate shocks, such as crises and wars, and also by idiosyncratic shocks to income groups; furthermore, it can be, and often is, modified by redistributive policies.

We have estimated the model on post-war US data by indirect inference and found a parameter set that matches the auxiliary model data behaviour without being rejected at the 5% level. The model when repeatedly hit by the shocks identified in the estimated equations implies that growth and inequality will tend to rise steadily over time. It is to be expected that this will prompt political demands for redistributive transfers. We have examined how a rising transfer system of taxes and tax credits would impact on inequality and growth, and find that it would reduce inequality but at a substantial cost in growth, and a rising cost in the rate of long-term growth. The model therefore suggests a strong reason why US governments have built only quite limited redistribution into their tax system.

This paper has following structure. The current introduction is followed by literature review. Chapter 3 sets out my model and the model data. Then chapter 4 describes the Indirect Inference method and the empirical results. Chapter 5 reviews the policy and welfare implications of the results, and discusses what light they shed on US political economy.

Chapter 2: Literature Review

In this section, we begin by outlining the development of heterogeneous agent model in macroeconomics and taking a review on relevant papers. We then go on to some recent work on the interactions of inequality and the macroeconomy. We end by considering the role of and constraints on political policy in more details and looking at various models of how policy affect the economy and their policy implications.

2.1 Literature review on Heterogeneous Agent Models

Representative agent model (RAM) has been widely used in macroeconomics, but it has several disadvantages which have been frequently found. For example, it assume that all economic agents are identical, which cannot capture how different agents react differently to complex economic shocks and limit its ability to fit some stylised facts. Unlike the homogeneity assumption in RAM, the heterogeneous agent model (HAM) assumes that that the economic agents are not homogeneous but rather have various individual behaviours and preference, which can provide a framework for understanding economic dynamics and analysing real-world economic phenomena.

The development of HAMs in macroeconomics has evolved significantly over the years.

In the early stage, the HAMs feature the idiosyncratic risks and incomplete market, and mainly considered how economic agents with different levels of wealth and income face varying investment opportunities and risks. Early work in heterogeneous agent modelling emerged from the work of Bewley (1977). He theoretically built the permanent income hypothesis that agents make consumption decisions based on both their current income and expected long-term average income, and assumed that consumers face a borrowing constraint (i.e. there is no borrowing in the capital market) and owns a random exogenous income. As a result, agents could learn the probability of the stochastic environment and smooth their consumption by taking a stable saving behaviour. Furthermore, Bewley (1980;1983) introduced heterogeneity into macroeconomic by setting multi-sectors (i.e. firms and workers) and multi-consumers (i.e. the role of uncertainty and the implication of imperfect information). Bewley (1986) also introduced the non-convex heterogeneous preferences of consumers on savings and consumption behaviours.

Afterwards, HAM is widely used to explore how heterogeneity can affect business cycle fluctuations and asset pricing. Hansen (l985) concerned one important labour property in business cycles that fluctuations of aggregate working hours are mainly caused by the fluctuations in employment status instead of the individual working hours of the employed workers. Hence, in his second model, consumers supply the indivisible random labour (either work full-time or do not work) and probably receive a subsidy “lottery” from firms to insure the unemployment. The steady-state allocations, however, are same as those in his first homogeneous-agent model because of the unemployment insurance (a full-insurance will be chosen in equilibrium). In the field of asset pricing, Aiyagari & Gertler (1991) use a numerical method to jointly study two asset puzzles: why the equity premium is extremely high and why the risk-free rate is extremely low, based on the Bewley model together with transaction costs. The crucial step of the computational algorithm is to find a constant real interest rate which guarantees a steady state. Huggett (1993) focuses on the second asset puzzle in a pure exchange environment where the individual borrowing constraint is expressed as the “credit balance” which is always greater than a given negative lower bound. The price of the credit balance plays a similar role to the interest rate in Aiyagari and Gertler (1991). Both find the puzzles could be explained by “self-insurance” behaviours that individuals demand much more risk-free assets than liquidity assets due to the uncertainty of income.

The first generation of HAMs shown above basically involves idiosyncratic risks, incomplete market, stochastic processes of uncertain states and equilibrium (convergent) market prices. Here we focus on consumer heterogeneity instead of producer heterogeneity. For example, Caballero (1990) studies the business cycle by introducing heterogeneity in the production sector where each firm among a continuum of firms has an idiosyncratic productivity shock, but consumers are identical. Although most first generation works yield some satisfactory findings that the income is generally less dispersedly distributed than the wealth in model simulations and the aggregate consumption is relatively smooth while the volatilities of individual consumption are significant, both consistent with truth, there are still some shortcomings. For instance, the proportion of individuals who touch the borrowing constraints is quite small, resulting in a limited effect of fluctuating shock distribution on the aggregate behaviour. Furthermore, the lack of aggregate uncertainty is implausible. Díaz-Giménez and Prescott (1992), and Krusell and Smith (1998, 2006) develop a new method to solve models by searching for an equilibrium law of motion for the wealth distribution around which some new numerical algorithms are developed. The distribution is generally described by finite order moments for simplicity and individual decisions are assumed to be made based on the distribution of moments. However, there might be an infinite-dimensionality issue if high order moments are considered for individual optimal decisions (Algan et al. 2014). The new generation of HAMs attempts to remove the dependence on aggregate laws of motion when solving for individual behaviour by searching for equilibrium cross-agent distributions in each period described by density functions (Algan et al. 2008; Reiter 2009; Young 2010).

In the second stage, HAMs are developed based on the work of Krusell and Smith (1998). Such model is used to discuss how heterogeneity affect business cycle fluctuations, suggesting that differences in agents' responses to shocks can amplify or dampen economic volatility. Krusell and Smith (1998) introduce heterogeneity with a continuum of agents into a basic RBC model where endogenous individual production takes the Cobb-Douglas form with both capital and labour as inputs. Individual production also depends on an aggregate productivity shock which follows Den Haan (1996) with two states, determined by a Markov process. Each agent inputs an indivisible labour as that in Hansen (1985) where the idiosyncratic shock $\epsilon \in \{0, 1\}$ with realisation probabilities such that the total unemployment rate in a bad state and that in a good state are always constant. They solve a decentralised problem where individual real wage and real interest rate are determined by marginal productivities respectively of aggregate labour and aggregate capital. They assume that agents predict the law of motion for wealth distribution to forecast prices and then make their optimal decisions. A simplification of this assumption is that consumers are only concerned with the first I moments of the wealth distribution, which is used to approximate an infinite dimensional wealth distribution and H is an undetermined evolution function. The kernel to find an equilibrium solution is to find a stationary H. With such assumptions, the differences in income and wealth among agents leads to varying consumption pattern and saving behaviours, and wealthier individuals tend to save a larger fraction of their income while poorer individuals may consume a higher proportion of their income. Then the role of income and wealth heterogeneity in shaping business cycle dynamics has been examined. Furthermore, the implications of income and wealth distribution for fiscal and monetary policy can be explored.

However, it is noteworthy that the existence of computational difficulties associated with solving HAMs requires the further research to develop more tractable solutions. The new generation HAMs emphasised solving models with a continuum of individuals. Algan et al (2008) (some ideas originate from Den Haan (1996) (1997)) develop a new computational algorithm using a projection method (PJM) to adopt Krusell and Smith (1998)'s model with a true continuum of individuals instead of a large sample approximation. They also use a finite number of moments, to describe the capital distribution and attempt to find a unique law of motion for aggregate states. The PJM could work on a true continuum of individuals and could diminish the dependence of individual solutions on an ad hoc aggregate law if the researcher does not focus on that law, although some algorithm details need to be improved.

Apart from the above framework, HAMs are now widely applied in New Keynesian framework and Wealth distribution studies (such as Florin, 2019; Kaplan et al., 2018; Tobias et al., 2023). Traditional Representative Agents New Keynesian (RANK) models simplify the features of individuals' economic behaviours, while Heterogeneous Agent New Keynesian (HANK) models emphasize the importance of heterogeneity on economic agents. In a HANK model, different economic agents have varying endowments such as income levels, wealth levels, and even consumption patterns; and the responses of different group's consumption to monetary and fiscal polies can differ markedly. Kaplan, Moll and Violante (2018) discussed the effect of monetary policy in a HANK framework, and examined how monetary policy through interest rate changes impacts of consumption, savings, and investment in an economy with heterogeneous households. The poorer households are more responsive to the changes in the interest rate, which suggests that monetary policy can have a stronger effect on aggregate demand. It emphasizes the role of income and wealth inequality in influencing monetary policy effectiveness, that poorer households respond more to interest rate changes than wealthier ones, affecting aggregate demand and economic outcomes. In addition, Bilbiie, Monacelli and Perotti (2024) also examined the interplay between monetary and fiscal policies in order to achieve economy stability and address income inequality. They argue that monetary policy could play a key role in stabilizing the economy during recessions , and that fiscal policy could play a crucial role for redistribution: a lower interest rate from monetary policy can simulate spending and investment, which is helpful for stabilizing output and employment, while tax adjustment and government spending from fiscal policy can address disparities in income and wealth among different groups in economy. They then discussed the trade-off between stabilization and redistribution, and suggested that a more integrated approach to monetary and fiscal policy can enhance both economic stability and equity.

2.2 Literature review on political economy

There have been intense studies of the effects of inequality on economic performance now. We first describes how inequality affect economic growth in details. The inequality, as reflected in income and wealth inequality, also has potential effects on economic growth. For example, some studies had indicated that economic inequality could promote efficiency, and an increasingly unequal income distribution would not cause negative effects on economic growth. In contrast, some recent research suggests that point to the negative consequences of rising economic inequality for the general development of the economy (eg. Albig et al., 2017). Although there is contradictory argument, relevant theories characterising the macroeconomic relations between inequality and economic growth can be summarised as some basic categories according to main features: imperfection of credit market, political economy, human capital accumulation, saving rate and innovation incentive (See Barro, 1997; Barro, 2000; Voitchovsky, 2005; and Cingano, 2014; Gabaix et al., 2016).

Generally speaking, the relationship between financial development, income distribution and economic growth is an important topic in economics. The role that financial markets play in the process has also received much attention, because individual's investment in financial market has direct effect on economic growth. With the economic growth as reference, the relationship between financial development and inequality has had attention (Demirgüç-Kunt and Levine, 2009). Different scholars hold different views on the role of financial markets. Some economists believe that financial development can reduce inequality and thus promote economic growth, and argue that there is a linear relationship between growth and inequality.

Speaking specifically, Banerjee and Newman (1993) divided the society into the poor and the rich, and only investment could help individual gain entrepreneurship. With an imperfect capital market, the poor could borrow money easily and become entrepreneurs by investment, and thus the inequality would decrease in the long term. With perfect financial markets, only the rich can invest and become entrepreneurs, which lead to the existence of inequality. Based on their research results, the development of capital markets would bridge the income gap and further stimulate economic development. Similarly, Galor and Zeira (1993) replaced the entrepreneurship with human capital, and pointed out that the impact of the capital market would be to raise borrowing cost, so inducing the poor to borrow less, thus the income gap would exist in the long term. And in their model, the development of financial markets could improve economic performance. A similar discussion of human capital accumulation also occurs in other work (De Gregorio, 1996). In addition, Piketty (1997) used a credit-constrained Solow model to research the effect of wealth distribution on economic performance. Similarly to that in an earlier literature, the initial wealth distribution would not affect the total output with the perfect financial markets; while the imperfect financial market exists, wealth distribution will affect the size of the loan, which would therefor influence individual status. And Piketty concludes: "aggregate output is higher in steady with lower rates and interest rate depended on the size of loan; and short-run interest rate can be self-sustaining and could have long-run effects on output through the induced dynamics of the wealth distribution and credit rationing." The recent research of Beck, Demirgüç-Kunt and Levine (2007) reaches a similar conclusion that the development of financial sector would reduce inequality. According to previous three studies, the equilibrium of income distribution is closely related to the initial status of wealth distribution. Therefore, the government's income distribution policy, such as tax, would affect the long-run income distribution and economic performance.

In contrary, some studies think the development of financial market could have limited effects in reducing inequality (Cagetti and De Nardi, 2005). Cagetti and De Nardi (2005) pointed out that the size of investment is based on the entrepreneur's initial wealth, and more restrictive borrowing constraints in the capital market would cause less wealth concentration, lower the average size of firms, reduce aggregate capital, and the fraction of entrepreneurs. Therefore, financial development would possibly generate inverse results: it makes more individuals choose to become entrepreneurs, expand the economic inequality; but it also expands the scale of their investment and promotes economic development.

Some studies argue there should not be a simple linear relationship between financial sector development and income inequality, and that instead there is an inverted U-curve relationship between the two variables, i.e. ‘Financial Kuznets Curve’ (Greenwood and Jovanovic, 1990). Although some studies disagree with this theory (Clarke, Xu and Zou, 2003), there was also support for it (Nikoloski, 2010). Starting with the research of Greenwood and Jovanovic (1990), they pointed out there is a nonlinear relationship between financial sector development and income distribution. This is because of the various individuals’ tolerance to borrowing cost. Considering the economic cycle, the attitudes and responses to borrowing costs are different between the poor and the rich: when there is an economic downturn, the rich could make use of the financial institutions’ service and invest , so widening the gap between the rich and the poor; when there is an economic upturn, the borrowing cost is affordable for both the poor and the rich, and both two social class could invest, which would narrow the income gap. That is to say, a financial Kutznets curve could be generated. In addition, their article also shows that financial intermediaries could promote economic growth. In addition, some economists reached a similar conclusion by changing the assumption on individual’s characters, such as entrepreneurial skills (Lloyd-Ellis and Bernhardt, 2000) and moral hazard (Chakraborty and Ray, 2007). Lloyd-Ellis and Bernhardt (2000) assumed that each economic agent’s wealth and entrepreneurial skills are different, and he/she should afford fixed cost to develop entrepreneurial skills in order to improve individual productivity. In the early stages the income gap will widen due to imperfect information that requiring the poor pay extra fee; while in the later period of economic development, entrepreneurship skills determine the status of the economy. If there were more entrepreneurs with strong capability, it presents a U-shaped shape; if there were fewer, it shows cyclical change. Furthermore, the research of Chakraborty and Ray (2007) also shows unequal income distribution would reduce the size of credit market and prevent the development of financial sector.

Some economists discuss mechanism from aspects of saving rates and investment. The paper of Barro (2000) mentions that some Keynesian economists believe individual saving rates rise with the increase of income. According to this view, redistribution of resources in an economy would be likely to lower the aggregate saving rate. Then a rise in inequality tends to raise investment according to this channel. Then more inequality means higher economic growth. The previous discussion of imperfect financial markets shows a similar mechanism. Here the theory of aggregate saving rates provides a reasonable interpretation for a positive effect of inequality on economic performance. As for investment behaviour, Voitchovsky (2005) pointed out that the rich have enough wealth to implement investment plans, and have less constraints on borrowing. The rich also represent the main social group creating savings in the economy. In addition, with enough wealth, they also could spread the risk of investments and could receive a higher rate of return. With this as the same channel mentioned by Barro (2000), the same conclusion follows: inequality may promote economic growth.

