A heterogeneous-agent model of growth and inequality for the UK

Thesis Submitted in Fulfillment of the Requirements for the Degree of Doctor of Philosophy of Cardiff University

Economics Section of Cardiff Business School, Cardiff University

by

Xiaoliang Yang

Primary Supervisor: Professor Patrick Minford

Secondary Supervisor: Dr. David Meenagh

September 2017

Acknowledgement

This thesis finished under the guidance of my supervisor Patrick Minford. I sincerely appreciate his help during my research. It was he who recommended me frontier research topics and relevant materials in macroeconomics and led me to the study on inequality. He gave me a lot of valuable suggestions and comments on topic selection, model setting and research method. I learned from him that as a researcher I should not only know how to do research but also should know how to present my idea and results. I am also deeply impressed by his profound knowledge, rigorous scholarship and more importantly his assiduousness. Every time when I saw this elder tired to move back and forth between the teaching buildings and conference rooms and every time when I saw him hurried to take a sandwich as lunch in a short break during a couple of hours seminar, I realised that perseverance and diligence were also quite important for a scholar.

I am also grateful to my secondary supervisor Dr. David Meenagh for his great help in my technique skills. I could always receive his responses and suggestions when I had programming problems no matter at weekends or even in holidays. I also want to give thanks to all faculty members, especially Dr. Helmuts Azacis who provide some precise suggestions as my annual meeting convener. I want to say thank you to my colleagues as I benefited a lot from our frequent discussions. Besides, I am indeed grateful to Cardiff Business School for offering me the research opportunity and the funding.

Last but not least, I give my special thanks to my families for their continuous encouragement and support these years.

Abstract

This paper analyses the effect of wealth inequality on UK economic growth in recent decades with a heterogeneous-agent growth model where agents can enhance individual productivity growth by allocating time to entrepreneurship. Entrepreneurship cost is negatively correlated to individual wealth which originates from the fact that the rich are more likely to undertake entrepreneurship than the poor. An appropriate wealth concentration to the rich theoretically stimulates their entrepreneurship incentives and then aggregate growth. Given UK quarterly data from 1978 to 2015, our model cannot be rejected to be true using the Indirect Inference method. The empirical study finds that our structural model could generate a stable relation between inequality and growth and model simulations could fit main properties of UK economy. Wealth inequality is found to stimulate economic growth, especially in a long term. Policy makers have to face a trade-off when conduct a redistribution policy like taxation because inequality reduction will be followed by a slow-down of economic growth. Moreover, as redistribution tax rate increases, growth reduction has a gradiently increasing trend and thus a moderate tax rate is a priority option for policy makers. Our comparison between tax regimes shows that the tax transferring income from the rich to the poor is preferred to others.

Key words: heterogeneous-agent model, entrepreneuring, aggregate growth, wealth inequality, redistribution, indirect inference

CONTENT

1 Introduction

Representative-agent model (RAM) with the assumption that economic agents have identical behaviour and expectation is the most important and prevailing idea of macroeconomic and financial modeling. However, it is frequently found incapable to explain some complex economic phenomena and fit some stylised facts. For instance, in asset pricing studies, the puzzles why the equity premium is extremely high and why the risk-free rate is extremely low, etc. can hardly be explained by RAMs. In business cycle studies, the questions why labour volatilities are relatively high and why wealth and income distribution across agents are dynamic, etc. can also not be answered by RAMs. Bewley (1980) (1983) firstly establishes a heterogeneous-agent model (HAM) to consider various individual behaviours by introducing idiosyncratic shocks (individual income endowments) and incomplete capital asset market (i.e. borrowing constraint where no borrowing is allowed in his model) which determine different state conditions across agents when optimal decisions are made and then generate heterogeneity. As the development of computational techniques and numerical algorithms as well as the popularity of microeconomic data in recent decades, HAM has been more widely used.

Development of HAM and the relevant numerical algorithm experiences three important stages. In the early stage, only idiosyncratic shocks (like individual income uncertainty and employment uncertainty) are employed for heterogeneity. For example, Hansen (l985) and Aiyagari & Gertler (1991), etc. use HAMs to explain the asset puzzles by “self-insurance” behaviours that individuals demand much more risk-free assets than liquidity assets due to the uncertainty of income. Aiyagari (1994), extremely renowned in this area, uses an HAM to explain why individual wealth and consumption are much more volatile than aggregate ones by defining a special equilibrium distribution where a constant real interest rate exists such that aggregate variables remain unchanged while individual ones could be time-variant. The numerical algorithm in first generation models concentrates on solving for equilibrium market prices like real interest rate. The second generation represented by Diaz-Gimenez and Prescott (1992), and Krusell and Smith (1998) (2006) aims to overcome the shortcoming of the first generation that aggregate volatility is lack in despite of sufficient individual volatilities by introducing both idiosyncratic shocks and aggregate uncertainties. They develop a new method to solve models by searching for an equilibrium law of motion for (wealth) distribution around which some new numerical algorithms are developed. Distribution is generally described by finite order moments for simplification and individual decisions are assumed to be made based on distribution moments. However, there might be an infinite-dimensional issue if high order moments are consider for individual optimal decisions. Hence, moment order is finally diminished to one which result in the so called "approximate aggregation" raised by Krusell and Smith that the mean of (wealth) distribution and aggregate shock could well determine the aggregate behaviour. The first two generations generally solve models by approximating a continuum of agents to a finitely large number of agents in practice. As some theoretical derivations are based on the assumption of continuous agents, this approximation might lead to variant sampling errors. The third generation emphasises on solving models with a continuum of individuals and moreover attempts to remove the dependence on aggregate law of motions when solve for individual behaviours. For instance, Algan et al (2008) and Bohacek and Kejak (2005) develop a projection method on Krusell and Smith's model. Reiter (2009) and Young (2010) adopt a combination of perturbation and projection methods. In most third generation algorithms, aggregate law is not essential to solution for individual behaviour, but instead they search for equilibrium cross-agents distributions in each period described by density functions, resulting in that "approximation aggregation" is not an inevitable finding.

Although the three generation models are helpful to explain some properties of business cycles, the long-run effects of wealth and income distributions on economic growth which takes increasingly more attention are not concerned.[1] Conversely, many other scholars establish theoretical models in order to derive an analytical relation between inequality across agents and aggregate growth rate, especially the inequality effect on growth rather than stimulate important individual behaviours and distributional properties as reality, with neither analytical or numerical solutions. Most of these growth models are developed from OLG models with different mechanisms of inequality effects on growth such as indivisible investment or labour, incomplete market barriers, "Median Voter Theorem" and security hazard. Given theoretical support of these models, many empirical studies on how inequality affects growth which generally adopt regression analysis are developed. Their regression methods include cross-country regression for which least squares estimation is frequently used and panel data regression where pooled least square, general GMM, difference GMM and system GMM are optional. After decades of debate, there is still no consensus on whether inequality could stimulate or impede economic growth has no consensus. For example, Alesina and Rodrik (1994) with cross-country OLS, Deininger and Olinto (2000) with panel system GMM and Bagchi and Svejnar (2015) with panel difference

[1] The effects of wealth or income distributions on growth are generally explained as stochastic ones due to exogenous shocks in all the three generation frameworks.

GMM find a negative effect of inequality (income or wealth) on economic growth. Contrarily, Perotti (1996) with cross-country OLS, Barro (2000) with panel random effects LS, Forbes (2000) with panel difference GMM and Ostry et al. (2014) with panel system GMM find a positive inequality effect (some only in developed countries). Overall as Halter et al. (2014) pointed out, both estimation method and sample employed have considerable influences on the estimated inequality effects. In fact 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 distributional characteristics of UK economy in recent decades and further study the wealth inequality effects on UK economic growth. We introduce heterogeneity into an endogenous growth model on the basis of Minford et al (2007)'s model by classifying population into two different groups for simplification, say the rich who own higher capital holdings and the poor. Individuals can allocate their same time endowment to leisure, labour and entrepreneurship incentive (time) where personal entrepreneurship incentive drives individual productivity growth and is also costly with per unit time cost . Our development is to relate individual entrepreneurship cost to capital distribution so that the rich have more entrepreneurship incentives than the poor. This idea comes from the fact that the rich have sufficient wealth to do surplus activity. We are not going to deny the success of a few entrepreneurs from impoverished backgrounds. However, it is easy to enumerate more successful entrepreneurs born in rich families or the middle class, such as John Pierpont Morgan, Rupert Murdoch and Warren Edward Buffett, William Henry Gates III and Steven Paul Jobs etc. Levine and Rubinstein (2013) investigate who are more likely to become Schumpeterian entrepreneurs measured by the incorporated self-employed using NLSY79 data in US and find that more entrepreneurs come from well-educated and high-income families. Since the macroeconomic behaviour comes from the aggregation of individual ones, our prospective effect of wealth inequality is that inequality concentrates wealth to the rich with more entrepreneurship incentives which enhances aggregate productivity growth and then economic growth. The traditional incomplete market assumption is released in the model as individuals have incentives to accumulate assets in order to pay entrepreneurship penalty and there is sufficient endogenous heterogeneity.