In these models of imperfect financial markets, there are usually the following assumptions: (1) The imperfect credit market reflects asymmetric information and limitations of the law. (2) The individual's initial status of wealth and income determines the borrowing limit, which would further affect individual's investment behaviour. In these models society usually is divided into two groups, the poor and the rich. Under these conditions, normally, the rich could borrow more and achieve a higher rate of return on investment. Nevertheless, further aspects come into play. The endowment of the individual could vary according to entrepreneur skills, learning ability and the accumulation of human capital; also economic behaviour and preferences. My work here assumes that rich and poor households only differ in their holdings of wealth and both have full access to perfect financial markets; however, for the rich the marginal utility of entrepreneurial costs is lower due to their greater wealth; this makes them more inclined to take the risks of entrepreneurial change. Hence wealth inequality has a positive effect on entrepreneurial effort; this in turn generates more growth, which in turn accrues mostly to the rich households, so increasing inequality.

I now turn to related work on political economy. Barro (2000) analyses the economic behaviour of social classes. In the economy, if the mean income is higher than the median income, then certain economic groups would tend to favour the redistribution of resource from rich to poor. And such redistribution of resources includes not only normal transfer payments but also involve public expenditure programs and even regulatory policies (Aghion and Bolton, 1997). In this case, Barro believed that inequality could have a negative effect on economic growth even if no redistribution of income takes place in equilibrium. Voitchovsky (2005) believed that such redistribution of resource is likely to have an ambiguous effect on growth. Specifically, more inequality means increasing taxation: the rich need to be taxed， but the productivity of poor agents would be increased, through the channels of gaining government public spending or reducing instability (Alesina and Perotti, 1996). In total, it is beneficial to the total economy. In addition, recent research (Van-Der-Weide and Milanovic, 2014) finds that high inequality reduces the income growth of the poor and increases that of the rich. If such political redistribution did excess damage to the rich, it would be harmful to the economy’s efficiency; but lack of redistribution would be seen as unfair, creating political instability which would be bad for economic growth (Alesina and Perotti, 1996). From such a perspective, it is often hard to determine the effect of inequality on economic growth. However, it is important to balance the relationship between efficiency and fairness. This work is focussed on the social unrest caused by inequality. According to Barro (2000), it economic inequality would motivate the poor to engage in crime, riot and other disruptive activities. Voitchovsky (2005) also pointed out that the economic inequality and poverty could be explanatory factors for crime behaviour; and a link between political instability, economic inequality and economic performance appears in numerous empirical studies. Alesina and Perotti(1996) believed that the increased risk due to insecurity, in turn, has negative effects on individuals' investment behaviour, and further affects growth. In total, a high level of inequality would induce social sociopolitical unrest, while the increased crime behaviour would in turn do damage to economic development, which would cause further internal imbalances in the economy. Barro (2000) suggested that the tendency for redistribution to reduce crimes could raise income equality and ultimately improve economic performance.

The model here also connects to the literature on entrepreneurship. There is a huge literature on modelling how entrepreneurship boosts growth via the innovations of entrepreneurs. Most 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; Grossmann 2009), while Gabaix et a (2016) highlighted the mechanism behind inequality in the context of emerging technologies and changing labour market. There are also studies that investigate the relationship between entrepreneurship and wealth, and the impact of redistribution policies on entrepreneurship (Cagetti and De Nardi 2006, 2009; Garcıa-Penalosa and Wen 2008; Atolia and Prasad 2011; Doepke and Zilibotti 2014). In all these models, risk insurance is vital because of the uncertain cost of Schumpeterian entrepreneurship. The model here assumes instead that households evaluate prospects without access to insurance, because the idiosyncratic risks they run cannot be nsured; and so consider the deterministic cost of entrepreneurship such as market regulatory barriers and government barriers like taxes. The assumption that households cannot insure is in effect an unavoidable friction that handicaps poorer households. The individual's innovation also play an important role in the relationship between economic growth and inequality. In an economy, individual seek for own profit-maximization. Specifically, talented individuals could expect a higher level of economic return based on professional skills. A concentration of talented and skilled individuals in the upper income ranks is conducive to technological progress, and therefore to economic growth (Voitchovsky, 2005). Recent research shows that based on a endogenous growth mechanism, individuals have entrepreneurship incentives that would induce the increase of individual's productivity and therefore aggregate growth. (Minford et.al., 2007) Thus, inequality would prompt social mobility: an individual who has talent or skills would use this endowment to pursue maximizing profit, and therefore drive the growth of individual productivity and aggregate growth.

A variety of empirical methods have been used to test the relationship between inequality and economic growth. Many studies adopt regression analysis, estimating reduced form equations on panel data- Pagano, 1993; Cingano, 2014). The initial endogenous growth Model embedding the financial sector was proposed by Pagano (1993), and Cingano (2014) expands it by adding human capital. Another approach takes a Dynamic Panel VAR model- thus Ferrara and Tirelli (2015) investigate the redistributive effects of a disinflation experiment in an otherwise standard medium-scale DSGE model augmented for Limited Asset Market Participation, implying that a fraction of households does not hold any wealth. They highlight two key mechanisms driving consumption and income distribution: the cash in advance constraint on firms; and the response of profit margins to disinflation, which is crucially dependent on the two most used pricing assumptions in the New-Keynesian literature. For the panel VAR model, Atems and Jones (2014) examines the effects of inequality on per capita income and the effects of per capita income on income inequality, by modelling a comprehensive cross-state panel that allows for the estimation of the dynamic responses of inequality and per capita income using panel vector autoregressive (VAR) models.

Such empirical methods face difficult identification problems: when an economy grows fast, whatever the reason that is triggering it, many accompanying features occur as well that enter the list of suggested causal/control factors: there is much R&D, more government spending on infrastructure, more education spending both public and private, better institutions and so on, including inequality. It is difficult to identify the factors that are driving growth and are driven by growth. Not surprisingly, we find in these panel data results a wide variation in the relationships between inequality and growth. But it is hard to disentangle causality from them. due to these identification issues. It is for this reason that we set up a tightly identified DSGE HAM model and test it by indirect inference; rejection of the model implies rejection of the identifying causal theory.

Chapter 3: Model setting and Model Data

We follow Yang et al. (2021)'s heterogeneous agent endogenous growth model. In it, the households face a trade-off between leisure, working and time spent on entrepreneurial strategy, and they can enhance individual productivity to promote economic growth by doing the last. The model assumes wealthy people are more likely to able to afford the costs of entrepreneurship, and then wealth concentration stimulates entrepreneurship and aggregated economic growth.

The economy is populated by households who own their own firm. Households are divided into groups, where these two groups are differentiated according to their initial wealth holdings . They consume and produce goods. Notice that all households share the same utility function; they differ only in their history of shocks that has led them to be by past endowment either in the rich group or in the rest.

3.1 The Model

3.1.1 Individual

Individual households own their firm, in which they engage their own inputs in the form of labour, capital and entrepreneur-driven productivity. We assume that the population in an economy is comprised of two groups with constant population weights $\mu_i; i = 1,2, \mu_1 + \mu_2 = 1$, where these two groups are differentiated according to their initial wealth holdings.

The representative agent i in each group maximizes an intertemporal utility function given by:

$$E_{i,t} \sum_{j=0}^{\infty} \beta^j U_{i,t+j} \tag{1}$$

where $E_{i,t}$ is the expectations of representative agent in group i in period t, β is the discount factor, and the utility function $U_{i,t}$ comprises consumption $\mathrm{C_{i,t}}$, leisure (defined as the normalised unit of available time for the household less labour input $\mathrm{N_{i,t}}$ and entrepreneurship time $\mathrm{Z_{i,t}}$)

$$U_{i,t} = \phi \frac{\epsilon_{i,t}^{C}\left(C_{i,t}\right)^{1-\psi_1}}{1-\psi_1} + (1-\phi)\frac{\epsilon_{i,t}^{N}\left(1-N_{i,t}-Z_{i,t}\right)^{1-\psi_2}}{1-\psi_2} \tag{2}$$

where ψ is related to the inverse of intertemporal elasticity of substitution of corresponding variable, ϕ measures the weights of consumptions and leisure in the utility function. Note that the two households experience household-specific shocks to consumption and labour supply preference: $\epsilon_{i,t}^{C}$ and $\epsilon_{i,t}^{N}$ are consumption and labour supply preference shocks.

Meanwhile, individual i owns the firm and also works for it. The real budget constraint is

$$\mathrm{C_{i,t} + b_{i,t+1} + K_{i,t} - (1-\delta_k)K_{i,t-1} \leq Y_{i,t} + (1 + r_{t-1})b_{i,t} - \pi_t Z_{i,t}} \tag{3}$$

Where $\mathrm{b_{i,t}}$ is the bonds purchase, $\mathrm{r_{t-1}}$ is the risk-free interest rate. And entrepreneurship has a unit cost $\mathrm{\pi_t}$ and the total cost of entrepreneurship for individual i is $\mathrm{\pi_t Z_{i,t}}$.

Individuals have a Cobb-Douglas production function (4) where the non-stationary individual productivity $\mathrm{A_{i,t}}$ envolves as the process (5) which depends on individual time spent on entrepreneurship as well as the aggregate productivity shock $\mathrm{\varepsilon_t^A}$.

$$\mathrm{Y_{i,t} = A_{i,t}\left(K_{i,t-1}\right)^{\alpha}\left(N_{i,t}\right)^{1-\alpha}} \tag{4}$$

and

$$\frac{A_{i,t}}{A_{i,t+1}} = \left(\theta_1 + \theta_2 Z_{i,t}\right)e^{\varepsilon_{a,t}} \tag{5}$$

The first order conditions for this optimization problem are:

$$\epsilon_{i,t}^{C}\left(C_{i,t}\right)^{-\psi_1} = \beta(1+r_t)E_{i,t}\left[\left(C_{i,t+1}\right)^{-\psi_1}\epsilon_{i,t+1}^{C}\right] \tag{6}$$

$$\frac{(1-\phi)\epsilon_{i,t}^{N}}{\left(1-N_{i,t}-Z_{i,t}\right)^{\psi_2}} = \phi\left(C_{i,t}\right)^{-\psi_1}(1-\alpha)\frac{Y_{i,t}}{N_{i,t}} \tag{7}$$

$$\epsilon_{i,t}^{C}\left(C_{i,t}\right)^{-\psi_1} = \beta\left\{E_{i,t}\left[\left(C_{i,t+1}\epsilon_{i,t+1}^{C}\right)^{-\psi_1}\epsilon_{i,t+1}^{C}\right]\left[\alpha\frac{Y_{i,t+1}}{K_{i,t}}+(1-\delta_k)\right]\right\} \tag{8}$$

$$\frac{(1-\phi)\epsilon_{i,t}^{N}}{\left(1-N_{i,t}\epsilon_{i,t}^{N}-Z_{i,t}\right)^{\psi_2}}+\phi\frac{\pi_t}{\left(C_{i,t}\right)^{\psi_1}} = \phi\theta_2[\left(\theta_1+\theta_2 Z_{i,t}\right)e^{\varepsilon_{a,t}}]E_t\left[\sum_{j=1}^{\infty}\beta^j\frac{Y_{i,t+j}}{\left(C_{i,t+1}\right)^{\psi_1}}\right] \tag{9}$$

where the first four equations are optimal rules for consumption, labour, capital, and entrepreneurship time respectively, and the log linearised equations are listed in Section 3.3[2]

3.1.2 Entrepreneurship Penalty Rate

Equation (9) is the optimal decision rule for $Z_{i,t}$; it can be approximated, by treating $\frac{Y_{i,t}}{C_{i,t}}$ as a random walk before the steady state[3] as:

$$(1-\alpha)\frac{Y_{i,t}}{N_{i,t}}+\pi_t = \frac{A_{i,t}}{A_{i,t+1}}Y_{i,t}\frac{\beta\theta_2}{1-\beta}e^{\varepsilon_{a,t}}. \tag{10}$$

where $(1-\alpha)\frac{Y_{i,t}}{N_{i,t}}$ indicates real wage rate $w_{i,t}$ in a perfect labour market. We define an 'entrepreneurship penalty rate' $\pi_{i,t}' = \pi_t/w_{i,t}$ to reflect the total cost, and re-arrange equation (10) as follows:

[2] Each consumption or labour supply preference shock emerges as separate two shocks linearised equation of two groups. for example, given rich group's consumption Euler equation $\left(C_{1,t}\right)^{-\psi_1} = \beta(1+r_t)E_{1,t}\left[\left(C_{1,t+1}\right)^{-\psi_1}\epsilon_{i,t+1}^{C}/\epsilon_{i,t}^{C}\right]$, the log linearization takes the form: $lnC_{1,t} = E_{1t}lnC_{1,t+1} - \frac{1}{\psi_1}(r_t + ln\beta) - \frac{1}{\psi_1}(\ln\epsilon_{1,t+1}^{C} - \ln\epsilon_{1,t}^{C})$, where $-\frac{1}{\psi_1}(\ln\epsilon_{1,t+1}^{C} - \ln\epsilon_{1,t}^{C})$ is the shock process. We denote the loglinarised errors respectively as $\epsilon_{i,t}^{C}$ and $\epsilon_{i,t}^{N}$.

[3] See Yang et al (2021) for the proof.

$$\frac{A_{i,t+1}}{A_{i,t}} = \frac{\beta\theta_2 \frac{Y_{i,t}}{w_{i,t}}}{(1-\beta)\left(1+\pi'_{i,t}\right)} e^{\varepsilon_{a,t}} \tag{11}$$

which indicates that as current income and consumption fall, the marginal utility of these costs rises, raising the disincentive to entrepreneurship. A log-linearized approximation to Eq.(11) is:

$$ln\left(A_{i,t+1}\right) - ln\left(A_{i,t}\right) = \phi_{1,i} - \phi_{2,i}\pi'_{i,t} + ln\left(Y_{i,t}\right) - ln\left(w_{i,t}\right) + \varepsilon_{a,t}, \tag{12}$$

where $\pi'_{i,t}$ can be specified as a function of the economy-wide money costs of entrepreneurship relative to the average wage. Via Eq.(12), it can be seen that individual penalty rate falls with rising income. We then assume that entrepreneurial costs are indexed to the general rise in income and so wealth. Thus, the entrepreneurship penalty rate evolves according to

$$\pi'_{i,t} = \rho_0^{\pi} + \rho_1^{\pi}\pi'_{i,t-1} - \rho_2 * Q\left(\frac{K_{I,t-2}}{K_{t-2}}\right) + \varepsilon_t^{\pi}, \tag{13}$$

where $\rho_1 \geq 0;\ \rho_0, \rho_2 > 0$.

3.1.3 Aggregation and Market clearing

In addition, each aggregate variable is the weighted sum of two group of individuals:

$$Y_t = \mu_1 Y_{1,t} + \mu_2 Y_{2,t}; \tag{14}$$

$$K_t = \mu_1 K_{1,t} + \mu_2 K_{2,t}; \tag{15}$$

$$C_t = \mu_1 C_{1,t} + \mu_2 C_{2,t}. \tag{16}$$

Equilibrium in the goods market is expressed by the resource constraint

$$Y_t = C_t + K_t - (1-\delta)K_{t-1} + \varepsilon_{Y,t} \tag{17}$$

where $\varepsilon_{Y,t}$ can be thought of as an aggregate demand shock.

3.2 Data

We use seasonally adjusted quarterly data in the US from 1970Q1 to 2018Q4. The aggregate output and consumption are GDP and household personal consumption expenditure respectively, measured by chain volume with base year 2012 from Federal Reserve Bank. The nominal interest rate is the interest rate reported by Federal Reserve Bank, and expected next-quarter inflation rate is 10-Year Breakeven Inflation Rate. The real interest rate is the difference between this and the expected next-quarter inflation rate. Aggregate labour is the labour Force Participation. The aggregate capital stock is capital stock at constant national prices. The demand for financial asset is the total financial asset holden by household.

For data of entrepreneurship, we follow Minford and Meenagh (2020): disincentives are determined by two factors, “labour market regulation” and the tax rate. LMR describes the degree of intervention in the labour market, which is measured by the average of two indicators “centralised collective bargaining” and “mandated cost of worker dismissal” reported by the Fraser Institute. The former describes the procedure for both employers and employees to make a collective agreement, and the latter reflects the cost of all social security and payroll taxes and other mandated costs. For the tax rate, we use here the corporation tax rate.

The two groups we consider are the two income deciles, the top 10% and the bottom 90%. Data on income and wealth distributions come from the World Inequality Database.

Table 1 contains all definitions and sources of data used in this study. Most US data are sourced from the Federal Reserve Economic Data, and others are from Organisation for Economics Cooperation and Development (OECD) and Fraser Institute. The inequality indicators are sourced from World Inequality Database (WID).

Figure 1 Model data

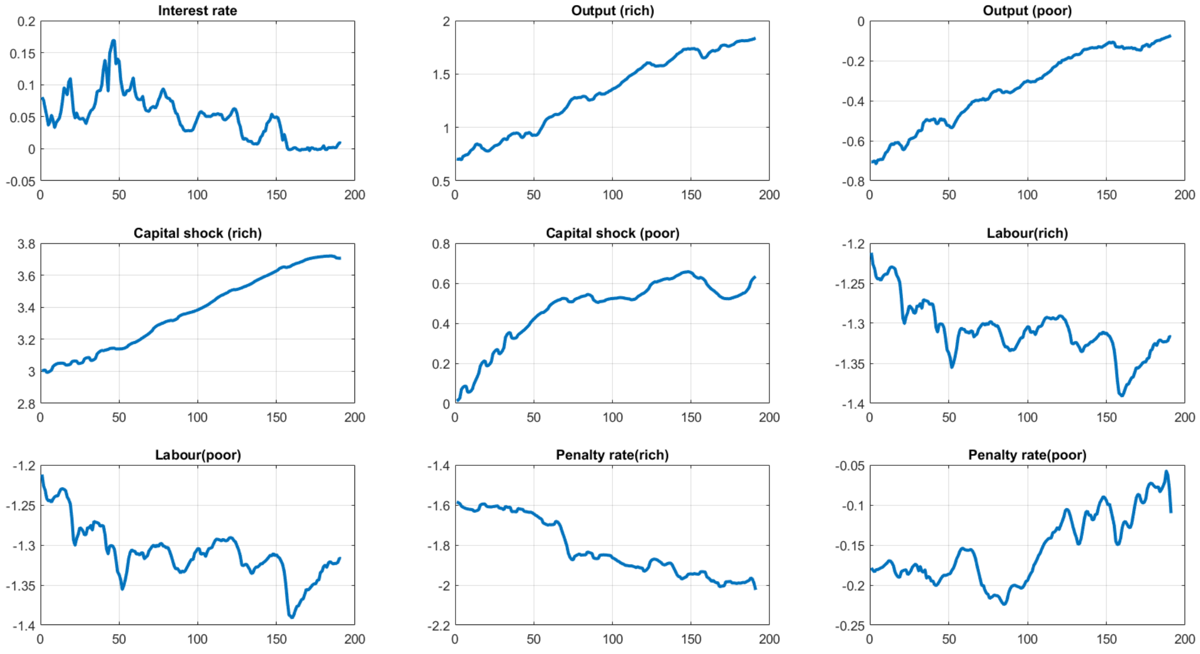


Table 1. Data description

Symbol	Variable	Definition and description	Source
Y	Output	Gross Domestic Product	FRED
K	Capital Stock	Capital Stock at Constant National Prices	FRED
C	Consumption	Personal Consumption Expenditure	FRED
r	Real interest rate	Calculated by the difference between nominal interest rate and next periods inflation	
	Nominal Interest rate	Federal Funds Effective Rate, Percent	FRED
	Inflation	Consumer Price Index: All Items Excluding Food and Energy	FRED
N	Labour Supply	Calculated by the ratio of Employment to Total workforce, and then adjusted by Weekly Working-Hours	
	Total workforce	Civilian Labour Force Level, Thousands of Persons	FRED
	Total unemployed	Unemployment Level, Thousands of Persons	FRED
	Weekly Working-Hours	Average Hours of Work Per Week, Total	FRED
$\boldsymbol{\pi'}$	Aggregate Penalty Rate	Estimated by the weighted average of 'LMR' and the 'corporation tax rate'.	
CTR	Corporation tax rate	Statutory Corporation tax rate	OECD
LMR	Labour market regulation	Estimated by the weighted average CCB and MCD	
CCB	Centralized Collective bargaining	Describes the procedure for both employers and employees to make a collective agreements	Fraser Institute
MCD	Mandated cost of worker dismissal	Describes the cost of all social security and payroll taxes and the cost of other mandated	Fraser Institute
TUM	Trade union members ratio	Calculated by the ratio of Trade union members to All Employers	OECD
CRT	Corporation tax rate	Statutory Corporation tax rate	OECD
A	Productivity	Estimated by Solow Residuals	
	Top10% Income Share	Pre-tax national income share	WID
	Top10% wealth Share	Net personal wealth share	WID
I	Real Investment	Real Government Consumption Expenditures & Gross Investment	FRED

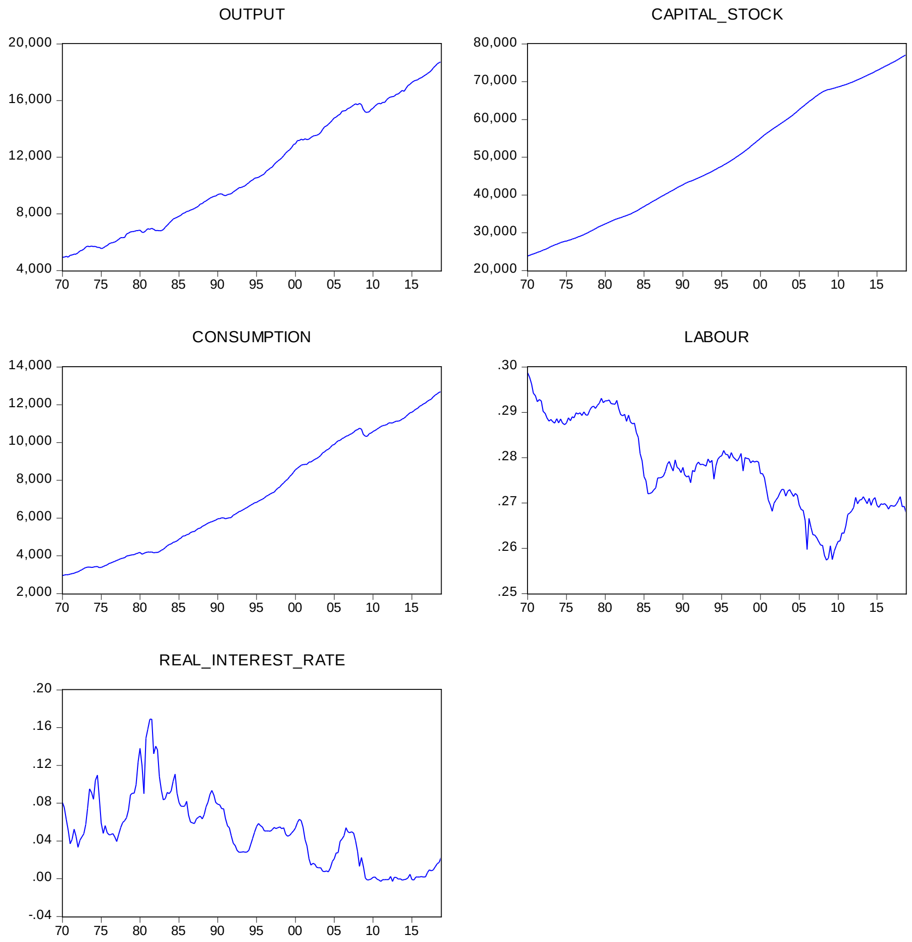
3.2.1 Data for Aggregate Data

In order to exclude inflation's effect, we collect the data measured by chain volume with a base year of 2012. For time series which have a different base year, we rescale them to 2012 data by first creating a new index diving each year of the series by its 2012 value, and then multiply each year's index by the corresponding 2012 current US dollar price value.

Aggregate output and consumption are quarterly GDP and Personal Consumption Expenditure respectively. Real interest rate is calculated by the difference between monthly nominal interest rate and monthly next-period inflation, and then converted into quarterly data: nominal interest rate is the target interest rate set by the Federal Open Market Committee, measured by monthly Federal Funds Effective Rate; and inflation is Consumer Price Index that considers all terms excluding food and energy. The aggregate labour is calculated by first diving the number of employment in US by the toral workforce, and then adjusting by weekly working-hours.

As for the available data of capital stock, the raw data published by the FRED is measured in millions of 2017 U.S. Dollars. By introducing Real Broad Dollar Index with the base year of 2006, we first covert convert the capital stock data into constant 2006 US dollars, and then rescale the 2006 data to 2012 by creating an index diving each quarterly data of the constant 2006 series by its 2012 value (therefore, index in 2012 will equal 1), and multiplying corresponding capital stock data in 2006 US dollars by each year's index.

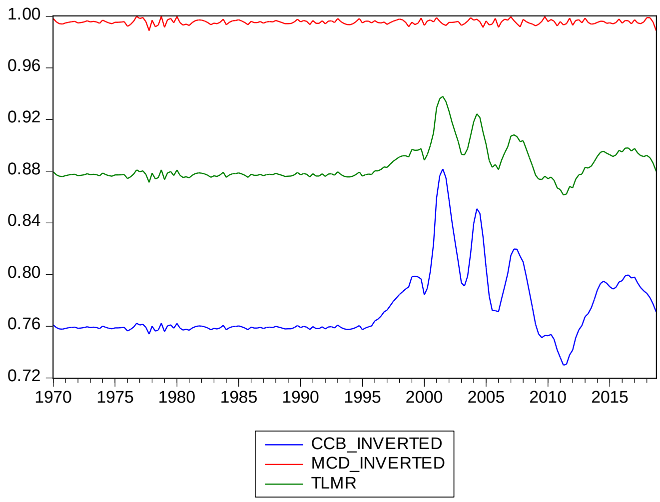
Figure 2. Data on aggregate variables (level unit: billion USD)



Aggregate Penalty Rate, π_t', measures the entrepreneurship cost. In specific, a decrease in Aggregate Penalty Rate indicates stronger entrepreneurship incentives and an increased time spending on entrepreneurship z_t, in which could further improve productivity. Following the ideas of Minford and Meenagh (2016) and Yang et al (2020), disincentives to Aggregate Penalty Rate are determined by 'labour market regulation' (LMR) and the tax rate.

To measure the effect of labour market regulation on π_t', two yearly indices are selected from Economics Freedom indicators reported by the Fraser Institute: the Centralized Collective Bargaining index (CCB) and the Mandated cost of worker dismissal index (MCD). CCB describes the procedure for both employers and employees to make a collective agreements, and MCD describes the cost of all social security and payroll taxes and the cost of other mandated. The higher scores of two indices indicates an increase in labour market flexibility. Both two indices are rescale to a [0,1] interval. Then we use a Denton proportionate variant adjustment method to interpolate quarterly series of CCB and MCD by introducing a higher frequency series, e.g. 'trade union member' rate (TUM), because the unionisation rate is correlated with bargaining power of unisons (or CCB), and correlated with increased protection of worker welfare with a common and stronger worker voice represented by unions (or MCD). The weighted average of resulting quarterly series for CCB and MCD are used to generate transformed 'labour market regulation' (TLMR), see Figure 3.

Figure 3. Generated Date on Labour Market Regulation index



On the other hand, we consider the role of corporation tax rate (CRT) because entrepreneurship activity is more sensitive to such tax. CRT reflects the tax environment. We then calculated Aggregate Penalty Rate π_t' by calculating the weighted average of CRT and TLMR, see Figure 4. And all the relevant date on entrepreneurship penalty rate are shown in Figure 5.

Figure 4. Generated Data on Entrepreneurship Penalty Rate $\pi_t{}'$

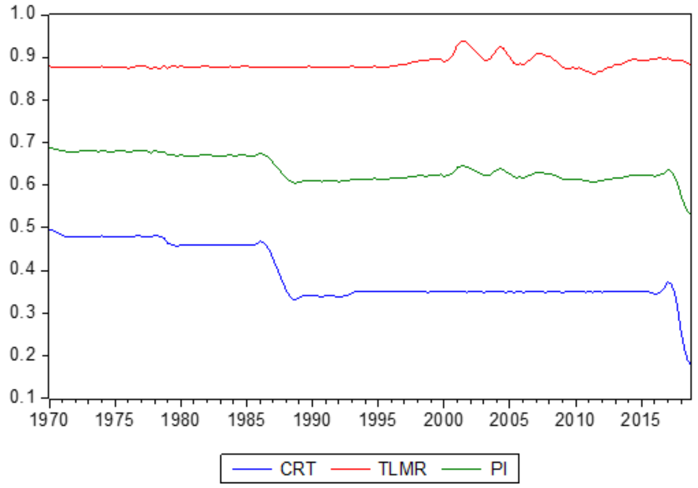
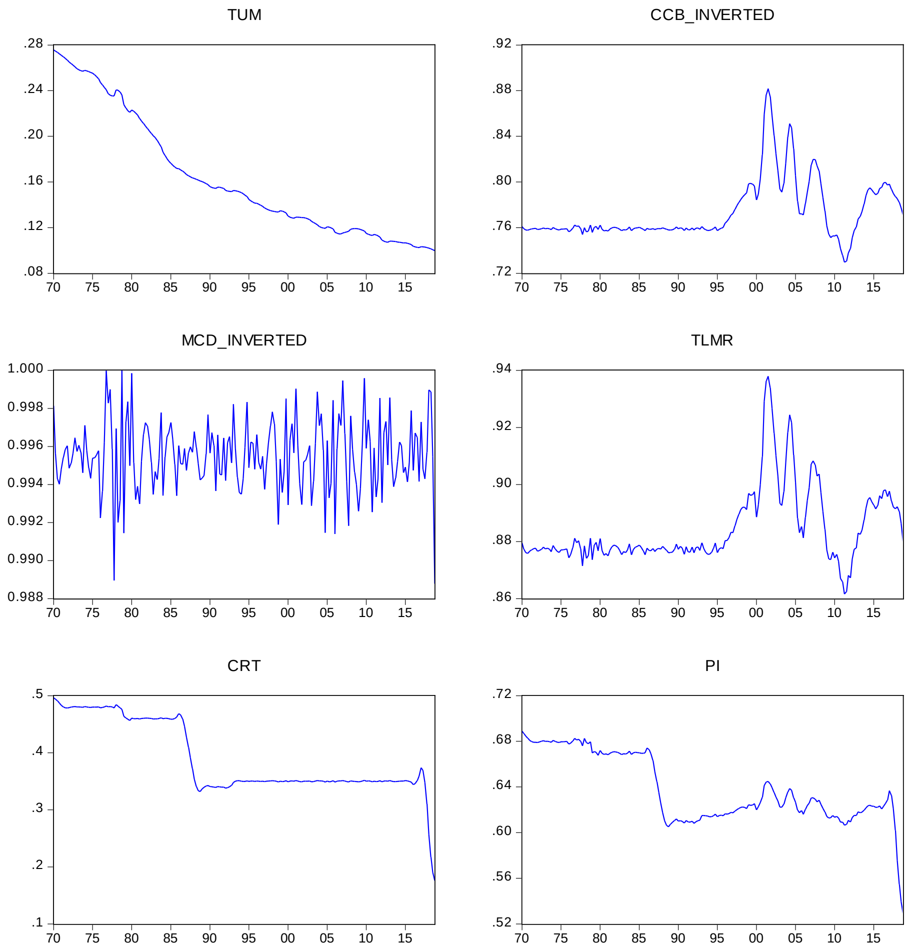


Figure 5. All relevant data on Entrepreneurship Penalty Rate



3.2.2 Individual data

We assume that there are two groups in the economy, in which the groups are classified by income deciles, i.e. top 10% and bottom 90% in US. We also hold the assumption about the existence of representative agent in two groups, and the agent in corresponding group has the typical decision-making behaviour regardless of the agents' movements across income groups in difference time periods.