We apply the growth model to UK economy with the quarterly data set after 1978Q1 due to data availability where inequality of wealth and income is measured by share of allocations to the rich. The research method can be simply adopted on other countries. To ensure the rationality of further analysis, we first examine whether benchmark model could artificially a stable tendency of capital inequality and economic growth co-movement and also whether policy intervention for redistribution like taxation is feasible by starting with two Adhoc identical individual groups where posterior heterogeneity solely comes from randomly idiosyncratic shocks. Afterwards we assess whether our model fits UK stylised facts and provide policy suggestions according to empirical findings. What we are most interested in is not whether our model could perfectly fit all the aggregate and individual behaviours which in fact is difficult to be realised due to the application of both individual and aggregate data, but instead whether the model could fit some important characteristics relevant to our topic such as the relation between wealth inequality and growth. Therefore, we employ indirect inference (Id-If) early formally used by Le et al (2011) to test model and the testing power is found very high. The basic idea of Id-If model testing is as follows. Given the null hypothesis that structural model is true, one could firstly achieve structural errors according to the difference between actual data and model computed values for variables and then bootstrap structural errors to obtain simulated samples.[2] As model is assumed true, simulated data should have same properties as actual data and those properties are described by an auxiliary model which could be impulse response functions, moments of data and so on. We choose a VARX as the auxiliary model in practice including key variables we are interested in. Parameter estimation is also on the basis of Id-If that parameter values are searched for until null hypothesis is not rejectable.

Our empirical study concludes three important findings. First of all, the structural model could generate a stable relation between inequality and growth and model simulation could fit main properties of UK actual data. Secondly, wealth inequality is found to have a stimulating effect on economic growth, especially in a long term. Lastly, although political redistribution intervention like income tax could reduce inequality to some extent, the corresponding cost is a slow-down of economic growth implying that policy makers have to face a trade-off. Moreover, as tax rate increases, growth reduction has a gradiently increasing trend and thus an appropriately low tax rate instead of a high rate is prior for policy makers. The comparison between tax regimes shows that an income transfer tax (transfer income from the rich to the poor) is preferred to other tax regimes.

This paper has following structure. The current introduction is followed by Literature review in next section. Our model setting is illustrated in section 3. Then section 4 describes Indirect Inference method including model test and estimation in detail. Section 5 introduces employed data. Empirical results are described in section 6, including detection on tendency, empirical study on actual data and redistribution.

---

[2] In practice, we bootstrap innovations which are residuals of AR regressions on structural errors. See details in section 4.

The last section concludes.

2 Literature Review

We firstly take a review on some HAMs and relevant solution algorithms which are aimed at fitting distributional characteristics in real economy and interpreting some micro-behaviours which can hardly be explained by traditional RAMs. The first generation of HAMs has two key features, idiosyncratic risks which cannot be fully insured and incomplete capital market (i.e. borrowing constraints) originating from Bewley (1977). He theoretically studies the permanent income hypothesis (this hypothesis indicates that the response of consumption to temporary economic volatilities e.g. prices and income is smooth, different from Friedman's long-run hypothesis). Consumer is assumed to face a borrowing constraint (no borrowing at all) and a random exogenous income. Consumer knows the probability of the stochastic environment and thus could change his savings to compensate the consumption. This mechanism is called "precautionary saving" nowadays. Then heterogeneity is introduced into macroeconomic models by Bewley (1980) (1983) to continuously study the permanent income hypothesis by setting multi-sectors and multi-consumers with random income endowments. Afterwards, HAM is widely applied in studying some important features of business cycles (e.g. volatilities of labour) and the asset pricing.

Hansen (l985) concerns one important labour property in business cycles that fluctuations of aggregate working hours are mainly caused by the fluctuations in the 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 does not work) and probably receives 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.[3] The crucial step of the computational algorithm is to find a constant real interest rate which guarantees a steady state.[4] 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 the similar role as the interest rate in Aiyagari and Gertler (1991).[5] 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 widespread adoption of heterogeneity in macroeconomic modeling benefits from the development in the computational algorithm. Aiyagari (1994) makes an honorable contribution to this development. In his Bewley model, each individual has budget constraint where individual asset

[3] There are two securities, treasury bills without exchange cost and stocks with exchange cost.

[4] The details of the computational algorithm of a sort are illustrated later in Aiyagari (1994).

[5] The algorithm in Huggett (1993) is similar as the one in Aiyagari & Gertler (1991) and Aiyagari (1994). However, different from using a bisection method to update the interest rate in Aiyagari (1994), the price of credit balance, q, will be raised if current q leads to an excess demand for credit while will be decreased if else.

[6] and the individual labor follows

.[7] The average labor is always a constant . The wage rate *w* is constant equal to the average marginal labor productivity. He focuses on the stationary equilibrium characterised by a constant real interest rate *r* which equals the marginal capital productivity. To consider the wealth movement, he defines individual total wealth (resource) as where the individual asset demand is . Solve individual's problem using the rewritten constraint to obtain individual optimal asset demand function while individual wealth evolves following . The key step to solve the model is to find the optimal interest rate for the stationary equilibrium with the algorithm details below.

1. It begins to approximate the given rule for individual labour using a 7-state Markov chain process in order to ensure the expected labour supply equals a constant (normalised to 1). is the state space of . Firstly, divide the real interval into the following seven ranges – and – corresponding to the 7 states of above. For each , randomly draw many times to calculate using its AR(1) process, given each pair of . Then use the obtained to calculate the Markov chain transition probability . Next use to calculate the unconditional probability of , denoted as and calculated the expected value of , denoted as . Then define to

[6] . *b* is a positive constant.

[7] where parameters can take on the values . The stochastic rule implies that has the mean 0, variance and serial correlation coefficient .

normalise the expected labour supply to one. Now, the 7 states of individual employment status can be updated as with the same transition probability above.

2. Use the Markov process above to generate a series of employment status.
3. Give an initial guess on the interest rate *r*, denoted as . Use the initial asset, the asset demand function and the series of employment above to obtain a series of the asset demand and also a series of the asset .
4. Calculate the sample mean of , denoted as
5. Solve the EQ condition to obtain .[8]
6. Then use the bisection method to set − and repeat the steps above to renew . If , replace by to yield − . If , replace by to obtain − . Keep doing so until close to .

Afterwards he uses the accurate simulations given the equilibrium interest rate to calculate indicators, like moments and skewness, etc. to quantitatively consider how idiosyncratic shocks affect aggregate saving behaviour. The equilibrium he defines could explain why individual wealth and consumption are much more volatile than aggregate ones.

Unlike the typical model setting with a continuum of individuals which in practice is generally approximated by a finite large number of agents, some works set a small number of groups of agents in order to avoid the possible variant sampling error (Discuss about this error later). Kydland (1984) classifies the consumers into 5 different types by their skills also with borrowing constraints to explain the stylised

[8] describes *k* as a function of *r* according to . should suffice for the condition for finite asset accumulation — where is called "time preference rate", otherwise replace by

fact that working hours fluctuate much more than the aggregate labour productivity and the real wage rate, which could not be well explained by the representative-agent models. Scheinkman and Weiss (1986) strengthen the restrictions on random indivisible labors in Hansen (l985) that there are only two groups of agents and only one group is productive (employed) at each period. Agents face a random duration of being productive. Since the real wage is assumed to be constant, the fluctuations in individual productivity could mimic the features in the business cycles.[9]

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 cycles by introducing heterogeneity in 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 fluctuated distribution on the aggregate behaviour. Furthermore, the lack of aggregate uncertainty is implausible.

The use of aggregate uncertainty in addition to idiosyncratic shocks is a symbol of the second generation of HAMs. Mankiw (1986) makes an early attempt to

[9] Some others, like Mankiw (1986), Telmer (1993) and Den Haan (1996), also use finite groups. However, I classify them into the second generation model with heterogeneity because they all introduce the aggregate uncertainty.

introduce both aggregate and individual shocks into a theoretical model with a continuum of individuals and no labour involved. The aggregate economy could have two possible states, good and bad, both with the realisation probability –. Individuals have an exogenous consumption space with values, a certain value when state is good and an uncertain value with unconditional expectation when bad where . Individual stochastic property mainly originates from the consumption at the bad state where consumption could be with probability while it could be —— with probability . He claims that consumers consider a utility maximisation problem by deciding how much of claims to buy. But agents in his model are essentially given an exogenous labour income where the "payoff" r can take on the value at a good state and at a bad state. And a central planner decides how to set y, given shocks.[10] Although he creatively relates individual shocks to the aggregate uncertainty, he ignores individuals' allocation trade-offs between consumption and some insurance assets. Telmer (1993) establishes two models, a non-tradable model and a tradable one, both with two groups of individuals whom are endowed with non-tradable labour income, depending on idiosyncratic shocks. Individual tradable bonds are supplemented into the latter model like in Huggett (1993). Aggregate income where the gross growth rate can take on two values (good state) and (bad state), is determined by a Markov process. Individual stochastic process is based on the one in Mankiw (1986) that both have an identical and certain income in a good aggregate state while have probabilities to be endowed with different incomes in a bad state. The difference is that two probabilities related to two individual income values in a

[10] I develop this idea and the central planner's problem is which leads to the solution same as equation (4) on Page 4 in Mankiw (1986).

bad aggregate state are not identical. Diaz-Gimenez and Prescott (1992) introduce an aggregate shock (reflects an exogenous real interest rate) into a Bewley model with a large sample of agents. The individual labour still has dummy values, also depending on an aggregate shock .[11] Individual shocks are not simply , but have the state space where the two sub-states and represent the human capital state space and the productivity state space respectively.[12] Individual labour income equals the product of and real wage which equals the marginal productivity when employed. The computational procedures, different from Aiyagari (1994), are shown below.

1. Guess the relations between prices (assume both prices of goods and bonds to be functions of) and .
2. Obtain the individual optimal rules, given prices.
3. Use individual rules and market clearing to obtain the law of motion where describes individual current asset and employment status.
4. Use the Markov process to update shocks to one-period ahead and repeat steps above. Keep doing so to get a time series.
5. Check if the obtained government spending (residual of market clearing) are all negative. If they are nonnegative, the initial given relations between prices and could reach equilibrium. Otherwise the equilibrium does not exist.

In some sense, the Markov process in this paper is a degeneration of the one in Aiyagari (1994) because the transition probabilities are artificially selected to match the US employment data rather than numerically approximated (i.e. the convergence is not proved). Furthermore, individual shocks are independent of the aggregate shock

[11] Some experiments are implemented in this paper according to the classification of the real interest rate regime where stands for a high rate regime while represents s a low rate regime

[12] See Page11 in Diaz-Gimenez & Prescott (1992) for details

for simplification, diminishing effects of the aggregate shock.

The second generation model has a wide spread after Krusell and Smith (1998)'s work. They 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 the way in 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 with realisation probabilities such that the total unemployment rate in a bad state and that in a good state are always constant, and respectively. They solve a decentralised problem where individual real wage and real interest rate are marginal productivities respectively in terms of the aggregate labour and aggregate capital. They creatively 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 only concern the first *I* moments of the wealth distribution, denoted by which is used to approximate an infinite dimensional wealth distribution where and *H* is an undetermined evolution function. The kernel to find an equilibrium solution is to find a stationary *H* with the numerical algorithm which takes the following steps.

1. Choose a value of *I* (could start from 1). Guess an initial functional form of with initial values of parameters.
2. Solve individual optimal rules, given and an initial wealth distribution.
3. Stimulate individual data using the individual rules.

4. Use the simulated data to calculate the simulated aggregate moments and then use least-square to estimate parameters in the function of . here is used to assess the estimation.
5. Check whether the estimators are close to the initial guess. If not, change *I* and repeat steps above; if else, stop.

One of their important findings is "Approximate Aggregation" (APAG) which means that the first moment of the distribution (mean of wealth) and the aggregate shock could well determine aggregate behaviour. Increasing the moments does not do it much better. Moreover, the initial distribution has neglectable effects on the results. Nevertheless, like Aiyagari (1994), this benchmark model also generates few agents who have low wealth levels, leading to same marginal propensities of consumption for many individuals. This is because foe the rich agents, wealth holdings are sufficient for full-insurance so that the borrowing constrains play no roles for these agents even if they have different individual states (employment status and wealth). To overcome this shortcoming, they develop a preference heterogeneity model with the assumptions that the unemployed also have income and consumers could have different preference discount factors. They find that the poor plays a more important role in the consumption aggregation and thus the aggregate consumption behaviour differs from the traditional permanent income hypothesis. The wealth aggregation, however, is still hardly affected by the poor. The heterogeneous preferences actually imply that some individuals must be poor since they are given lower discount factors, which seems like a self-fulfilling result. Krusell and Smith (2006) enrich their previous model by introducing another asset, bond, in addition to capital, into a two-period model and endogenous labours combined with an idiosyncratic employment shock following the same Markov process as that in Krusell and

Smith (1998). Constraint takes the form

. They use a slightly different algorithm to solve the equilibrium. To find a stationary law of motion *H* for the wealth distribution measurement (e.g. log-linear of the 1st moment), they also start with an initial guess on an functional form of *H*. But they do not set initial values for the undetermined parameters in *H* this time. They solve the individual optimal rules then and stimulate the 2nd period individual saving and wealth combined with the drawn shocks, given the initial distribution . Now the 2nd period can be calculated using the simulated individual data and the drawn 2nd period aggregate shock. They keep doing the iteration this way to obtain a time series of . Then they directly use least-square to estimate the undetermined parameters in *H*. After that, they replace the undetermined parameters by the estimators and repeat the steps above until parameter values converge. Note that endogenous and makes prices (*r*, *w* and *q*) more difficult to clear the markets and thus they use a dichotomy method. That is, firstly suppose that consumers make decisions based on the given market prices. Then in the simulation, market prices are allowed to vary in order to guarantee all markets clearing. But this dichotomy is debatable. Consider the aggregate labour which is the mean of all the individual labours. Since is achieved from individual optimal problem, there is no rule to guarantee a constant mean of all . Overall, their simulated data could capture most cyclical features except that the poor takes a low proportion as in the benchmark model in Krusell and Smith (1998) and simulated wealth inequality (Gini-coefficient) is counter-cyclical while the US data shows a cyclical inequality. They use R^2 to evaluate the accuracy of the *H* approximation in both papers, although several methods are discussed about in Krusell and Smith (1998). Den Haan (2010a) and Algan et al (2014) both argue that R^2 is problematic to test the accuracy of an approximated aggregate law because same law of motion expressed by different ways might result in different values of R^2. Consider Krusell and Smith's approximated law of motion . Rewrite it by minus on both sides as . Both equations are identical, but R^2 in the latter equation estimation is smaller, since the variance of is lower than variance of . Krusell and Smith (1998) (2006)'s work is developed by many scholars. For instance, Mukoyama and Sahin (2006) basically use Krusell and Smith (1998)'s model, also with a linear aggregate law of motion to study the welfare cost of business cycles across agents. A finite group of individuals faces two types of idiosyncratic shocks, individual employment status which depends on an aggregate productivity shock and individual skill status which is independent of the aggregate shock. Moreover, a special transition probability such that the probability of being unemployed depends on the skill status (unskilled agents are more likely to be unemployed in future) is adopted. They find their model is helpful to explain that unskilled individuals have more welfare lost during recessions. Chang and Kim (2007) also employ the same model, but with a different algorithm. They also assume a linear aggregate law of motion for aggregate capital and assume that both wage and interest rate are linear functions of aggregate capital and aggregate productivity shock. Then choose 11 grid points for aggregate capital and together with individual optimal rules they simulate time series on 200,000 agents. Afterwards, the sample of aggregate capital can be obtained from the aggregation of individual capital and is used to OLS estimate the proposed aggregate law.

The utilization of an aggregate law of motion in early second generation models is a creative way to make individual and aggregate behaviours independent to some extent and thus it could simplify some algorithm when consider both aggregate and individual shocks. It, however, seems like a self-fulfilling procedure to firstly assume a law and then to show it cannot be rejectable. It is hard to say whether an aggregate law exists, although aggregation is composed of individual variables. Additionally, an infinite dimensional issue when use the first i moments of capital distribution as state variables to determine individual policy rules is raised by Algan et al (2014). They follow Krusell and Smith's assumption that individuals consider the first I^{th} order moments denoted by where is the employment status to make decisions and approximate optimal capital rule to the following N^{th} order polynomial.

where coefficients depend on the aggregate state . For convenient notations, individual superscripts are omitted. Decentralised moments (i.e.) are considered later on instead of decentralised moments for simplification.[13] Consider Krusell and Smith's linear policy rule, i.e. . Aggregating (AL1) with powers from 1 to I directly yields the aggregate state in next period where . This indicates is sufficient to obtain when individual optimal rules are linear w.r.t. state variables. Now consider a nonlinear individual rule, for example . Aggregates (AL1) with the power from 1 to I, yielding the followings.

[13] To transfer decentralised moments to centralised moments, one just need to change parameter values.

That is, to obtain , one has to firstly increase dimension of the aggregate state from to . However, after that, one has to continuously increase the dimension of moments from 2*I* to 4*I*. The continuously increasing dimensional requirement leads to an infinite dimensional issue and this issue exists as long as individual optimal rules are nonlinear about state variables, even if individuals are assumed to only consider the first order moment.

There are still some scholars make an attempt to avoid using an Adhoc aggregate law of motion like Den Haan (1996) and Preston and Roca (2007). The former uses a basic Bewley model in which *N* (a finite number) consumers have the constraint where is the bond purchased at time *t* with price .[14] Individual income is determined by the employment status with probability of being employed identical across individuals. could be positive, i.e. the unemployed agent still has an income and this assumption is also adopted in the second model in Krusell and Smith (1998). The probability of individual employment status also depends on the economic state (good and bad) which follows a Markov process. Individual employment probabilities should satisfy the condition that the expected total unemployed proportion when state is bad greater than that when state is good where the total unemployed proportion is . One of his contributions is the use of Parameterised Expectation Algorithm (PE)

14 _ and _ , same as that in Huggett (1993).

which refers to approximate the rational expectation in the individual optimal rules using a certain parameterised function with individual states. To use PE, he firstly optimises individual utility ———, given the budget constraint and the borrowing constraint, to obtain the following optimal rules, using the KKT condition.