The two groups are classified as the rich and the poor by income deciles, i.e. the top 10% and the bottom 90%. Date on income and wealth inequality are sourced from World Inequality Database (WID). The individual data on representative agents in two groups are transformed by the aggregate date from the aggregate data according to corresponding group population weights, denoted by μ_i (e.g. $\mu_1 = 0.1$; μ_2 =0.9). Then individual output and capital can be generated directly by multiplying aggregate output and capital shock by population weights respectively, and individual consumption can be obtained from individual budget constraint (See figure A5).

To estimate individual labour supply, we calculate the ratio of Employment to Total workforce, and then adjust it by Weekly Working-Hours. We assume that unemployment only occurs in the poorer group, and so we can obtain the individual labour supply of two groups. And we can also simply assume representative agents in different have the same labour supply preferences. Individual entrepreneurship time cannot be directly observed, but we do observe the Entrepreneurship Penalty Rate $\pi'_{i,t}$, i.e. $\pi'_{i,t} = \frac{\pi_t}{w_{i,t}}$ and $\pi'_t = \frac{\pi_t}{w_t}$. We can estimate the individual group $\pi'_{i,t}$ by $\pi'_{i,t} = \frac{w_t}{w_{i,t}}; \pi'_t = \frac{\mu_{i*}w_t}{\mu_i{*}w_{i,t}} = \frac{\mu_i}{\omega_{w.i}}\pi'_t$, where $\omega_{w.i} = \frac{\mu_{i*}w_t}{w_{i,t}}$ is the wage share or labour share. Lastly, we estimate individual productivities by using Solow residuals.

Figure 6: Data on individual output, capital and consumption

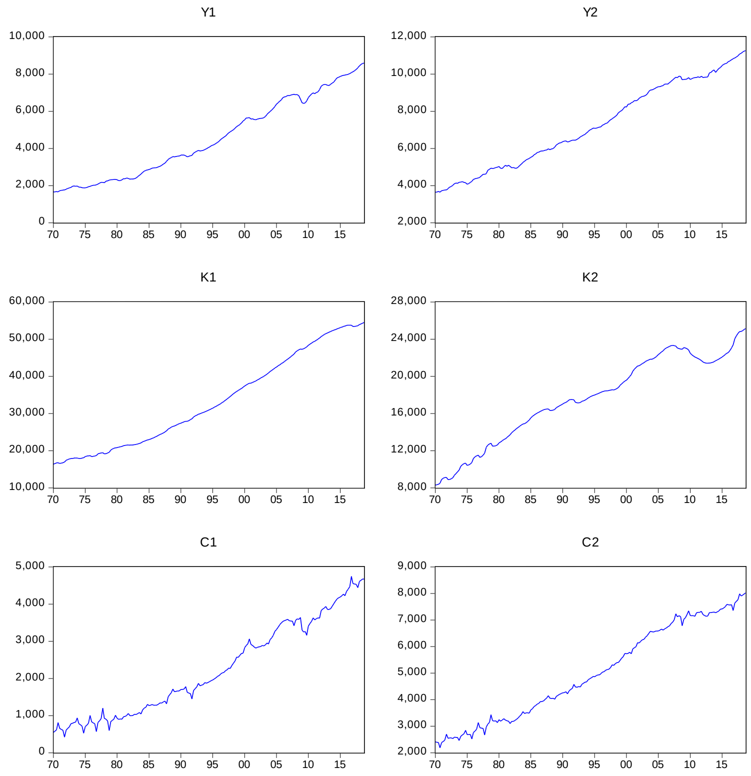


Figure 7: Estimates of individual labour and entrepreneurship penalty rate

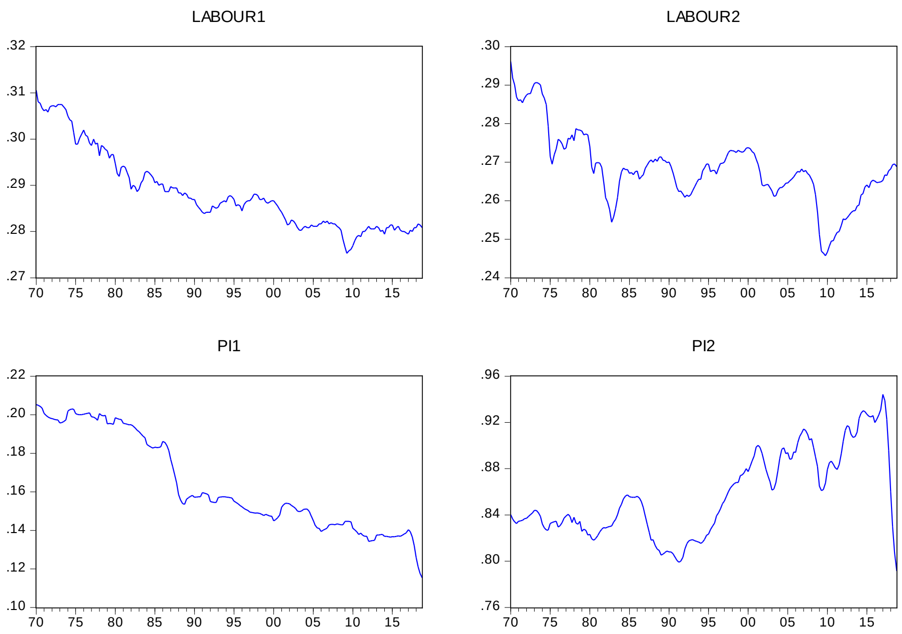
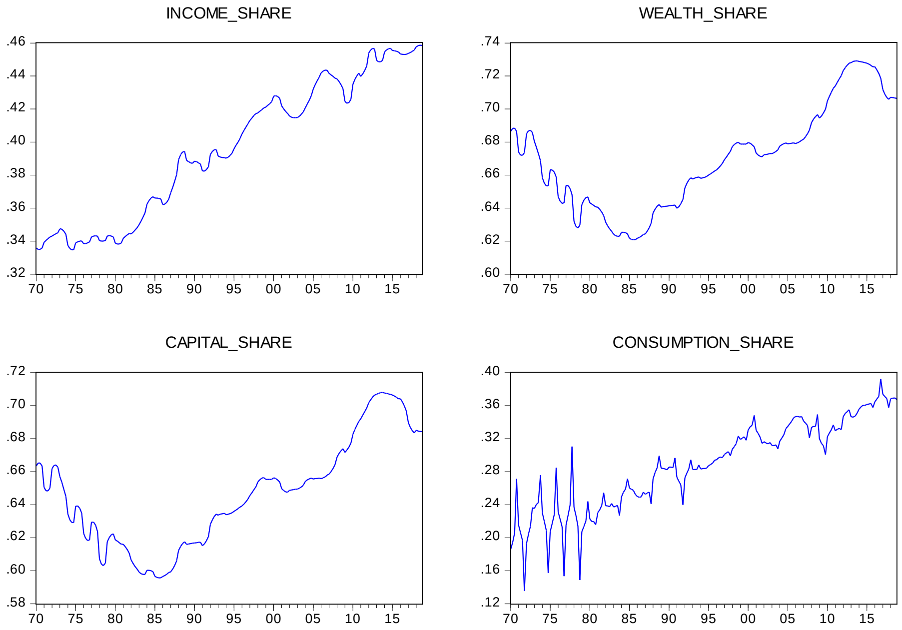


Figure 8: Some Inequality Indicators



3.3 List of linearised equations

$$r_t = \psi_1\left(E_t lnC_{2,t+1} - lnC_{2,t}\right) - ln\beta$$

$$lnY_t = \omega_{Y,1} lnY_{1,t} + \omega_{Y,2} lnY_{2,t}$$

$$lnK_t = \omega_{K,1} lnK_{1,t} + \omega_{K,2} lnK_{2,t}$$

$$lnC_t = \frac{Y}{C} lnY_t - \frac{K}{C}[lnK_t - (1-\delta)K_{t-1}] + \varepsilon_t^C$$

$$lnY_{1,t} = \alpha lnK_{1,t-1} + (1-\alpha)lnN_{1,t} + lnA_{1,t}$$

$$lnY_{2,t} = \alpha lnK_{2,t-1} + (1-\alpha)lnN_{2,t} + lnA_{2,t}$$

$$lnK_{1,t} = \ ln\beta + lnY_{1,t} - \frac{1}{\alpha}\frac{K}{C} r_t$$

$$lnK_{2,t} = \ ln\beta + lnY_{2,t} - \frac{1}{\alpha}\frac{K}{C} r_t$$

$$lnC_{1,t} = E_t lnC_{1,t+1} - \frac{1}{\psi_1}(r_t + ln\beta) + \epsilon_t^{C_1}$$

$$lnC_{2,t} = \frac{1}{\omega_{C,2}}\left(lnC_t - \omega_{C,1} lnC_{1,t}\right) + \epsilon_t^{C_2}$$

$$lnN_{1,t} = \frac{lnY_{1,t} - \psi_1 lnC_{1,t} + 2\,\frac{\psi_2 \phi_{2,1}}{\theta_2}\pi'{}_{1,t}}{1+\psi_2} - \frac{2\psi_2(1-\theta_1+\phi_{1,1})}{\theta_2(1+\psi_2)} + \epsilon_t^{N_1}$$

$$lnN_{2,t} = \frac{lnY_{2,t} - \psi_1 lnC_{2,t} + 2\,\frac{\psi_2 \phi_{2,2}}{\theta_2}\pi'{}_{2,t}}{1+\psi_2} - \frac{2\psi_2(1-\theta_1+\phi_{1,2})}{\theta_2(1+\psi_2)} + \epsilon_t^{N_2}$$

$$\ln A_{1,t+1} - lnA_{1,t} = \phi_{1,1} - \phi_{2,1}\pi'_{1,t} + lnY_{1,t} - lnw_{1,t} + \varepsilon_{a,t}$$

$$\ln A_{2,t+1} - lnA_{2,t} = \phi_{1,2} - \phi_{2,2}\pi'_{2,t} + lnY_{2,t} - lnw_{2,t} + \varepsilon_{a,t}$$

$$\pi'_{1,t} = \rho_0^{\pi} + \rho_1^{\pi}\pi'_{1,t-1} - \rho_2 * Q\left(\frac{K_{1,t-2}}{K_{t-2}}\right) + \varepsilon_t^{\pi}$$

$$\pi'_{2,t} = \rho_0^{\pi} + \rho_1^{\pi}\pi'_{2,t-1} - \rho_2 * Q\left(\frac{K_{2,t-2}}{K_{t-2}}\right) + \varepsilon_t^{\pi}$$

Chapter 4: Indirect Inference method and estimated results

4.1 Indirect Inference Method

We use the approach of Indirect Inference proposed by Le et al. (2011) to assess the model's ability to match the data. Le et al. (2016b) provide a full description of the procedure. To generate a description of the data against which the theory's performance is indirectly evaluated, the approach uses an auxiliary model that is fully independent of the theoretical one. The estimated parameters of the auxiliary model, or functions of these, might be used to summarise such a description; we name them as the 'data descriptors'. While they are viewed as 'reality,' the theoretical model under consideration is simulated to determine its suggested values. In estimation the parameters of the structural model are chosen so that when the model is simulated, the auxiliary model estimates are similar to those obtained from the real data. The structural model parameters that minimise the distance between a given function of the two sets of estimated coefficients of the auxiliary model are the best.

When evaluating the model's data fit, the structural model is simulated, and the auxiliary model is fitted to each set of simulated data, yielding a sample distribution of the auxiliary model's coefficients. A Wald statistic is computed to see if functions of the auxiliary model's parameters calculated on actual data fall within a confidence interval given by the sampling distribution.

The auxiliary model should be a process that describes how the data evolves under any applicable model. It is well known that the reduced form of a macro model with non-stationary data is a VARMA, in which non-stationary forcing factors are used as conditioning variables to accomplish cointegration (i.e. ensuring that the stochastic trends in the endogenous vector are picked up so that the errors in the VAR are stationary). This in turn can be approximated as a VECM. As an auxiliary model, we utilise a VECM with a temporal trend and the productivity residual inserted as an exogenous non-stationary process, which we re-express as a VAR(1) for the three macro variables of interest (interest rate, output, and inflation). The two exogenous elements have the effect of achieving cointegration. The VAR coefficients on the lagged dependent variables and the VAR error variances are treated as the data descriptors, and the Wald statistic is computed from them. Thus, we are essentially determining whether the observed dynamics and volatility of the selected variables can be explained by the simulated joint distribution of these variables at a given confidence level. The Wald statistic is given by:

$$W = (\hat{\theta} - \overline{\theta(\beta)})'\Omega(\beta)^{-1}(\hat{\theta} - \overline{\theta(\beta)})$$

where $\hat{\theta}$ is the vector of VAR estimates of the chosen descriptors yielded in each simulation, with $\overline{\theta(\beta)}$ and $\Omega(\beta)^{-1}$ representing the corresponding sample means and variance-covariance matrix of these calculated across simulations, respectively.

The joint distribution of the Φ is obtained by bootstrapping the innovations implied by the data and the theoretical model; it is thus a small sample distribution estimate. For small samples, this distribution is usually more accurate than the asymptotic distribution.

This testing procedure is applied to a set of (structural) parameters put forward as the true ones (H_0, the null hypothesis). The test then asks: could these coefficients within this model structure be the true (numerical) model generating the data? We extend our procedure by a further search algorithm, in which we seek other coefficient sets that minimise the Wald test statistic --- in doing this we are carrying out indirect estimation.