– (Den 1)

(Den 2)

– (Den 3)

(Den 4)

Additionally guarantees the bond market clearing. Individual future consumption depends on current state, including economic state and wealth distribution (described by a function of individual wealth levels). Then can be approximated by an undetermined function where represents a measurement of the wealth distribution.[15] is an undetermined parameter vector.[16] The key to solve the model is to find the equilibrium and an optimal . Given a functional form, the latter is equivalent to find an optimal with main procedures below.

[15] First of all, wealth distribution can be described as a function of all the individual wealth levels. Then one can separately approximate this function as a measurement, e.g. moments and shares, etc. Haan states that the wealth statistic covering *I*-1 agents, but it could cover all the agents. That is, could replace

[16] He chooses an exponential form of a polynomial with order n to replace in practice.

1. Use the Markov processes to generate time series of economic states and individual employment status.
2. Calculate the distribution measurement for each individual, given the initial individual wealth values.[17]
3. Guess an initial value of , say so that is determined for each individual.
4. Replace by from (Den.1) and (Den.2) to yield:

_ (Den 5)

(Den 6)

5. Suppose _. Given an initial guess on , can be solved from (Den 5) (i.e.). Calculate by (Den 4). If the achieved _, let _ and use (Den 4) to calculate .
6. Check the bond market equilibrium condition. If , replace the initial by − and repeat the steps above. If , replace by − and repeat the 5 steps. Keep using this bisection method until .
7. Choose any one agent *i*'s simulated series based on the initial value with an equilibrium obtained from step 6. Regress on and denote the nonlinearly estimated parameter vector as . Then update to by and repeat all the steps until a convergent is achieved.

This algorithm has an advantage in the flexibility of the wealth distribution

[17] Actually, one only need to be given initial values of individual bonds, because the other component of wealth is known after generating individual employment status (recall)

measurement as there is no law of motions for a certain measurement so that not only moments of wealth distribution, but also deciles, quintiles and even shares could be used. Nevertheless, issues also exist. Firstly, the algorithm excessively relies on the selected functional form of and the initial value of . Particularly, convergence of in step 7 above depends on the artificially chosen . Secondly, a time-consuming step to find the equilibrium has to be embodied into the algorithm of searching for an optimal just because there is no direct way to solve for in step 6 above and the reason for the latter is that an exact value of is required in step 5 when compare with _.

Preston and Roca (2007) make an development by introducing a utility penalty of borrowing taking the form of ——— where and the lower bound of capital holding is into individual objective functions in order to find certain forms of first order conditions (uncertain first order conditions come from the inequality restriction, i.e. the borrowing constraint), different from the way using KKT condition in Den Haan (1996). That is, the utility decreases as individual borrows more (is closer to). They contributively adopt perturbation methods (PBMs) in HAMs for model solutions. Prices (*r* and *w*) equal marginal productivities of an aggregate production with the aggregate technological shock . The individual labour is where employment shock .[18] Theoretically, Markov processes with a finite number of states are not necessary to approximate the stochastic processes above in a PBM, although in practice, they are used for simplification. To use a PBM, they firstly solve for the first order conditions which give an analytical solution of their model, namely optimal

[18] and and are mean of and .

rules for control variables (and) depending on predetermined variables (and) and shocks (and).[19] To save words, I call both predetermined variables and shocks as exogenous variables below. Then they use a second order Taylor expansion to approximate the optimal rules around a special equilibrium where neither an aggregate shock nor individual shocks exist and thus wealth is equally distributed across agents. Note that the unknown coefficients in the approximated optimal rules are now the first and second order derivatives with respect to (w.r.t) exogenous variables with their equilibrium values. Next, they differentiate the raw model equations w.r.t exogenous variables directly and set exogenous variables to their equilibrium values in order to get equations with the unknown coefficients in the approximated optimal rules. Solve these equations to figure out the unknown coefficients. The special equilibrium assumed previously is important to this method. Consider the term where is the aggregate capital, for instance. With idiosyncratic shocks, one could only get and the term is uncertain, leading to unsolvable approximated optimal rules. Given the special assumption above, one have with no uncertainty. Following this way, both individual optimal rules and aggregate capital rule finally depend on the sate space where is the variance of individual capital and is the covariance between individual capital holdings and employment shocks. Given that assumption, both and equal zero in equilibrium. Unlike checking R^2 in Krusell and Smith (1998) (2006), they calculate the Euler equation errors to assess the accuracy of their

[19] See details of the model equations on Page 11 in Preston and Roca (2007).

method. Denote as the Markov transition probability of states. Given the first order condition

, they replace consumptions by their optimal rules and to yield

. Then they solve the inverse marginal utility to obtain

and define the Euler equation error as which also depends on the states. Given a series of states, they examine statistics of , like mean, standard deviation and extreme values which all show accurate results. Nevertheless, their calibrated model also shows APAG as in Krusell and Smith (1998) and many other studies. Additionally, use of the utility penalty is debatable because the penalty makes agents almost impossible to touch the borrowing constraint.

Except some studies classifying individuals into types with common features (like skilled and unskilled) in each type, most theoretical models illustrated above are implemented in practice by approximating on a continuum of agents in the original assumption to a finitely large number of agents. Since some theoretical derivations are based on the assumption of continuous agents, this approximation might lead to variant sampling errors. Therefore, some algorithms emphasised on solving models with a continuum of individuals are raised, which is called the third generation HAMs by me. 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.

But different from the essentiality of the aggregate law to solutions for individual optimal rules in Krusell and Smith's model, the aggregate law in their algorithm can be completely neglected when solve individual rules. The reason for still using the aggregate law is that individuals only consider simple measurement for wealth distribution to predict prices rather than consider a complicated time-variant cross-agents distribution period by period. I clarify some notations before illustrate their new algorithm. They distinguish the *j*th order moments of capital distribution at the beginning and the end of *t*, denoted by and respectively where the moment involves the agents with current employment status . Assume individuals are only concerned with the first *N* moments, denoted by where is the proportion of agents who are unemployed and simultaneously touch the borrowing constraint boundary ( ) out of total unemployed agents at the end of period *t*-1.[20] All the other 2*N* moments consider individuals with positive capital because the moments when at the beginning of each period can be calculated. Define the aggregate state as with the aggregate technological shock .[21] Their algorithm starts from describing the aggregate law as a function (a polynomial in practice) such that with parameters where . The idea to solve for and to confirm the polynomial order is to find the best responses (whether individual optimal rules corresponding to also result in the same (or close)). Hence, given an initial guess on , say , they use the

---

[20] since the employed agents will not run out of capital at the end of one period. See details on how to derive moments at the end of one period from those at the beginning on P33-35 in Algan et al (2008).

[21] Actually there are two alternatives to express , but the other one is more complicated because of the extra components, . See details on P10 in Algan et al (2008).

following 5 steps (first time to use a PJM) to solve individual rules for consumption and capital . First of all, first order conditions are solved using KKT condition as Den Haan (1996) to deal with the inequality constraint and a polynomial function with undetermined parameters, denoted as is used to approximate the rational expectation terms. Secondly, given an initial value on , say , they solve individual optimal rules with wage . Note that both wage and interest rate are marginal productivities of the aggregate production and thus only rely on aggregate states, but disappears when approximating the rational expectation to a polynomial. Thirdly, assume both capital and moments are allocated in some ranges and choose several grid points in each range using the Chebyshev polynomial method.[22] Calculate and on all grids. Fourthly, use the aggregate law to get and then drive to yield . Next, to get the values of expectation terms, denoted by *E(X)*,[23] they primarily calculate *X* using and , given any possible value of aggregate states. Then calculate the mean of all *X*s as *E(X)*. Fifthly, use least-squares to estimate polynomial parameters (minimise the sum squares of the gap between achieved *E(X)* and the polynomial with unknown parameters) and yield the estimators . Replace by − and repeat previous steps until is convergent. The next stage of the whole algorithm is to derive the aggregate law from the achieved policy rules using a PJM again. The idea is that they use optimal rules and cross-agents probability density functions (PDFs) to get new time series of moments at the beginning and the end of periods, denoted

[22] They choose 50 grid points for capital within a range [0, 99] and 5 grid points for the moments within different ranges.

[23] *X* in this model takes the form of .

as and respectively. Then minimise sum of the squares of where includes to yield the estimators . During the procedure of generating moments, the time-independent cross-agents PDFs need to be confirmed in advance at the beginning of each period. They firstly approximate PDFs with employment status by the following polynomials.[24]

where the higher-order moments is called "reference moments" and used for a more accurate approximation. If reference moments at each period are given, optimal parameters can be obtained by solving the first order conditions from the optimal problem . Optimal ensures . The illustration on the algorithm to generate moments, including updating and reference moments, and transferring moments from the beginning to the end of each period, is skipped here.[25]

As I mentioned before, the PJM could work on a true continuum of individuals and could diminish the dependence of individual solutions on an Adhoc aggregate law if the researcher does not focus on that law, although some algorithm details need to be improved. The moments of positive capital holdings are updated following the rules below.[26]

[24] k in the PDF is not a special because this PDF is cross-agents instead of cross-time for one agent

[25] See details on P16-17 in Algan et al (2008).

[26] See details on P35 in Algan et al (2008).

where represents the number of agents who have employment status at t and have at t+1, and it depends on and . The first equation, for example, describes that the j^{th} order capital moment of the unemployed agents at the beginning of period t+1 equals a weighted sum of the j^{th} order capital moment of the unemployed and that of the employed at the end of current period t. This actually is a debatable approximation instead of an accurate moment.

Many algorithms are developed to consider a continuum of individuals and one key step is the dynamics of the cross-agents distribution along grid points of capital. For example, Bohacek and Kejak (2005) adopt a PJM method on the Krusell and Smith (1998)'s model. Their algorithm is similar as Algan et al (2008)'s one but simpler since they focus on the cross-agents distribution without the step to find a convergent aggregate law. Different from Algan et al (2008), they does not directly approximate the rational expectation term in individual first order conditions to a function, but firstly use the sample mean of all the possible forms of the term which are expected corresponding to all the possible realisation of shocks to replace expectation terms. For instance, suppose the stochastic shock has J discrete realisations with transition probability from one period to next for any individual i. Then individual first order conditions w.r.t consumption can be approximated by where means that current consumption depends on current capital holdings and gives next period capital. The realisation of next period consumption depends on the transition probability. Afterwards, they approximate individual rules for consumption as Chebyshev polynomial functions of individual capital. They also approximate one stationary distribution condition (similar as Algan et al (2008), but simpler due to time-independent parameters) as a Chebyshev polynomial. Then a PJM is used to estimate the parameters. Although they consider a continuum of agents, the aggregate shock is neglected.[27] Reiter (2009) brings in a lump-sum government tax rate on Krusell and Smith (1998)'s model in order to introduce more volatilities of individual wealth at period *t*, denoted by which is composed of the gross return on the capital at the end of period *t*-1, denoted by , stochastic wage income where is a continuous random labour productivity with mean 1, and the lump-sum tax subsidy where is the tax rate. Individual has the budget constraint . He adopts a combination of a PBM and PJM to solve the model (the kernel step is the transition of PDF on grid points and thus Algan et al (2014) name his method as the grid method). First of all, a PJM is used to solve for individual optimal rules and then an approximating cross-agents wealth distribution with idiosyncratic shocks but no aggregate uncertainty. In the first stage, unlike Algan et al (2008), he approximately describes cross-agents distributions by fractions of agents holding a certain amount of capital instead of moments. To obtain the dynamics of cross-agents distributions at

[27] I still classify this paper into the 3rd generation because the aggregate shocks are not difficult to be introduced using this type of algorithm by renewing the state variables.

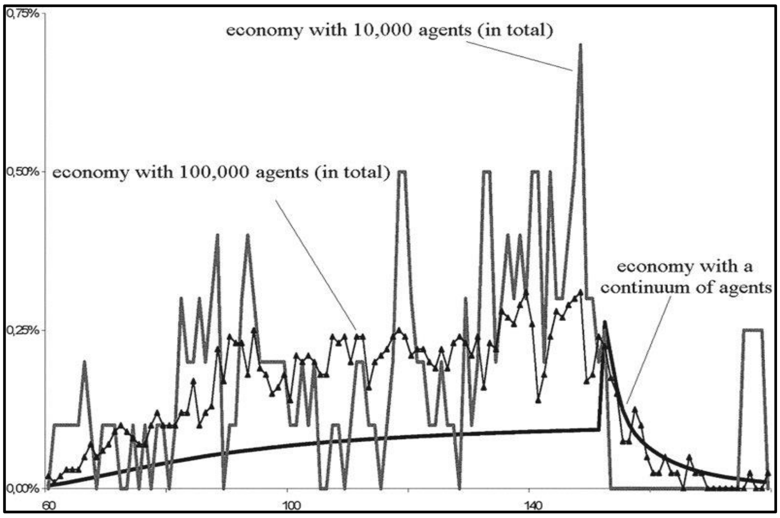
the beginning of each period, he chooses 1,000 grid points on an interval of beginning-period capital and assumes agents are uniformly distributed between grid neighbours. Note that although employment status at the end of period *t*-1 might be different from that at the beginning of period *t*, capital holdings at the end of *t*-1 is always same as those at the beginning of *t* for each individual. Since individual optimal rule for capital at the end of current period *t* (equivalent to capital at the beginning of *t*+1) depends on the current wealth and current state , denoted as

. The next target is to derive the cumulative distribution function (CDF) of and it starts from writing wealth as a function of grids by inversing individual optimal capital decision, denoted by where has same grid points as . Individual budget and positive consumption imply and thus the mass of agents whose wealth is less than a certain value (actually a CDF value) should at least contain the mass of agents whose capital at the end of t is less than that certain value. Suppose is the *i*th grid point of capital interval and is the PDF of . The CDF of wealth is defined by

. Then the PDF of cross-agents at the beginning of next period can be easily calculated using the achieved CDF values of wealth distribution and the number of agents with a certain employment status. In their second stage, a PBM is used for the final solution associated with aggregate shocks. Young (2010) modifies Reiter (2009)'s algorithm when transit the PDF of a cross-agents distribution period by period on grids. In the first stage, given the grid points (i.e.) of , he uses individual optimal rules to find the values of with same employment status such that where and then obtain the PDF . Next, use individual policy rules, the achieved and the grid points of (same grid points as , i.e. also) to calculate . Define some new grid points as where the function is individual optimal rule. He sets a mechanism to decide whether allocate the mass of agents hold into the PDF at the end of period *t* by comparing the distance between and . The CDF can easily be obtained and follow same second stage as Reiter (2009) to yield .

Compared with the second generation, some third generation models may not inevitably lead to the APAG because one could separately describe dynamics of the cross-agents distributions using PDFs for individual behaviour solution. Although Algan et al (2008) also find the APAG in practice, it might be a result from a simple use of the first order moments as aggregate state variables, not the result of incapability. Reiter (2009) find high volatilities of aggregate capital in their model with both aggregate productivity shock and tax rate shock, which imply moments may plan a more important role in individual policy decision than that in Krusell and Smith's model. Moreover, compared with a large number approximation, the third generation models could efficiently remove the sampling-invariant issue. Algan et al (2014) compare the simulated fractions of unemployed agents close to the borrowing constraint when the economy moves from a recession to a boom with a large sample and those with a continuum of individuals. Obviously, the dynamics show significant uncertainty, inconsistent with the stylised fact.

Figure 1: Uncertainty Comparison between Continuum and Finite Sample



Source: Algan et al (2014) Figure 3, chapter 6 in Schmedders and Judd (2014) 3rd Edition

There are many other scholars who attempt to establish theoretical frameworks to show the long-run effect of inequality on growth, with no interest in whether model could generate similar micro-behaviours as reality. Nevertheless, most of them investigate how income inequality instead of wealth inequality influences aggregate growth. For instance, the earliest study can be traced back to Kuznets (1955) who initially raises a Kuznets Curve that aggregate growth is found to be a concave function of income inequality. As we are more interested in the wealth inequality effects, we only briefly review some relevant studies on wealth distribution or on income distribution but with ideas that could be easily applied to wealth inequality. These growth models are mostly developed from OLG models with some mechanisms of inequality effects on growth such as indivisible investment or labour, incomplete market barriers, "Median Voter Theorem" and security hazard, etc.[28]

Some scholars use investment or labour indivisibility to show the positive effect

[28] Some authors like Ehrhart (2009) and Voitchovsky (2011), etc. have more detailed classifications on the mechanisms of inequality effects. Since many studies in fact apply multi-mechanisms classified in their papers, we do not follow that way.

of inequality on growth. For example, Galor and Tsiddon (1997) consider an economy including many production sectors with different technologies but same goods. Each individual endowed with a unit of time chooses a producing sector *j* (can only choose one sector due to labour indivisibility) to supply constant proportion of time endowment as labour and the rest is the sector requirement for training. Individual labour efficiency in sector *j* relies on his ability and the ability of his parent in this sector (simply say “like father, like son”). They prove that there are an upper threshold of ability and lower threshold in each sector. If individual ability is higher than the upper threshold, his generation will choose a sector with higher technology while his generation will chose a sector with lower technology if his ability is less than the lower threshold. Generation stays in same sector if parent’s ability is between two thresholds. That is, given initial inequality, individual ability will be allocated more optimally to different sectors (high ability labours are allocated to sector with high technology) and thus aggregate technology will be enhanced and so will aggregate growth. Bhattacharya (1998) assume some individuals as capitalist can engage in investment with high return rate but also high risk while individuals with more bequests from parents could internally finance investment with lower cost which generates income inequality but bequests also make the poor easier to obtain external financial credit for investment setting-up (investment indivisibility). Hence aggregate capital level is enhanced by inequality and beneficial to economic growth. Barro (2000) raises an idea that inequality could stimulate growth to some extent because it can make some individuals accumulate sufficient capital to satisfy minimum quantity requirement of both production and human capital investment. Galor and Moav (2000) establish a model based on Galor and Tsiddon (1997), but with new human capital accumulation way which depends on individual physical capital and also parent's bequest. They consider a situation that marginal saving rate of the rich is generally higher than the poor. Hence initially inequality could increase aggregate capital stock and further increase aggregate growth. The poor could also benefit from this process because wage income is higher following the capital rise. This process is also called "Trickle-down" process. However, as economic development, marginal return of physical capital is gradually lower than that of human capital. Equality might enhance growth since credit restriction limits human capital accumulation of the poor.