Thus we calculate the minimum-value Wald statistic using a powerful algorithm based on Simulated Annealing (SA), in which the search takes place over a large range around the initial values, with the search being optimised by random jumps around the space. The benefit of this extended method is that when we ultimately compare model compatibility with data, we use the best possible version of the model.

Alternative methods are available for estimation of the structural DSGE model here, notably FIML and Bayesian ML. Le et al (2016b) show, using Monte Carlo methods, that in small samples such as the one here, FIML is badly biased and tests based on it have weak power, whereas Indirect Inference gives a small bias and tests based on it have high power. As for Bayesian methods, they work well when the priors are accepted to be true, since they push the estimated parameters close to the priors. However, when there is controversy over the model parameters as here, the priors may bias the estimates away from the unknown true model; Indirect Inference estimates the model parameters with low bias around the unknown true model. Accordingly, it is the method best used here.

The steps below outline the framework involved in implementing the Wald test by bootstrapping:

Step 1: Calculating the shock process.

The first step is to back out the structural errors from the observed data and parameters of the model. We count the amount of independent structural errors as less than or equal to the number of endogenous variables. These structural errors can therefore be calculated as the difference between the LHS value (actual data) and the RHS value. For models without expectation variables the errors are calculated by taking the RHS from the LHS. For models that exhibit expectations variables we need to estimate a VAR of all the expected variables and use this to calculate the expectations following the robust instrumental variable methods of McCallum (1976) and Wickens (1982). Specifically, the VAR process will be utilised to estimate and generate fitted values one period ahead of expectations, the residuals are then calculated by the differences between LHS and RHS of the equation. We compute the corresponding coefficients and innovation of the shock process by OLS regression on the generated residuals series.

Step 2: Generate the simulated data by bootstrapping

Once the simulated disturbances are drawn from the structural errors the simulated data is then generated by a bootstrapping procedure that involves randomly drawing from the set of i.i.d innovations with replacement and solving via a project method.[4] This random selection and replacement process preserves any simultaneity between each disturbance.[5] Once the model has been solved the process is repeated N times[6], drawing each sample independently. The specific process is detailed as follows:

Step 2.1: In the first step, an initial vector of shocks is drawn and inserted into the models' 'base run', the model is then solved via the projection method with the solution becoming the lagged variable vector for the next period $t = 2$.

Step 2.2: After replacement of the initial set of innovations, the second vector of shocks is then drawn and becomes the solution for the first period $t = 1$. The model is then solved for period $t = 2$ and this in turn becomes the lagged variable vector for the next period $t = 3$.

Step 2.3: After replacement of the $n\ th$ set of initial innovations, the N^{th} vector of shocks is drawn and becomes the solution for the time period $t - 1$. The model is then solved for t and becomes the lagged variable vector for $t + 1$.

The process is then repeated for the full sample size N. And the deviations between the data in the simulation and in the original data-set are estimated in order to obtain the effects of these bootstrapped innovations. In the final step the procedure involves adding back the effects of the deterministic trends on the effects of the shocks and estimating the auxiliary model on all pseudo-samples. The full sample size of simulated data and the actual data must be consistent.

[4] The method described here follows Minford et al. (1984, 1986) and is similar to that of Fair and Taylor (1983).
[5] This process assumes that the structural errors are generated by an autoregressive process rather than being serially independent. If the structural errors are correlated with prior values then they need to be estimated.
[6] In this model $N = 1000$

Step 3: Compute the Wald statistic

Under the null hypothesis, the true economic model is the structural model with the given estimates. Deciding whether to reject or not reject the null hypothesis requires the estimation of the auxiliary model with simulated data. Here, a Wald test statistic is chosen to be the test statistic. One can apply the OLS estimates to the auxiliary model and compute both parameter vector from the actual data and the set of parameter vectors of pseudo samples and to obtain their distribution, from which one obtain corresponding estimated coefficient $\hat{\theta}$ and $\theta_s(\beta)$, respectively, where define $\overline{\theta(\beta)}$ as the average value that is computed from:

$$\overline{\theta(\beta)} = \frac{1}{1000}\sum_{s=1}^{1000} \theta_s(\beta)$$

The Wald statistic is to choose a suitable metric for measuring the distance between two set of parameters and the formula is specified as:

$$W = (\hat{\theta} - \overline{\theta(\beta)})'\Omega(\beta)^{-1}(\hat{\theta} - \overline{\theta(\beta)})$$

where $\Omega(\beta)$ the variance and covariance matrix of $(\theta_s(\beta) - \overline{\theta(\beta)})$. This process measures the distance that the actual estimated paremeters are from the average of the simulated ones. The following step is to access the combinations of all estimated coefficient the model can fit. For the model to fit the data at the 95% confidence level, it requires the Wald statistic for the actual data to be less than the 95% confidence level of the Wald statistics from the simulated data. One can present a straightforward statistic by either a P-value or transforming the Wald result a normalised t-statistic.

4.1.1 Handling Non-Stationary Data

Data used and generated by DSGE models are often non-stationary with at least part of their movement each quarter being random, this is largely the cause for uncertainty surrounding forecasting and modelling the economy's long-term future. Since Whittaker (1922), methods of data filtering or data smoothing have been designed in an attempt to remove the potential effect of such measurement error and reveal the underlying trend in the data. However, it is well known that using filtered data may distort the dynamic properties of the model in undesirable ways. In fact, many statistical filters such as the popular Hodrick-Prescott Filter (HP), can be represented as a symmetric two-sided moving average of the raw data. This alters the lag dynamic structure, generating cycles where possibly none exist.

As noted in Meenagh et al. (2012), this could have serious implications in the estimation process of a DSGE model, where both the expectations structure and the impulse response functions are usually matters of considerable interest. In a study by Cogly and Nason (1995), the authors show that the HP filter when applied to difference stationary series, is likely to generate a spurious cyclical structure at business cycle frequencies.

Another common tool employed throughout the DSGE literature is the Band Pass filter (BP), Canova (2014) however points that such filtering mechanisms only roughly capture the power of the spectrum at certain frequencies in small samples while taking growth rates greatly amplifies the high frequency content of the data, such side effects may result in containment errors that taint final estimates. Alternatively, one could map the data to be stationary by detrending the time series data. However, this process involves the use of linear detrending or first differencing. Canova (1998) shows that transforming data in this manner prior to estimation, does not maintain the inherent fluctuations with the same periodicity. In-fact, first differencing often magnifies the high frequency noise component in data. Andrle (2008) also concluded that detrending data could not explain the 95 movements of data, particularly when the permanent shock has a significant impact on the business cycle.

In addition to this, the data generated by DSGE models can also be non-stationary. The reasons for the presence of non-stationarity vary, this could be because the model structure causes non-stationarity, for example by making state variables functions of predetermined variables that depend on accumulated shocks or because shocks are permanent as is commonly assumed in DSGE models for productivity (real) and money supply shocks (nominal), as is the case in in this model. Given the ambiguity of processing non-stationary data, the case for the potential preserving effects of using unfiltered non-stationary data becomes stronger. I follow the methods developed by Le et al (2011) and later extended in Meenagh et al (2012) in which we bypass the issue of non-stationarity by use of a VARX as an auxiliary model.

As mentioned earlier, the state-space representation of a log-linearised DSGE model can be represented as a (VARMA) with some restrictions (Wickens, 2014). It can then be approximated by a finite order reduced form (VAR) model. A levels VAR can be used if the shocks are stationary, but a VECM may be needed if the data generated by the model is nonstationary. For example, if productivity shocks are permanent, then the production function is not cointegrated and as a result the associated VAR representation in levels would have nonstationary disturbances. Meenagh et al (2012) show that a VECM can be used as an auxiliary model if the shocks or exogenous processes are non-stationary. For example, if there are unobservable non-stationary variables, such as a money supply shock, then the number of cointegrating vectors will be less than the number of endogenous variables. As we have estimates of all of the coefficients of the model, we can therefore construct 96 residuals from the data. Treating these residuals as observable variables, we would then have as many cointegrating relations as endogenous variables. This would allow us to represent the solution of the estimated model as a VECM in which the non- stationary residuals appear as observable variables, we can then use an unrestricted version of this VECM as our auxiliary model. Following the developments of Meenagh et al. (2012) and Le et al. (2015a), I demonstrate that the chosen auxiliary model is an approximation of the reduced form of the DSGE model when under the null hypothesis of cointegration and can be represented as a cointegrated VARX.

4.1.2 The Choice of Auxiliary Model

The II test criterion is determined by the difference between the empirical auxiliary Wald statistic from the observed data and the simulated auxiliary Wald statistic from the simulated data. Those parameters (θ) of an auxiliary model may be not an accurate description of the data generating process, but they can be estimated easily by conventional estimation methods. Therefore, there is no simple rule to identify the best auxiliary model. A natural choice of auxiliary model is an unrestricted VAR, because a VAR is the reduced form of a DSGE model, however Minford et al. (2016) test if there are more powerful choices for the auxiliary model or 'data descriptors' and compare the power against auxiliary models derived from Impulse Response Functions and the Simulated Moments Method. Evaluating the power of these different methods in small samples using Monte Carlo simulations, they find that in a small macro model there is no difference in power, however in large complex macro models the power with Moments rises more slowly with increasing misspecification relative to the other two which remain similar. The greater the power the less the range of uncertainty about how wrong their models could be. These findings suggest that VAR coefficients and average IRFs are more or less interchangeable for this purpose; but that Moments give less power in testing large complex macroeconomic models. When VAR coefficients are used as the data descriptors, the estimated parameters are used to describe the dynamic property of the data whilst the variance of the errors are used to capture data volatility. With IRFs served as the data descriptions, the IRF can be transferred as a nonlinear combination of VAR coefficients and the error covariance matrix[7].

Le et al (2016) also show that the DSGE models we are examining may be over-identified, therefore the addition of more VAR coefficients by raising the order of the VAR can increase the power of the test. Analogously, adding more elements to the IRF descriptors or to the moment descriptors should do the same. However, Le et al (2016) also points that increasing the power in this way also reduced the probability of finding a tractable model that would pass the test, an inherent trade-off between power and tractability. Additionally, empirical results in Le et al. (2011, 2015, 2016) show that when including a broader set of endogenous variables in the auxiliary model, it usually results in a strong rejection. Le et al. (2015) points out that the power of the full Wald test increases as more endogenous variables are added. This is also true when the lag order is raised and can lead to uniform rejections. Meenagh et al (2012) also argues that such attempts usually lead to rejection when the model in question appears to share with too many elaborate structures.

Based on the above information, I employ a VARX(1) as the auxiliary model for model simulation and estimation and choose output Y_t, capital K_t and capital inequality IQ_t as the key variables from which to base the Wald test. Since these three variables can represent a general inner relationship of the model as well as describe the economy in full. A VARX(1) with three endogenous variables takes the form of:

$$\begin{bmatrix} Y_t \\ K_t \\ IQ_t \end{bmatrix} = B \begin{bmatrix} Y_{t-1} \\ K_{t-1} \\ IQ_{t-1} \end{bmatrix} + CX_t + \varepsilon_t$$

[7] Minford et al. (2016) show that the error from a VAR model can be specified as a: et = Bvt , where vt are the structural innovations and B denotes the error covariance matrix.

where $B = \begin{bmatrix} b_{yy} & b_{yk} & b_{yiq} \\ b_{ky} & b_{kk} & b_{kiq} \\ b_{iqy} & b_{iqk} & b_{iqiq} \end{bmatrix}$ is a 3 by 3 coefficient matrix. $X_t = \begin{bmatrix} Trend \\ A_{1,t-1} \\ A_{2,t-1} \end{bmatrix}$ contains the two exogenous non-stationary productivity shock $(A_{1,t-1}, A_{2,t-1})$ and time trend (Trend), $\varepsilon_t = \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{kt} \\ \varepsilon_{iqt} \end{bmatrix}$ denotes error vector.

The auxiliary coefficient vector to compute the Wald statistic take account into B which contains the 9 elements and the 3 variances of the VARX regression residuals: VAR(ε_t). Given the actual data, the estimated B and VAR(ε_t) are shown below.

Table 2 Coefficients of the auxiliary model; and those simulated by the estimated model of chapter 4

Auxiliary coefficients	**Actual**	**Simulated**		
		Mean	**2.5% Lower**	**97.5% Upper**
b_{yy}	0.712	0.514	0.077	1.112
b_{yk}	-0.492	-0.449	-1.049	0.102
b_{yiq}	0.267	0.206	-0.017	0.446
b_{ky}	0.113	0.278	-0.041	0.669
b_{kk}	0.901	0.602	0.116	1.093
b_{kiq}	0.154	0.192	-0.138	0.519
b_{iqy}	-0.029	-0.133	-0.268	-0.018
b_{iqk}	0.036	0.149	0.029	0.279
b_{iqiq}	1.005	0.976	0.364	1.556
$Var(e_y)$	2.316×10^{-5}	0.003	0.002	0.004
$Var(e_k)$	6.941×10^{-7}	0.004	0.003	0.005
$Var(e_{iq})$	7.951×10^{-6}	1.218×10^{-4}	8.941×10^{-5}	1.459×10^{-4}

4.2 Estimated results

4.2.1 Calibration

Shown in Table 3, the capital share in production α is set to 0.386, which is the average of annual labour share in US from 1970 to 2018 reported by FRED economic date. Utility discount factor β is set to 0.97, which is in order to math the average US risk-free rate 0.03. The capital depreciation rate is set to 0.045, which is based to depreciation of Business capital for United States. Share of consumption preference in CRRA utility is set to 0.5, which reflect that individual has an idential weights of consumptions and leisure in the utility function. The ratios of steady-state are calculated from the average ratioin the priod of 1970 to 2018.

Table 3 Calibrated parameter values

Description	Coefficie nt.	Yang et al (2021) UK model	US model
Share of capital in Cobb-Douglas Production	α	0.300	0.386
Utility Discount rate	β	0.997	0.970
Capital depreciation rate	δ	0.003	0.045
Share of consumption preference in CRRA utility	Φ	0.500	0.500
Steady state output share by 10% income decile	$\omega_{y,1}$	0.380	0.406
Steady state capital share by 10% income decile	$\omega_{k,1}$	0.640	0.674
Steady state consumption share by 10% income decile	$\omega_{c,1}$	0.200	0.302
Steady state aggregate output/consumption ratio	Y/C	1.710	1.535
Steady state aggregate capital/consumption ratio	K/C	2.696	6.865

4.1.2 Estimation and The properties of residuals

Here we exanine whether the model could minic main characterisitcs and tendency of economic growth and distributions of capital and income in rencet decades in US. Our model contains 10 shocks. The Table below shows what we used for simulating the model. We also show charts of the residuals and innovations. Our assumptions based on theory are that the productivity shocks are non-stationary but that the others are all stationary. Each residual is fitted with an AR process by OLS, with the result shown in the Table 4. The indirect inference procedure tests these assumptions jointly with the other parameter assumptions based on theory.