Since 1990s, more scholars turn to support that inequality is harmful to economic growth. For example, Aghion and Bolton (1997) and Piketty (1997) highlight that investment barriers to the poor caused by inequality, given incomplete market is harmful to growth. They assume individuals can either only supply labour with no capital investment and very low return or invest both capital and labour as entrepreneurs with a capital requirement. Entrepreneurship investment, however, has a probability to succeed which is determined by labour working efforts. Capital market is assumed incomplete and if too many poor agents need to borrow for investment, interest rate is higher which will reduce the poor's incentives to make efforts (i.e. moral hazard). The rich as lenders could anticipate the poor's behaviours and set borrowing restrictions which is called "credit-rationing". Consequently a society with more unequal wealth distribution has lower growth due to lack of poor's effort incentives and redistribution could enhance aggregate growth by reducing borrowing needs for the poor to invest and further increasing their incentives. Some scholars like Dahan and Tsiddon (1998) consider that human capital development will also be impeded by inequality, leading to unequal skills and finally worsen aggregate growth. One important issue of this kind of models is that lack of effort incentives when agents are poor in fact is based on an implicit assumption that efforts can never make the poor become rich which is debatable. Some scholars use indivisibility to show a negative inequality effect. For example, Fishman and Simhon (2002) similar as Aghion and Bolton (1997) also consider inequality effects on entrepreneurship, but assume that individuals can be either entrepreneurs or workers but not both simultaneously. Entrepreneurship activity requires both capital and labour which is much easier for the rich than the poor due to incomplete capital market.[29] Individual's optimal allocation between consumption and bequests (savings) in each generation will lead to an extremely unequal economy given initial inequality, i.e. the rich will be always rich generation by generation while the poor will be always poor. Inequality may not distort aggregate growth in an economy starting with relatively equal distributions, but is harmful in an extremely unequal economy because many poor individuals do not have sufficient capital for entrepreneurship.

Many political economists, such as Alesina and Rodrik (1994), Persson and Tabellini (1994), Bénabou (1996), Aghion et al (1999) and Acemoglu and Robinson (2002), attempt to show negative inequality effects on growth by redistribution pressure caused by the Median Voter Theorem (MVT).[30] For convenience, we use Persson and Tabellini (1994)'s model as an example. Given budget constraints in youth and in age and where is political instrument like tax. Individual in youth solves the optimal problem by allocating current consumption and future consumption . Individual in youth has an income endowment where *s* is constant skill and is individual specific

[29] There are two types of products, final goods and intermediate goods used for final goods production. Entrepreneurs employ labours and capital (his own or borrowing) to produce intermediate goods.

[30] Bénabou (1996) uses MVT in his first model.

multiplier with zero mean and non-positive median.[31] The solved individual optimal rules for , and depend on predetermined , aggregate and parameters. Aggregation on individual production implies that aggregate income growth equals aggregate capital accumulation growth which depends on parameters, especially political instrument . The optimal for individuals solved by first order condition in terms of finally is a function of steady-state interest rate *r*, skill and individual exogenous and the equilibrium is the one when where is the median individual specification. Therefore the equilibrium aggregate growth rate could be found positively correlated to the median individual specification which measures inequality in this model. That is, aggregate growth is higher if individuals are more equal, i.e. is higher. This kind of models has two issues. First, higher (closer to zero) may not represent higher equality. Suppose population consists of 5 people. Consider scenario I where are with mean 0 and median . In scenario II, are with mean 0 but median 0. Although median increases to 0 in scenario II, we can hardly say scenario II has higher equality. For simplification, suppose average skill . The richest person takes only 10% out of total income in scenario I while takes 24% in scenario II. Second, as the argument by Zweimüller (2000) and Kucera (2002), inequality has no significant effect on taxation policy or transfer regime in many across-country statistics. Redistribution pressure by the poor might be uninfluential if the rich has sufficiently high political power (Stiglitz, 2012).

Some others like Benhabib and Rustichini (1996), Keefer and Knack (2002) and Burdett et al. (2003) consider security reason due to inequality. An extremely unequal

[31] Some authors like Bénabou (1996) use nonlinear production and next period individual human capital (skill here) depends on current production and exogenous, but they share same essentiality.

wealth distribution results in illegal return rate for the very poor much higher than their legal return rate which will induce an increase in crime which will discourage the incentives of capital accumulation for the rich. Different from many scholars who apply OLG models, Benhabib (2003) consider a model where two types of infinitely living individuals are classified by their individual productivities. They compare two games, cooperative game where a central planer (e.g. a government) optimises social utility which is a weighted sum of individual utilities by controlling individual labour inputs and consumptions, and a non-cooperative game where individuals solve their own optimal problem by controlling own labours and consumptions. Cooperative equilibrium is found preferred to non-cooperative equilibrium for central planner. Given the implicit assumption that saving ratio across agents is constant (by a special preference form) and consumption and leisure are complement, they show that high inequality (more capital allocated to agents with higher productivity) results in a lower growth rate because agents with higher productivity make less effort. However, a completely equal allocation of capital is not optimal since agents with higher productivity who are extremely unequally treated will also make less effort (by policy rent-seeking in cooperative equilibrium). Hence the optimal situation is a moderate inequality which could stimulate agents with higher productivity. Foellmi and Zweimüller (2004) also classify infinitely living individuals to two types, the rich and the poor with constant population weights but different wealth endowments. Each individual is a monopolist to produce a type of. Individuals have willingness to buy goods which depend on individual wealth and thus the rich has higher willingness to pay higher. Consumption demand relies on both willingness and group size.[32]

[32] The consumption demand for one good equals the population size if its price is less than the poor's willingness while it equals zero if prices is greater than the rich's willingness. If the price is between two groups' willingness, goods will be shared by group sizes.

Individual productivity progress relies on individual innovation which only costs time input and innovation incentive is induced by consumption demand. They compare the effects of equality and inequality on the equilibrium growth rate and show that given initial unequal wealth allocation, whether inequality or equality is beneficial to growth depends on the redistribution method. If income or wealth is transferred from the rich to the poor with same population weights, then growth will be reduced by this equalisation because lower profits (the rich is unwilling to pay as high as before for some goods and the effect on the rich is greater than the wealth effect on the poor) drains innovation incentives. However, if the redistribution behaves by increasing the population weight of the rich, growth may rise or fall because more individuals now could afford goods with mid-priced goods (i.e. "market size effect") while fewer agents could afford goods with extremely high prices (i.e. "price effect").

These models are generally used as theoretical support for empirical regression, different from the three generation HAMs which are generally used for simulation. The estimation methods in relevant empirical studies include cross-country regression for which least squares estimation methods are frequently used and panel data regression such as pooled least square, general GMM, difference GMM and system GMM. Cross-country regression generally regresses average growth rate in the time horizon of data set on regressors at the beginning of the time horizon. Difference GMM estimation firstly takes difference on original regression equation to get rid of the (country specific) fixed effects and use lagged level variables as instruments of differenced variables, generally taking the form of where contains all the other regressors. System GMM estimates both level and difference equations respectively using lagged difference variables and lagged level variables as instruments and the combination of both moments (weight matrix) yields system GMM estimators. The debate on whether inequality could stimulate or impede economic growth has no consensus until now.

One common issue of most existing studies is that income inequality instead of wealth or capital inequality is employed as a regressor due to data limitations which might lead to misspecified effects on growth because wealth and income distributions do not always have same dynamic trend (Rodríguez et al. 2002). Next we briefly review some typical empirical studies despite the use of income inequality in some papers.

In the early debates, more empirical studies support the negative inequality effect on growth. For example, Alesina and Rodrik (1994) conduct cross-country OLS regressions regressing average aggregate growth rate from 1960 to 1985 (also consider 1970 to 1985) on GDP, primary schooling index, income inequality measured by Gini-coefficient of income, wealth inequality measured by Gini-coefficient of land and a dummy variable describing democracy all in 1960. They find a significantly negative effect of wealth (also income) inequality on aggregate growth.[33] Persson and Tabellini (1994) also use cross-country OLS regression (2SLS estimation in "sensitivity analysis") in two different time horizons of data set, 1830-1985 and postwar respectively. Besides same regressors as Alesina and Rodrik (1994), they introduce development level as a regressor which is measured by the gap between GDP in a certain country and the greatest GDP among all the countries. Additionally, they use middle class income share as income inequality indicator instead of Gini-coefficients in Alesina and Rodrik (1994). Growth is found significantly negatively rely on inequality in all sample countries and in

[33] TLSL regression is also considered when wealth inequality measured by land is excluded.

democracy countries, but not in non-democracy economies. Perotti (1996) also conduct cross-country OLS regressions (WLS is used to test robustness for a certain regressor) from 1960 to 1985 by regress average growth rate on regressors in 1960. Different from studies above, they measure income inequality by income share of middle class (shares of 3rd and 4th quintiles). They also consider investment deflator measured by PPPI as a regressor and separately consider human capital of male and female by their years of schooling. Lastly, the population share of people aged over 65 as a proxy of fertility is considered in one regression. They finally find significantly positive effects of inequality on growth in whole sample and also in democracy economies for all regressions. The effects are still positive but insignificant in non-democracy countries.