Table 4 AR coefficients of structural errors

Shock Name	Shock Process	AR (1) coefficient	Stdardard Error	P-value
Demand	Stationary	0.012	0.008	0.0000
Capital	Stationary	0.001	0.004	0.0000
Consumption (aggregate)	Stationary	0.053	0.031	0.0000
Consumption (rich)	Stationary	0.926	0.070	0.0030
Consumption (poor)	Stationary	0.229	0.067	0.0000
Labour supply (rich)	Stationary	0.047	0.045	0.0000
Labour supply (Poor)	Stationary	0.052	0.041	0.0000
Entrepreneurship Penalty shock	Stationary	0.937	0.046	0.0000
Productivity (rich)	Non-Stationary	1.000	0.060	0.0000
Productivity (poor)	Non-Stationary	1.000	0.067	0.0000

It should be noted that the two individual productivity shock $ln(A_{i,t+1})$ follow non-stationary processes as in many stochastic growth models, but we give it an endogenous growth following the process below which relies on the entrepreneurship time $\pi_{i,t}$:

$$ln(A_{i,t+1}) = ln(A_{i,t}) + \theta_1 + \theta_2 \pi_{i,t} + v_{A,t}$$

where θ_1 measures the natural productivity growth and the $v_{A,t}$ is the innovation from the aggregate productivity shock so that both individual productivities are related to aggregate productivity is exogenous. Other structural shocks follow an AR(1) process, and coefficients are obtained from OLS estimation. The model backed structural shocks and estimated innovations from OLS are shown in figure 9 and figure 10.

We then apply stationarity tests to analyse the properties of residuals and then describe the basic characteristics of shocks and residuals. We apply three types of stationarity test on the residuals from shocks, i.e. Augmented Dickey-Fuller Test, and Phillips-Perron Test. The null hypothesis of ADF and PP tests that series has unit roots. Table 4 reports the results of the ADF test. For the stationarity of ten series, eight series tend to be stationary at 99% significance level after calculating fist difference, while the other series are stationary at 95% significance levels. Table 5 reports the PP test, which shows all series tend to be stationary at 99% first difference stationary level. From these partly conflicting results, we estimate the model with all the schocks stationary; all are estimated as AR processes. The Table below shows what we used for simulating the model. We also show charts of the residuals and innovations above.

Figure 9 Model shocks

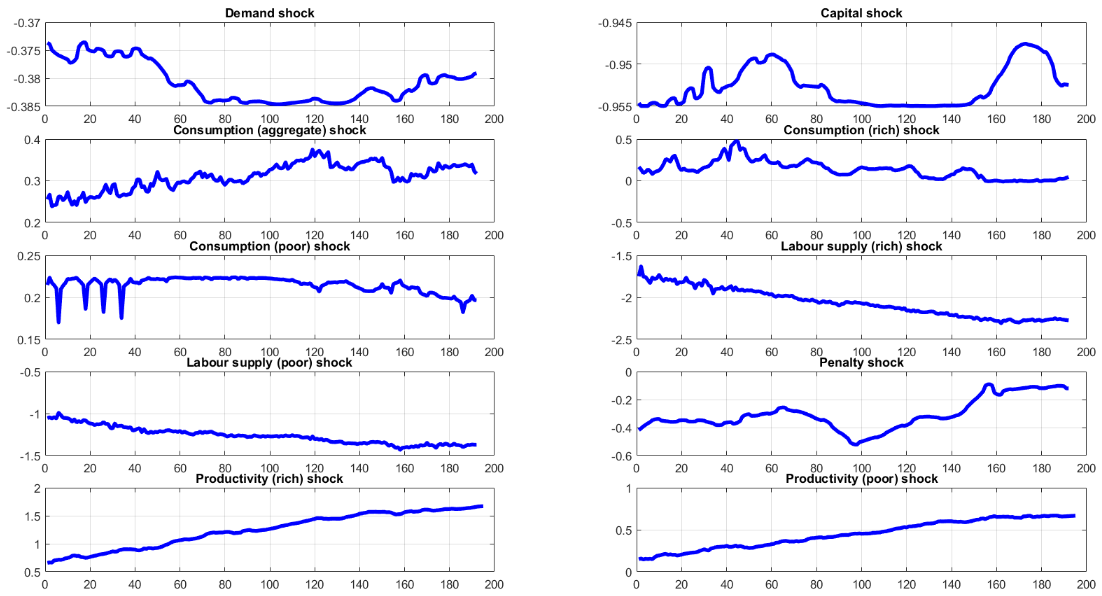


Figure 10 Model innovations

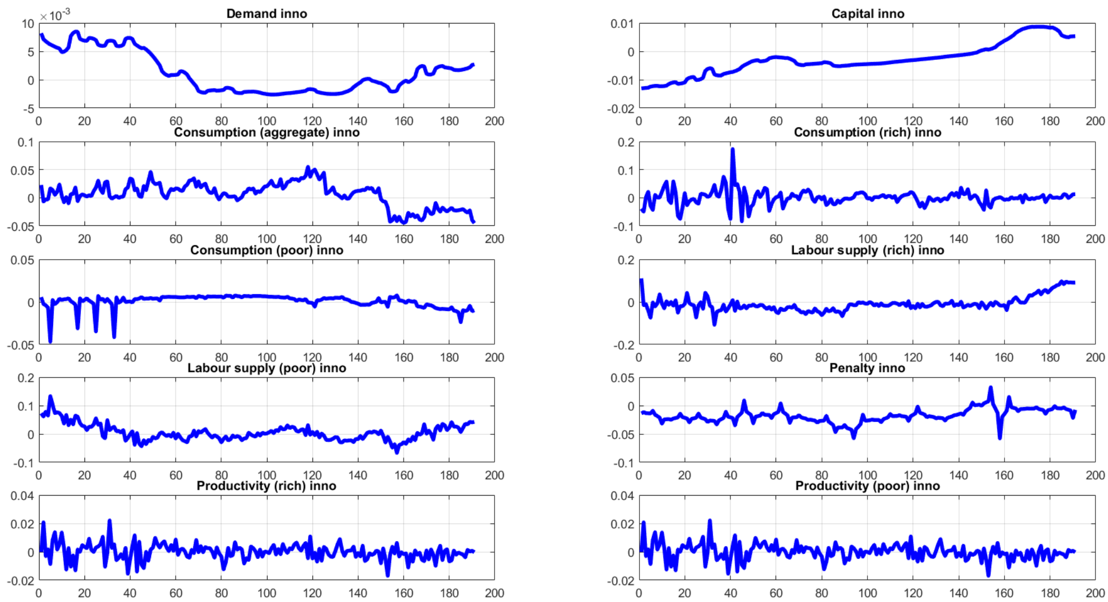


Table 5. Augmented Dickey-Fuller(ADF) Test results

	Level		First Difference	
	Intercept	Intercept and trend	Intercept	Intercept and trend
Demand Shock	-2.507991 (0.1153)	-0.906237 (0.9520)	-3.104479** (0.0280)	-4.162496* (0.0062)
Capital shock	-1.321891 (0.6192)	-2.333100 (0.4136)	-3.709358* (0.0047)	-3.719158** (0.0234)
Consumption Aggregation shock	-2.590267 (0.0967)	-3.205456*** (0.0865)	-12.87114* (0.000)	-12.91824* (0.0000)
Consumption shock for the rich	-11.10322* (0.0000)	-11.08473* (0.0000)	-13.58029* (0.0000)	-13.54371* (0.0000)
Consumption shock for the poor	-0.187983 (0.9364)	-0.698168 (0.9711)	-7.225609* (0.0000)	-7.630433* (0.0000)
Labour input shock for the rich	-0.659532 (0.9910)	-0.558722 (0.9799)	-11.27293* (0.0000)	-11.56284* (0.0000)
Labour input shock for the poor	-2.895683** (0.0477)	-2.629458* (0.2678)	-14.10067* (0.0000)	-14.24130* (0.0000)
Aggregate Penalty shock	-1.668848 (0.4454)	-2.325987 (0.4174)	-7.155462* (0.0000)	-7.150809* (0.0000)

Table 6. Phillips-Perron (PP) Test results

	Level		First Difference	
	Intercept	Intercept and trend	Intercept	Intercept and trend
Demand Shock	-1.957461 (0.3056)	-0.884207 (0.9546)	-8.323777* (0.0000)	-7.771870* (0.0000)
Capital shock	-1.096962 (0.7169)	-1.426138 (0.8504)	-7.483800* (0.0000)	-7.501588* (0.0000)
Consumption Aggregation shock	-2.189687 (0.2108)	-2.889268 (0.1684)	-17.31947* (0.0000)	-17.39355* (0.0000)
Consumption shock for the rich	-11.01242* (0.0000)	-10.99125* (0.0000)	-57.60480* (0.0001)	-58.17514* (0.0001)
Consumption shock for the poor	-10.03737* (0.0000)	-10.15297* (0.0000)	-97.47446* (0.0001)	-122.9067* (0.0001)
Labour input shock for the rich	-5.072370* (0.0000)	-6.952015* (0.0000)	-32.77509* (0.0000)	-48.76352* (0.0001)
Labour input shock for the poor	-3.942854* (0.0021)	-3.938086* (0.0123)	-26.25348* (0.0000)	-29.05033* (0.0000)
Aggregate Penalty shock	-6.112130* (0.0000)	-6.416844* (0.0000)	-31.72838* (0.0001)	-32.98655* (0.0001)

Note: Statistic with ***, ** and * indicate a rejection of the unit root process at 10%, 5% and 1% significant level respectively. Values in the parentheses are p-values.

Table 7 Indirect inference estimated parameter values

	Coefficient	Yang et al (2021) UK model	US model
Elasticity consumption in CRRA utility	Ψ_1	1.000	0.348
Elasticity of labour in CRRA utility	Ψ_2	1.000	0.445
Marginal effect of entrepreneurship time on individual productivity growth	θ_2	0.510	0.398
Marginal effect of capital term Q on individual opportunity cost of entrepreneurship	ρ_2^{π}	0.001	0.029
Marginal effect of penalty rate on productivity growth for the rich	$\phi_{2,1}$	0.548	0.307
Marginal effect of penalty rate on productivity growth for the poor	$\phi_{2,2}$	0.220	0.169

The model as estimated did not fit the data behaviour at the 5% probability level; only at the 2% level (with 12 degrees of freedom, being the number of auxiliary model parameters). I could not find a parameter set that gave a closer match. However, this is still a reasonable match, and, as it is the closest I could find, I use it to explore the implications of the model for the effects of different policies and shocks.

Table 7 Indirect inference Wald test result

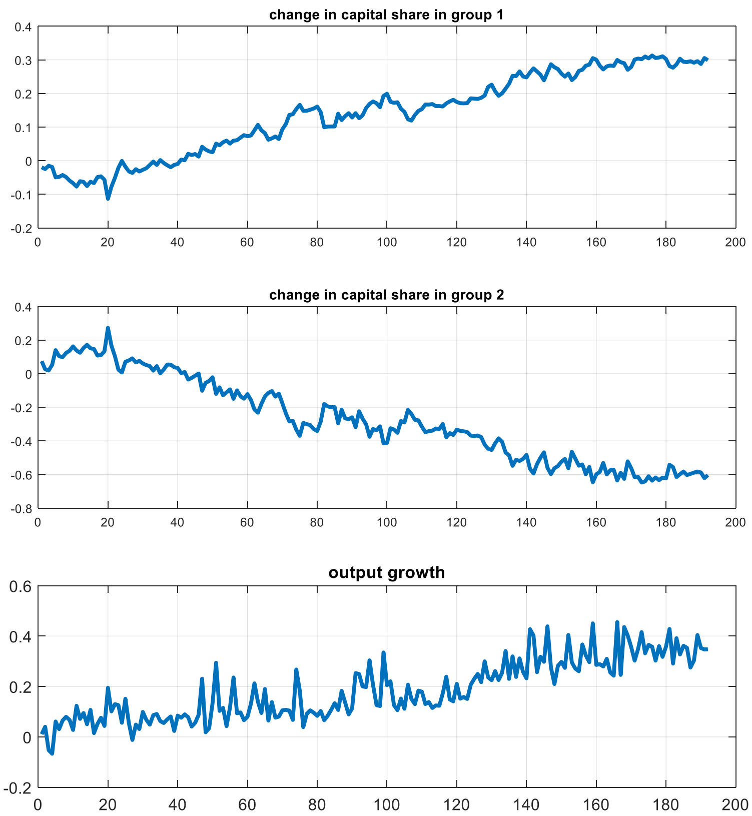
Description	Number of Simulations	Transformed MD (t-stat)	P-value
Y_t, K_t, IQ_t	100	2.27	2.3%

The estimated model's behaviour

I now turn to a discussion of how the model behaves in response to shocks. My first question concerns the effects of all the shocks in the model on the longer term behaviour of growth and inequality. This can tell us what to expect in the way of US long term trends in these two key variables.

To simulate efforts by the government to reduce inequality, we assume a transfer tax, t, on capital of the rich, refunded to the poor on their capital. Thus we reduce the capital of the rich according to K'1/K=[K1/K]x(1-t) and we raise that of the poor according to K'2/K=[K2/K] x (1+tK1/K2). We show the trade-off between average growth and average inequality over the simulation period as we raise the transfer rate, t.

Figure 11 Tendency caused by all shocks



What we see from these shock simulations of this model is a tendency for growth and inequality to rise over time steadily. In this respect the model resembles the one estimated for the UK by Yang et al (2021). The estimated parameters in the two models are similar- see Tables of estimated model parameters for those of Yang et al also.