Barro (2000) makes an early use of panel data with time horizon 1965-1995 which is divided into three 10-year episodes. 3SLS with random effects is used in pooled regressions with inflation, schooling years and some lagged variables as instruments.[34] He respectively considers quintile shares and Gini-coefficients of expenditure rather than income following Deininger and Squire (1996) as income inequality indicators. He finds that income inequality has positive effects on growth in developed countries while negative effects in undeveloped economies and insignificant effects in the whole sample. Knowles (2001) conduct a study on the basis of Barro (2000) by trying different but consistent across countries inequality measurements. He argues that some previous studies result in misleading findings because inconsistent data on inequality are used. For example, some countries statistic Gini-coefficient of income using gross income (essentially before redistribution) but some use expenditure (essentially after redistribution). He concludes that income

[34] Country specific effects are considered as fixed effects only when evaluate inequality determinants.

inequality has a significantly positive effect on growth if it is consistently measured by income share of the 3rd quintile, but a negative effect if inconsistently measured by Gini-coefficient in the whole sample. A significantly negative effect could be found with consistent measurement in high-income economies only if an ex post redistribution inequality is used, i.e. expenditure inequality. Forbes (2000) adopts a difference GMM in a panel data fixed effects regression using data from 1965 to 1995 with each 5 year as an episode. His sample does not include undeveloped countries like African and sub-Saharan countries. With similar regressors raised by Perotti (1996), his estimation results show significantly positive effects of income inequality on growth in those middle and high income economies. Bagchi and Svejnar (2015) use panel fixed effects regression (including both country effects and time effects) using IV (essentially similar with Forbes (2000)'s estimation method) with four 5-year time intervals from 1987 to 2007. They consider the effect of wealth inequality which has three measurements, the share of capital held by billionaires over national capital stock, billionaire capital per GDP and billionaire capital over population with two instruments available, billionaire capital per GDP over per capita income and exchange rate. In addition to usual regressors, they also introduce initial poverty and a dummy variable describing whether a billionaire exists as regressors. Negative wealth inequality effects are found, although income inequality effect is insignificantly positive. Furthermore, poverty has an insignificant effect on growth together with income inequality but a significantly positive effect if income inequality is excluded.

Traditional least square and GMM estimations with lagged level variables as IVs generally lead to bias due to unobservable country effects. However, difference GMM used by Forbes (2000) also ignores some cross-country variations. Some inequality indicators within a country do not have high variants across time, such as human capital and income inequality indicators. Most volatiles of these indicators in a sample of countries come from the cross-country variation and thus difference GMM will eliminate the cross-country variation (Voitchovsky, 2005 and Castelló-Climen, 2010). Compared with difference GMM, system GMM is more efficient. Deininger and Olinto (2000) make an early use of system GMM estimation in a panel regression using data from 1966 to 1990 with 5-year episodes. Particularly, they consider the effects of both income equality and asset inequality measured by Gini-coefficient of land on growth. They find asset (land) inequality is significantly harmful to growth in both whole sample and rich countries in spite of that income inequality has significantly beneficial effects in rich countries. Voitchovsky (2005) uses system GMM estimation in the panel data from 1975 to 2000 with same episode interval in high and middle income economies. He focuses on effects of different inequality measurements, such as Gini-coefficient, share ratio of 90% over 75% (named by "top ratio") and share ratio of 50% over 10% (named by "bottom ratio"). If Gini-coefficient is the unique inequality measurement, the effect is insignificant. If Gini together with share ratios are used to measure inequality, findings are significant where top ratio has positive effects but bottom ratio has negative effects. His finding on the poverty effect is contrary to Bagchi and Svejnar (2015)'s finding. Castelló-Climen (2010) develops Barro (2000)'s analysis by using system GMM with data from 1965 to 2000. His basic finding is as Barro (2000)'s that both income inequality and human capital inequality have positive effects on growth in rich economies but negative effects in poor countries. Additionally, he compares seven different time horizons and also different inequality measurements and finds that the sign of inequality effects is more likely to be influenced by sampled time horizons. Ostry et al. (2014) estimate the panel data from 1960 to 2010 across maximum 153 countries with system GMM to study both the sign of inequality effects on growth and the duration of the effects. In particular, they focus on the ex post redistribution inequality which is called "net inequality" similar as Knowles (2001). They find net inequality has a significantly positive effect on growth while higher net inequality also increases the probability of a short duration of growth both OECD and non-OECD economies. Halter et al. (2014) investigate both short term (5-year episode) and long term (10-year episode) effects of inequality on growth from 1966 to 2005 with system GMM. In the whole sample, they find inequality has a significantly positive effect in short term while a negative effect in long term. Inequality has a significantly positive effect in high income economies in both short term and long term. They also conclude that the estimated direction of inequality effect is affected by both estimation method and sample. More precisely, difference GMM estimation is more likely to show a positive effect while system GMM estimation is more likely to give a negative effect.

As shown above, most existing studies emphasise on the linear relationship between inequality and growth. However, the relation is possible to be nonlinear which is also what we believe and are interested in. Banerjee and Duflo (2003) argue that a negative wealth inequality effect on growth in some empirical studies is found because only linear relation is considered. Hence, they conduct a panel system GMM from 1965 to 1995 by considering inequality (Gini) square as a regressor and find a "U-shape" effect that growth is worsened as long as inequality changes no matter the direction. Kolev and Niehues (2016) also consider the nonlinear effect of net inequality with system GMM in a panel from 1960 to 2010 across 113 countries where the key regressor is square of ex post redistribution inequality. Their first finding is that linear inequality effect (regardless of the nonlinear relation) is significant at 5% confidence level only if no other regressors except GDP are considered no matter which estimation method like OLS, fixed effects pooled least square, difference GMM and system GMM, is used in OECD countries (1985-2010). Contrarily, nonlinear effect is significant with regressors like human capital and investment included in the whole sample (1960-2010). Secondly, nonlinear inequality effect is likely to be negative whereas linear effect is likely to be positive especially in post-communist OECD countries. They also evaluate the effects of different taxations on growth and find most have negative effects, especially the marginal income tax rate which support the argument against that redistribution contributes to economic growth.

Neves and Silva (2014) also review some empirical papers on examining different theoretical mechanisms of inequality effect which are omitted in this paper. As more studies indicate that it is difficult to have a consensus on the inequality effect across different countries, this paper prefers to a within-country study instead of a cross-country study on this topic. Furthermore, both lack of volatilities of some inequality indicators and the complex interactions between inequality and growth limit the use of regression analysis. Hence, in this paper we attempt to use a reliable model which will firstly be tested by comparing simulation and actual data to analysis inequality effects on growth.

3 Model Setting

This section illustrates our endogenous growth model with a growth mechanism originating from Minford et al (2007) that individual entrepreneurship incentives could affect productivity growth and further aggregate growth. In their model, entrepreneurship incentives depend on its cost described by an exogenous penalty rate. We introduce heterogeneous agents into this growth model and relate the penalty rate as well as entrepreneurship incentives to individual wealth holdings. Hence wealth distribution could endogenously affect economic growth. This idea comes from the fact that rich people who have adequate wealth to support surplus activities generally have more incentives to do entrepreneurship or innovation than the poor whose priority is survival. We are not going to deny the success of a few entrepreneurs from impoverished backgrounds. However, it is easy to enumerate more successful entrepreneurs born in rich families or the middle class, such as John Pierpont Morgan, Rupert Murdoch and Warren Edward Buffett, William Henry Gates III and Steven Paul Jobs etc. Levine and Rubinstein (2013) investigate who are more likely to become Schumpeterian entrepreneurs measured by the incorporated self-employed using NLSY79 data in US and find that more entrepreneurs come from well-educated and high-income families. We regard in spite of this evidence as merely suggestive of how an attractive structural model can be specified relating inequality and growth. Our aim here is to develop a full and rigorously-based DSGE model and to test it by the powerful method of indirect inference to check on how well it matches the observed data behaviour of a suitable major economy-here that of the UK.

3.1 Individual behaviour

We firstly consider individuals in the economy. Suppose the whole population is comprised of a finite number of groups. Agents in each group behave as a representative agent. To simplify, assume only two groups (individuals) with constant population weights denoted by and . Each individual has a centralised optimal problem where individual utility function takes same form across agents as well as individual budget constraint. Each individual has same time endowment, normalised to unity, which could be allocated to labour input , entrepreneurship time (incentive) and leisure. The idiosyncratic shocks in our model are individual labour supply preference shocks and individual consumption preference shocks instead of individual employment shocks in traditional HAMs. Agents are assumed to be fully employed. Individual *i*'s utility takes CRRA form with three arguments, individual consumption , and as follows.

where and are utility elasticity of consumption and leisure. and are weights of consumption preference and leisure preference in utility. This benchmark model does not consider general tax which will be introduced in the redistribution section. Individual entrepreneurship has per unit time cost, denoted by and the total cost of entrepreneurship for individual *i* is then . Both perfect labour market and capital market are assumed to be perfect and thus individuals have a centralised income equal to individual production which can be used for consumption, capital accumulation and bonds purchase. Individual *i*'s budget constraint is now given below.