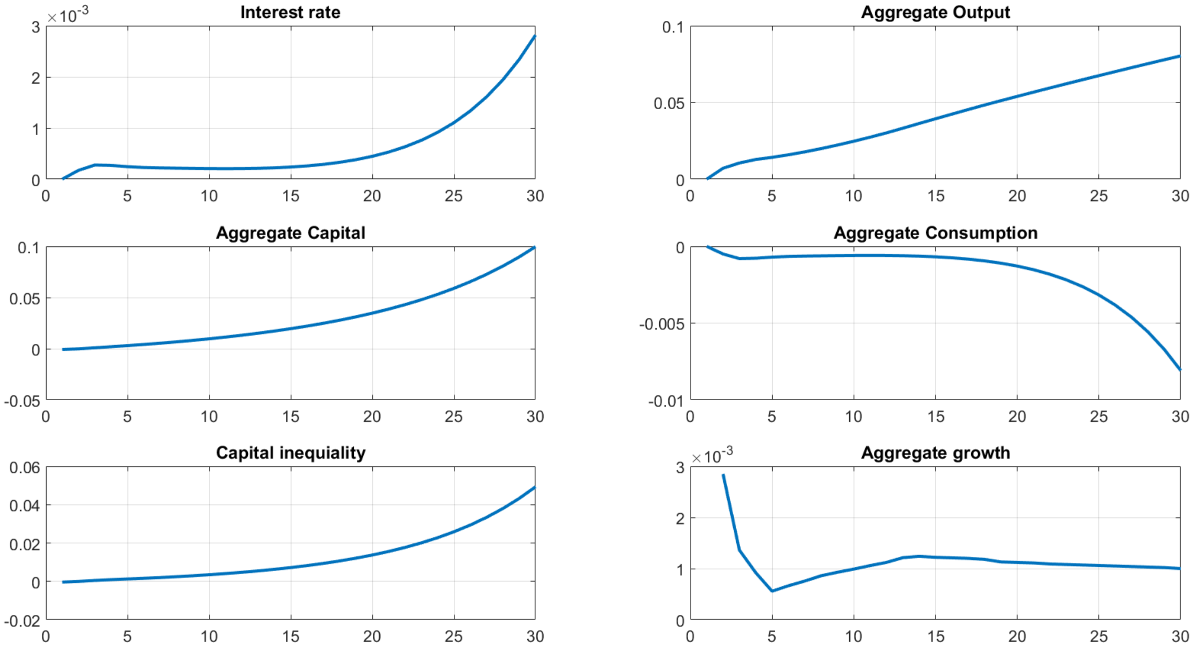
Another question I pose is which shocks contribute most to inequality and growth; this can be seen from the variance decomposition of the model. This, like the above long term tendencies, is also similar to that of Yang et al for the UK. The shocks contributing to growth are diverse, while that contributing most to capital inequality is the penalty rate (i.e. the tax/regulative costs laid on entrepreneures)

Table 8 Variance Decomposition (T=30)

	Penalty shock	Prod. Shock	Demand shock	Cons. Shock (rich)	Cons. Shock (poor)	Labour shock (rich)	Labour shock (poor)
Interest rate	42.8	27.2	4.0	4.0	0.8	15.9	5.3
Aggregate output	70.8	20.0	0.4	3.1	0.1	3.6	2.0
Aggregate capital	69.4	26.5	0.5	0.8	0.1	2.0	0.6
Aggregate cons.	78.3	5.5	3.3	7.5	3.9	1.2	0.3
Income of the rich	75.5	10.6	0.8	0.9	0.0	9.8	2.4
Income of the poor	41.7	25.4	2.9	3.8	0.7	8.0	17.5
Capital of the rich	77.4	15.9	2.1	0.6	0.0	2.9	1.1
Capital of the poor	69.8	1.2	5.4	0.2	1.3	14.5	7.7
Cons. of the rich	55.5	5.7	0.8	35.3	0.3	1.8	0.6
Cons. of the poor	80.1	5.2	4.7	4.0	0.8	1.9	3.3
Labour of the rich	72.5	10.6	0.6	0.2	0.0	14.7	1.4
Labour of the poor	53.3	19.9	1.0	2.3	0.2	2.7	20.6
Capital inequality	80.1	0.1	0.9	0.0	0.0	12.5	6.4
Growth rate	20.6	23.9	10.9	12.5	9.4	12.4	10.3

We also examine the impulse response function of the model with respect to its major policy variable, the entrepreneur tax rate- next Figure. As can be seen, a fall in the tax/regulative cost of innovation pushes up inequality, growth and output over time. The rising demand for capital raises the real interest rate, reducing consumption.

Figure 12 IRF to one standard deviation negative aggregate penalty rate shock



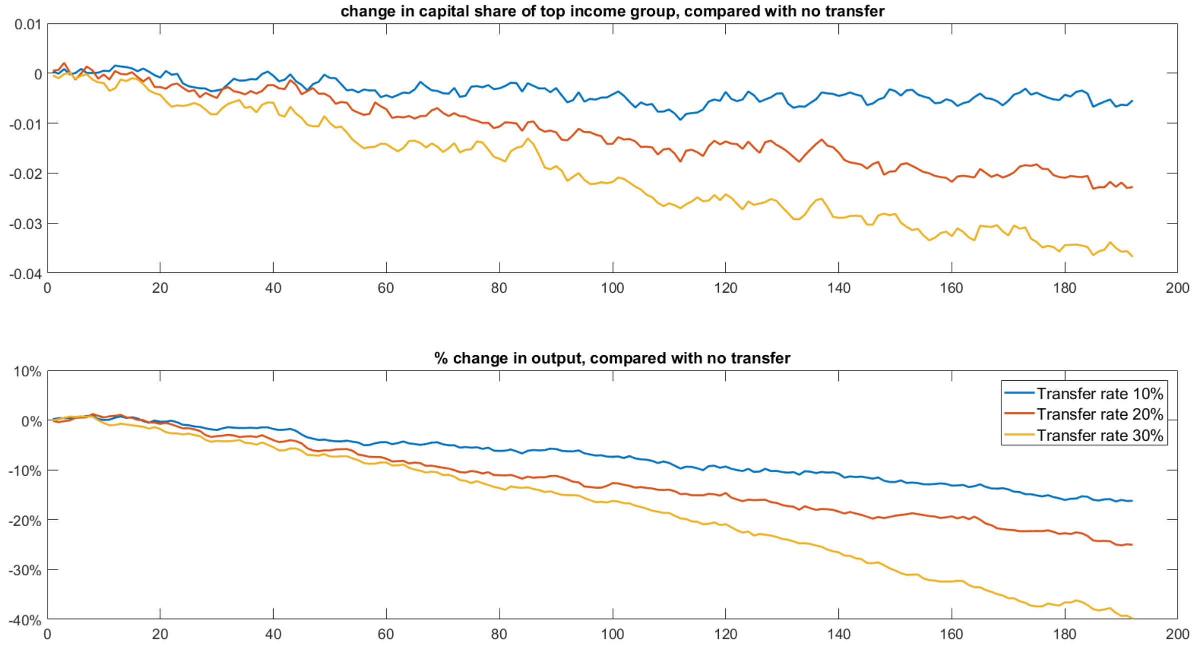
Chapter 5: The trade-off in reduced growth from redistributive policies to reduce inequality

In this chapter I ask whether government policies could reduce inequality and at what cost to growth? Clearly both limiting inequality and sustain high growth are objectives of social and government policy. However the model we have estimated here suggests they are in clear conflict. The issue I wish to explore in this final section of my thesis is the trade-off between these two objectives. I do so by adding a transfer policy to the model under which the government redistributes income from the rich group to the poor group. As the transfer rate rises, both inequality and growth fall. The question I wish to answer is by how much. If we then could establish social indifference curves, we could guide the government to an optimal transfer rate. I also investigate what light my results shed on the political economybof redistribution in the US, i.e. how tax an redistribution actually evolveddue to the political process, in the postwar US.

The realisation of the income transfer needs approximation in our model; otherwise the transferred income to the poor will not affect their individual capital accumulation and consumption because the individual budget constraints themselves are not used as model equations. I use an approximation to avoid this trap as follows.

Suppose a constant income tax rate τ_Y is enforced on the rich. The tax revenue per capita across the whole population now is $\tau_Y \mu_1 Y_{1,t}$ which is transferred to the poor at the end of period t. Since the original linearised individual capital equation can be written as $\ln\left(\frac{\mathrm{K_{i,t}}}{\mathrm{Y_{i,t}}}\right) = \exp\left[\frac{1}{\Psi_1} - \mathrm{K}/((1-\tau)\alpha \mathrm{Y})\right] \equiv \mathrm{f}(\mathrm{rt})$, individual capital per capita $\mathrm{K_{i,t}}$ after income transfer can be written as $\mathrm{K}'_{\mathrm{i,t}} = \mathrm{f}(\mathrm{rt})\mathrm{Y_{i,t}}(1-x_i)$, where $x_i$is the equivalent income tax rate for group i. Given the population weights μ_i, $x_1 = \tau_Y$ and $x_2 = -\tau_Y \frac{\mu_1 Y_{1,t}}{\mu_2 Y_{2,t}} \approx -\tau_Y \frac{\omega_{y,1}}{\omega_{y,2}}$

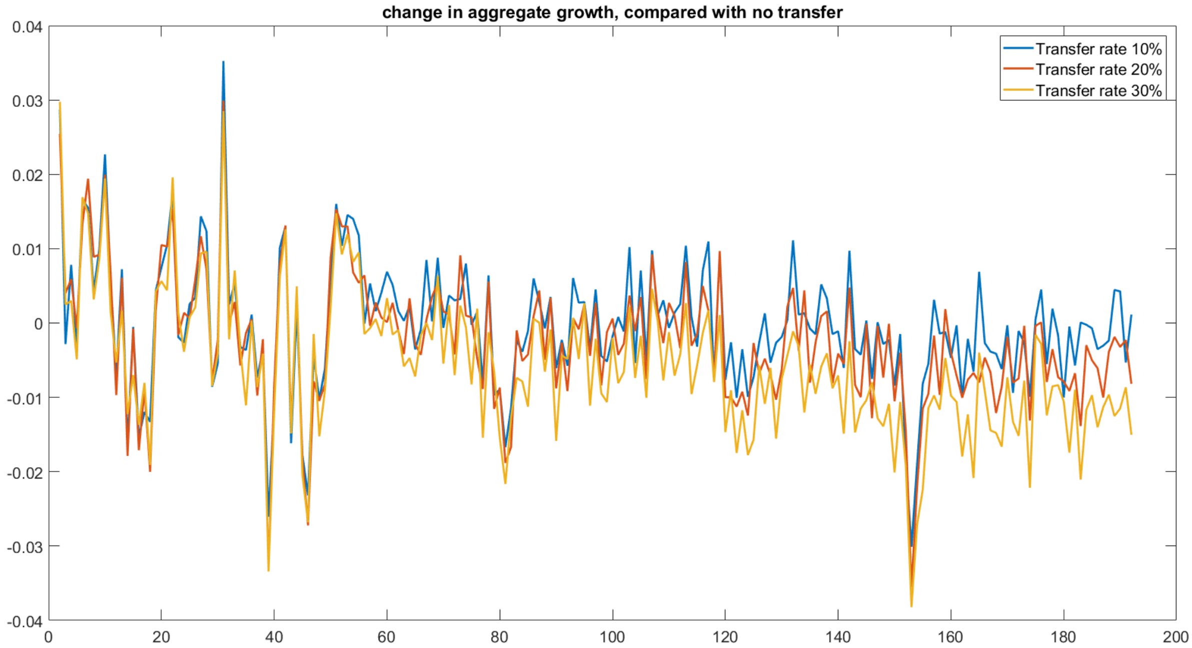
Figure 13 Redistributive effects by different income transfer rates



The results show a trade-off between growth reduction and inequality reduction in which to gain a 1 percent rise in the capital share of the bottom group sacrifices about 0.4% p.a. growth, on average, implying a cumulative loss of output of about 12% over a 30-year period. This ratio is roughly constant as the transfer rate rises. However, the marginal effect on the terminal growth rate is rising as the transfer rate rises- see next chart; hence long-term growth increasingly suffers as inequality is reduced by transfers.

These negative effects of largescale transfers on growth seem to account for why the US tax system has is quite limited in the extent of its redistribution.

Figure 14 Redistributive effects on growth by different income transfer rates:



The political economy of redistribution in the US in the light of my results

In this thesis I have found that there is evidence for my model of growth and inequality in US post-war data. As this chapter has shown, there is a trade-off between growth and inequality created by redistribution. With a policy of tax and transfer, growth is reduced while inequality is also reduced. This raises a central issue of political economy: would governments be induced to redistribute income in their competition for power?

A natural starting point for this analysis is the median voter theorem of Downs (1957) and much subsequent work (early examples are Meltzer and Richard, 1981, 1983; for later work see Alesina, Roubini and Cohen, 1997) suggesting that governments should favour voters at the centre of the income distribution in order to swing their votes to their side; it also suggests that different party platforms would converge towards this common agenda. This would suggest a tendency to favour redistributive tax/transfer policies such as have become the norm in European countries, where the tax/GDP ratio has risen to great heights in the post-war period. We can see broadly supportive evidence of this in the upward trends in taxation in most OECD countries- see Figure 15 and Table 8 below.

However, we do not observe this tendency in the US, at least to the same degree.

Figure 15 average tax/GDP ratio from 1965 to 2022

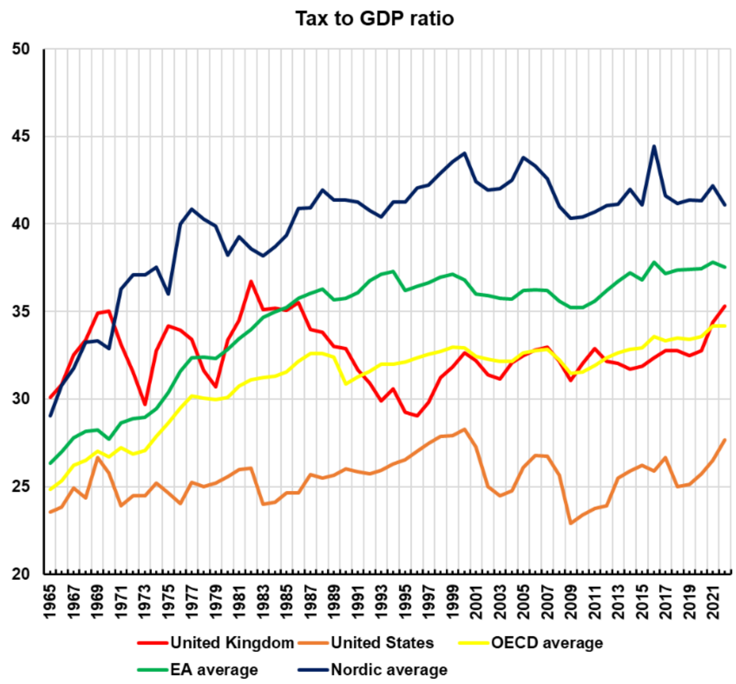


Table 9 Average Tax-GDP ratio in major OECD countries from 1965 to 2022

Country	Average Tax/GDP ratio	Country	Average Tax/GDP ratio
United States	25.5	Italy	35.4
Australia	26.5	Japan	25.1
Austria	39.8	Korea	20.0
Belgium	41.3	Latvia	29.4
Canada	32.1	Lithuania	29.6
Chile	19.8	Luxembourg	34.6
Colombia	17.7	Mexico	13.5
Costa Rica	21.8	Netherlands	37.3
Czechia	33.8	New Zealand	31.6
Denmark	43.0	Norway	39.8
Estonia	32.3	Poland	34.2
Finland	39.9	Portugal	26.8
France	40.9	Slovak Republic	32.8
Germany	35.6	Slovenia	37.8
Greece	27.8	Spain	27.9
Hungary	38.4	Sweden	43.0
Iceland	33.1	Switzerland	23.9
Ireland	29.0	United Kingdom	32.5

Yet the simulation results from my model imply that the welfare of the average household, which will broadly coincide with that of the median, is maximised by redistributive policies reducing inequality, even though this reduces growth substantially. Table 9 sets out these results.

Table 10 Simulated effects of different redistributive policies on welfare and growth

	Measures	Transfer rate 10%	20%	30%
Welfare	Accumulated utility from consumption-average household	19752.25	22410.91	23746.16
Welfare inequality	Percentage difference from when tax=10%	-	13.46%	20.2%
Growth	Accumulated output growth over simulation period*	273.80%	198.92%	168.2%
Growth inequality	Percentage difference from when tax=10%	-	-27.35%	-38.6%

* simulation period = 190

Figures 16 and 17 below depict simulations of output and consumption under different transfer rate when I use the identical structural shocks. It is evident that as transfer rates increase, growth decreases, but welfare improves.

Figure 16 Output simulations using same random shocks under different transfer rate

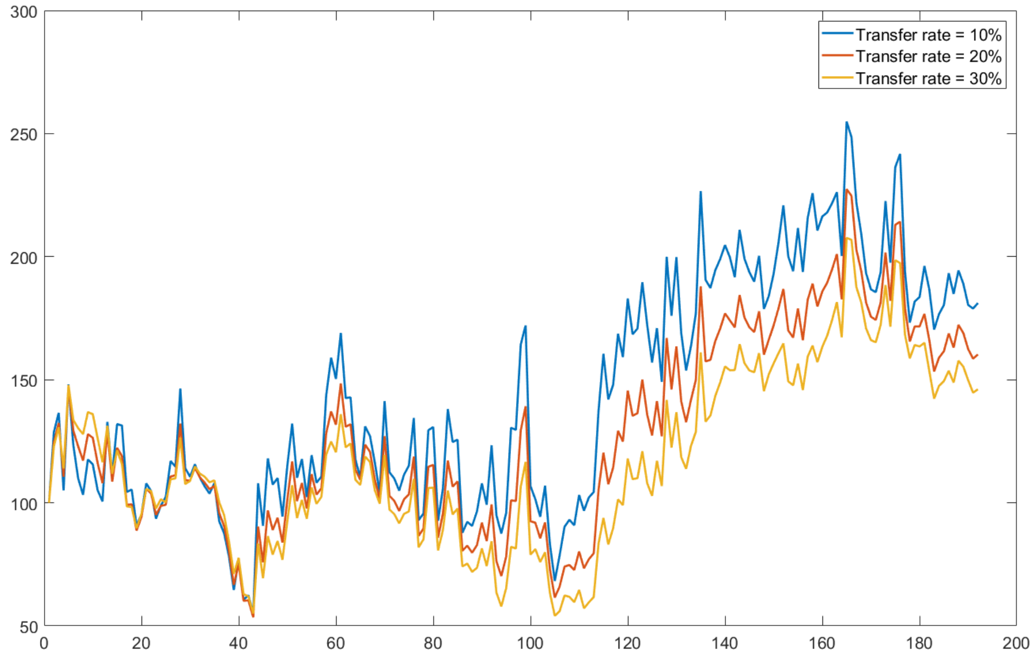
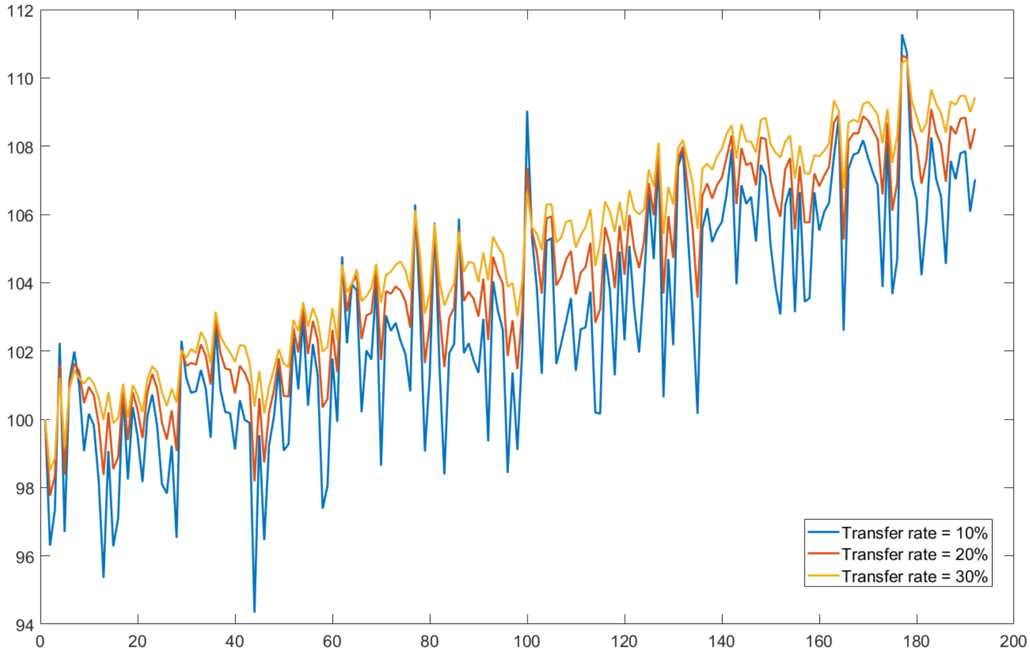


Figure 17 Consumption (average household) simulations using same random shocks under different transfer rate



The US is therefore an exception to the median voter tendency across the OECD. This is not surprising however given the checks and balances created by the US constitution, which was set up by agreement between US states in 1789 to prevent strong central power at the federal level. This reflected two main factors: the desire of states to preserve power and the fear of 'majoritarian tyranny' on the part of the founding fathers. States compete to attract business, which favours growth against equalisation of income; as for the median voter household, its power to elect the president and Congress is limited by the large weighting of states in the Presidential Electoral College and the allocation of Senate seats entirely by state.

My model results confirm that US political economy favours growth over inequality, even though the median voter household would benefit from less inequality via more redistribution. These results are in line both with the observed facts of postwar taxation around the world, in which the US has systematically low tax rates relative to other OECD countries, and with what one would expect the US Constitution to deliver.

It is of interest to compare my US results with those for the UK, found in recent work by Yang et al (2021).This paper found a similar trade-off between growth and inequality from redistribution. In a recent unpublished PhD thesis Hankui Wang (2023) found that this model estimated over annual data from 1870 implied that welfare of the average household would be raised by redistribution, so accounting for the steady rise in the UK tax rate over the period from then to modern times. In the UK political system, the median voter has considerable power, as the governing party is determined by the swing voter in a first-past-the-post voting system. Hence the theorem has good predictive power for the UK.

Conclusions

In this thesis I have put forward a model of endogenous growth in which entrepreneurial innovation by household-owned firms depends on their wealth and the extent to which their efforts are taxed and regulated. Firms are heterogeneous by wealth and the model implies that wealthier firms will innovate more than others, even if facing the same marginal tax rates and regulative constraints. I have estimated the model on post-war US data by indirect inference and found a parameter set that matches the auxiliary model data behaviour without being rejected at the 2% level. The model when repeatedly hit by the shocks identified in the estimated equations implies that growth and inequality will tend to rise steadily over time. It is to be expected that this will prompt political demands for redistributive transfers. I have examined how a rising transfer system of taxes and tax credits would impact on inequality and growth. I find that it would reduce inequality but at a substantial cost in growth, and a rising cost in the rate of long-term growth. The model therefore suggests a strong reason why US governments have built only quite limited redistribution into their tax system.

References

1. Aghion, P. and Bolton, P. (1997). A Theory of Trickle-Down Growth and Development. The Review of Economic Studies, 64(2), p.151.
2. Alesina, A. and Perotti, R. (1996). Income distribution, political instability, and investment. European Economic Review, 40(6), pp.1203-1228.
3. Alesina, A., Roubini, N. and Cohen, G. (1997) Political Cycles and The Macroeconomy, MIT Press, Cambridge, Massachusetts.
4. Algan Y, Allais O, Den Haan WJ (2008) Solving heterogeneous-agent models with parameterized crosssectional distributions. J Econ Dyn Control 32(3):875–908
5. Algan Y, Allais O, Den Haan WJ, Rendahl P (2014) Solving and simulating models with heterogeneous agents and aggregate uncertainty. In: Schmedders K, Judd KL (eds) Handbook of computational economics, vol 3. North Holland, Oxford, pp 875–908
6. Aiyagari SR (1994) Uninsured idiosyncratic risks and aggregate saving. Q J Econ 109(3):659–684
7. Aiyagari SR, GertlerM(1991) Asset returns with transactions costs and uninsured individual risk. JMonet Econ 27(3):311–331
8. Atems, B. and Jones, J. (2014). Income inequality and economic growth: a panel VAR approach. Empirical Economics, 48(4), pp.1541-1561.
9. Atkinson, A. (2015). Inequality: What Can Be Done?. 1st ed. Harvard University Press.
10. Atolia M, Prasad K (2011) Relative wealth concerns and entrepreneurship. Economica 78:294–316
11. Barro, R. (2000). Inequality and Growth in a Panel of Countries. Journal of Economic Growth, 5:5-32 (March 2000).
12. Banerjee AV, Newman AF (1993) Occupational choice and the process of development. J Polit Econ 101(2):274–298
13. Beck, T., Demirgüç-Kunt, A. and Levine, R. (2007). Finance, inequality and the poor. Journal of Economic Growth, 12(1), pp.27-49.
14. Bewley, T. F. (1977). The Permanent Income Hypothesis: A Theoretical Formulation. Journal of Economic Theory, 16(2), pp.252-295.
15. Bewley TF (1981) Stationary equilibrium. J Econ Theory 24(2):265–295
16. Bewley TF (1983) A difficulty with the optimum quantity of money. Econometrica 51(5):1485–1504
17. Bewley, T.F. (1986). Stationary Non-Convex Preference.

18. Bilbiie, F. O., Monacelli, T., & Perotti, R. (2024). Stabilization vs. redistribution: The optimal monetary–fiscal mix. Journal of Monetary Economics, 147, 103623.
19. Brei, M., Ferri, G. and Gambacorta, L. (2018). Financial structure and income inequality. BIS Working Papers No 756.
20. Boadway R, Marchand M, Pestieau P (1991) Optimal linear income taxation inmodels with occupational choice. J Public Econ 46(2):133–162
21. Bohacek R, Kejak M (2018) Optimal government policies in models with heterogeneous agents. J Econ Theory 176:834–858
22. Broer T., Krusell P., and Oberg E. (2023). Fiscal multipliers: A heterogeneous-agent perspective. Quantitative EconomicsVolume 14, Issue 3 p. 799-816
23. Cagetti M, De Nardi M (2006) Entrepreneurship, frictions, and wealth. J Polit Econ 114(5):835–870
24. Cagetti M, De Nardi M (2009) Estate taxation, entrepreneurship, and wealth. Am Econ Rev 99(1):85–111
25. Cingano, F. (2014). Trends in Income Inequality and its Impact on Economic Growth. OECD Social, Employment and Migration Working Papers.
26. Colciago, A., Samarina, A. and Haan, J. (2019). Central Bank Policies and Income and
27. Wealth Inequality: A Survey. Journal of Economic Surveys, 33(4), pp.1199-1231.
28. Díaz-Giménez J, Prescott EC (1992) Liquidity constraints in economies with aggregate fluctuations: a quantitative exploration. Federal Reserve Bank of Minneapolis Staff Report No. 149.
29. Díaz-Giménez J, Castaneda A, and Rios-Rull J. (2003) Accounting for the U.S. Earnings and Wealth Inequality. Journal of Political Economy, 2003, vol. 111, issue 4, 818-857
30. Doepke M, Zilibotti F (2014) Culture, entrepreneurship, and growth. In: Aghion P, Durlauf SN (eds)
31. Downs, A. (1957) An Economic Theory of Democracy, New York: Harper and Row. Handbook of economic growth, vol 2. Elsevier, Amsterdam
32. Fair RC, Taylor JB (1983) Solution and maxiumum likelihood estimation of dynamic nonlinear rational expectations models. Econometrica 51:1169–1186.
33. Ferrara, M. and Tirelli, P. (2015). Disinflation and Inequality in a DSGE Monetary Model: A Welfare Analysis. SSRN Electronic Journal
34. Florin O.B., (2019). Monetary Policy and Heterogeneity: An Analytical Framework.
35. Gabaix, X., Lasry, J. M., Lions, P. L., & Moll, B. (2016). The dynamics of inequality. Econometrica, 84(6), 2071–2111.

36. Garc´ıa-Pe˜nalosa C,Wen J (2008) Redistribution and entrepreneurship with Schumpeterian growth. J Econ Growth 13(1):57–80.
37. Kaplan, G., Benjamin M., and Giovanni V.. (2018). "Monetary Policy According to HANK." American Economic Review, 108 (3): 697–743
38. Grossmann V (2009) Entrepreneurial innovation and economic growth. J Macroecon 31(4):602–613.
39. Hansen GD (1985) Indivisible labor and the business cycle. J Monet Econ 16(3):309–327.
40. Krusell P, Smith AA (1998) Income and wealth heterogeneity in the macroeconomy. J Polit Econ 106(5):867–896
41. Krusell P, Smith AA (2006) Quantitative macroeconomic models with heterogeneous agents. In: Blundell R, Persson T, Newey WK (eds) Advances in economics and econometrics: theory and applications, ninth world congress, vol 1. Cambridge University Press, London
42. Le VPM, Meenagh D, Minford P, Wickens M, Xu Y (2016) Testing macro models by indirect inference—a survey for users. Open Econ Rev 27(1):1–38
43. Levine, R. (2008). Finance and The Poor. Manchester School, 76, pp.1-13.
44. Levine R, Rubinstein Y (2017) Smart and illicit: who becomes an entrepreneur and do they earn more? Q J Econ 132(2):963–1018
45. Matthews K, Minford P, Blackman S (1994) An algorithm for the solution of non-linear forward rational expectations models with current partial information. Econ Model 11(3):351–358
46. Meenagh D, Minford P, Wang J (2007) Growth and relative living standards—testing barriers to riches on post-war panel data. Cardiff Economics Working Papers No. E2007/12.
47. Minford L, Meenagh D (2020) Supply-side policy and economic growth: a case study of the UK. Open Econ Rev 31(1):159–193
48. Meltzer, A.H. and Richard, S.F. (1981) .A rational theory of the size of government., Journal of Political Economy, 89, 914-27.
49. Meltzer, A.H. and Richard, S.F. (1983) .Tests of a rational theory of the size of government., Public Choice, 41, 403-18.
50. Minford, L., and Meenagh, D. (2020) ‘Supply-Side Policy and Economic Growth: A Case Study of the UK , Open Economies Review, 2020, vol. 31, issue 1, No 7, 159-193.
51. Naceur, S. and Zhang, R. (2019). Financial development, inequality, and poverty: Some international evidence. International Review of Economics & Finance, 61, pp.1-16.

52. Pagano, M. (1993). Financial markets and growth: An Overview. European Economic Review, 37(2-3), pp.613-622.
53. Peterson, E. (2017). Is Economic Inequality Really a Problem? A Review of the Arguments. Social Sciences, 6(4), p.147.
54. Pietrovito, F. (2009). Financial Development and Economic Growth: A Theoretical and Empirical Overview. SSRN Electronic Journal.
55. Piketty, T. (2014). Capital in the Twenty-First Century. Translated by Arthur Goldhammer.. Cambridge, Massachusetts, England: Harvard University Press.
56. Piketty T, Zucman G (2014) Capital is back: wealth-income ratios in rich countries 1700–2010. The Quarterly Journal of Economics, 129(3):1255–1310
57. Reiter M (2009) Solving heterogeneous-agent models by projection. J Econ Dyn Control 33(3):649–665
58. Van-Der-Weide, R. and Milanovic, B. (2014). Inequality is Bad for Growth of the Poor (but Not for That of the Rich). The World Bank Economic Review.
59. Voitchovsky, S. (2005). Does the Profile of Income Inequality Matter for Economic Growth?. Journal of Economic Growth, 10(3), pp.273-296.
60. Wang, H. (2023) "UK economic growth and inequality from 1870", unpublished Cardiff PhD thesis submitted Sept 2023
61. Yang X., Minford P., and Meenagh D. (2021). 'Inequality and Economic Growth in the UK'. Open Economies Review, 37, pp 37-69.
62. Young, ER (2010) Solving the incomplete markets model with aggregate uncertainty using the Krusell–Smith algorithm and non-stochastic simulations. J Econ Dyn Control 34(1):36–41