(2)

Since we have inequality generation mechanism illustrated later and perfect capital market assumption, borrowing constraint is not applied. In fact as Algan et al (2014) criticised, individuals could accumulate a relatively high wealth level so that borrowing constraint is rarely touched in practice. Preston and Roca (2007)'s utility penalty is also not used because no significant role of penalty is found as discussed in the literature section.

Individual *i* has a Cobb-Douglas production below which also determines an implicit real wage equal to marginal productivity of labour.

(3)

where individual productivity is non-stationary as in many stochastic growth models, but we give it an endogenous growth following the process (4) below which relies on the entrepreneurship time.

where measures the natural productivity growth and the shock is an aggregate productivity shock so that both individual productivities are related to aggregate productivity which is exogenous and equal to the Solow residual of an implicit aggregate Cobb-Douglas production.[35] We focus on the interior solution. That is, we assume neither nor is zero; otherwise the optimal rules will be equivalent to those in an exogenous growth model. We simply adopt identical parameters across agents to avoid that inequality is generated by too many exogenous factors different across agents. One could release this as a development.

Individual's first order conditions on bonds, capital holdings and labour, together with first order condition on current and future consumptions, yield intertemporal

[35] Since aggregate economy actually is an aggregation of individual behaviours plus some errors, there is no necessity to apply an aggregate production function. Instead, we apply an exogenous aggregate productivity to reflect the effects of aggregate production on individual ones.

Euler equation (5), optimal capital rule (6) and optimal rule for individual labour supply (7) respectively.

(5)

Given individual production (3) and individual productivity progress (4), the differential of in terms of is

Therefore first order condition of is[36]

Substituting out the first term in (8) by (7) yields

[36] Since the technological progress (4) is irrelevant to future shocks, the expectation could be removed from the term .

Equation (9) actually is an accurately optimal decision rule for . However, it involves an infinite number of expectation terms which makes it impossible to solve without approximation. We follow Minford (2007)'s way by approximating — as a random walk before steady state and the proof is shown in Appendix 1. Equation (9) can be written as

As shown in Appendix 1, is set to unity for simplification and this value will also be used in our empirical study. Then Use to simplify (10) as[37]

Equation (11) gives an approximately optimal decision rule for with a finite number of terms. According to the perfect labour market assumption, — is the individual implied real wage rate and we define — to rewrite (11) as

If individual is engaged in entrepreneurship, he will pay in each unit of time and he will also lose unit wage and thus is a scaled variable like a tax rate. We call it "entrepreneurship penalty rate" in rest of this paper. Substituting

[37] If one estimate different from but close to unity, the approximation could be used for simplification of (10) where is the steady-state gross growth rate of individual consumption to yield

out —— from (12) using (4) gives us the reduced equation of . However, it is not necessary to do so because is essentially an intermediate variable which could be substituted out from those equations involved in it. And one advantage of substituting out is that the difficulty of collecting data on it could be avoided. Similar as Minford et al (2007), by relegating — into the error term, (12) can be linearised as

(13)

where .[38] Note that the error term in (13) is an aggregate error instead of an individual one because firstly — is close across agents given the perfect labour market assumption and secondly measurement errors when collect data on individual productivity could be reduced and the aggregate error is related to aggregate productivity which is easy to collect by aggregate Solow residuals.

3.2 Entrepreneurship penalty rate

Now consider the form of individual penalty rate — which should follow the idea that the rich are more likely and capable to do entrepreneurship and policies are biased towards the rich to some extent. Hence we assume individual penalty rate is negatively correlated with the ratio of group average capital over population average capital — which also describes inequality despite that group capital share — is the inequality indicator in this paper. Group capital share is not used because it cannot make policy biased towards the rich due to the small population

[38] One could also linearise equation (12) directly by taking logarithm on both sides to achieve different and drift. However, there are only quantitative effects, but no qualitative effects of all the other parameters on

weight of the rich. We also assume current penalty rate depends on lagged instead of current capital ratio since policy in reality is pre-set.[39] Lastly we assume individuals can observe but cannot observe the exact relationship between and and thus is predetermined when individuals make their optimal decisions.

The evolution of is temporarily denoted as follows.

where and is the aggregate entrpreneuring penaly rate which is a time-stationary AR process which describes some common properties of individual processes. The empirical finding of Banerjee and Duflo (2003) and Kolev and Niehues (2016) indicates that inequality has significant nonlinear effects on growth where they consider the effect of inequality square. We hence set where represents a perfect equality while either greater or less than 1 implies inequality. The coefficient firstly aims to avoid the penalty policy too much beneficial to the rich as the poor generally has a greater population weight relative to their average income share. For example, this relative ratio in our data set for the poor is while for the rich is . Second, this setting simplifies aggregation. is finite in spite of its quadratic form and so is because has both upper bound and lower bound due to the capital aggregation equation. According to

[39] Note that capital level at the beginning of period t is denoted by

the derivation in Appendix 2, — has a minimum at perfect equality where —

— , penalty rates are identical across agents. It is also shown that the expected aggregate growth next period has a reduced form depending on current aggregate output, current aggregate capital and both current and lagged inequality. Capital inequality on the long-run growth rate is clear that minimised growth rate is reached when distribution is always perfectly equal. The temporary effect is more complicated due to the opposite directions of current and lagged inequality effects. One important reason is that labour input might be temporarily reduced by an increase in entrepreneurship incentives because of the total time endowment allocation which will have a currently negative effect on output while the rising incentives have a stronger positive next period instead of current period. This short-run behaviour is insignificant when the existing inequality is relatively high like in reality, but significant if capital distribution is close to equality. More details of the relevant analysis are provided in the empirical section.

3.3 Aggregate Economy

Since the individual variable measures the representative value in each group, the aggregate variable in the whole economy is the weighted sum of individual ones.

(15)

(16)

(17)

To simplify, we assume UK is a closed economy; although international trade in

UK takes a considerable share of national economy. This is equivalent to say net export is captured by market clearing error. We also assume government spending entirely funded by individual entrepreneurship charges and they can be substituted out from the aggregate market clearing condition below.

(18)

where the simple assumption on the perfect bonds market get rids of the aggregate bonds from the aggregate market clearing.

An aggregate production function here is not directly employed to create aggregate output, but used to back out aggregate productivity shock which follows an I(1) process

(19)

where is close to 1 which will be estimated later. captures the deterministic trend of non-stationary .

3.4 Linearised model equations

There are 2 individual Euler equations, i.e. (5). One is chosen as the path of real interest rate and the other determines the path of its individual consumption, yielding the linearised equations (20) and (28) below. The consumption path of the rest individual, equation (29) below is obtained from consumption aggregation equation (17). measures individual consumption preference error. Output aggregation equation firstly can be log-linearised to —— —— and since and are steady-state values of output per capita in group one and aggregate output per capita, —— is then the steady-state output share of group one, denoted as . Output share of group two is . Linearised output aggregation and capital aggregation equations are (21) and (22). Linearised market clearing condition (23) gives the path of aggregate consumption. Individual capital decision rules are linearised to (26) and (27) by substituting out expected consumption using Euler equation (5).[40] To linearise labour equation (7), we take first order Taylor expansion on both sides regardless of constant terms to obtain

. Using approximation − and − based on the actual aggregate data yields .

Since plays no more roles than an intermediate variable, substituting it out using equation (4) and (13) finally yields (30) and (31) where error measures individual labour supply shock.

Three more points need to be clarified. Some linearisation is conducted around the steady state in despite of non-stationarity because the steady state in our model is the one relying one non-stationary predetermined variables or exogenous shocks ("conditional steady state") instead of a constant value which is explained in the Indirect-Inference section. Secondly, some constant terms when log-linearise are omitted because of our numerical algorithm also illustrated later in Indirect-Inference section which has a "Type II Fix" procedure to capture the omitted constant terms. Lastly, individual bonds are helpful to derive the real interest rate and Euler equations, but are removed from equation list primarily because they take small share over individual capital resource allocations which we are not interested in and also because

[40] We also use the approximation —— ——. If one does not use this approximation, then will be replaced by .

their applications will bring difficulties in individual data collection. However, we still use individual budget constraints when derive first order conditions and also when calculate terminal conditions where individuals have .

List of linearised equations

(20)

(21)

(22)

(24)

(25)

(32)

(33)

4 Indirect Inference

In this section, we illustrate the methodology, Indirect Inference, applied to test our model and also estimate key parameters, involving 4 subtopics, the basic idea and statistic we use, auxiliary model, simulation method and estimation method.

The concept of Indirect inference (In-If) are firstly raised by Smith (1993) and Gourieroux et al (1993) in order to estimate structural parameters which are difficult to be directly estimated from a structural model. To briefly summarise their idea, I resort to a simple structural model, an MA process given by Davidson and MacKinnon (2004). To estimate structural parameters and , one could set an auxiliary model which could well specify the structural model, for example an AR whose parameters are simpler to be estimated.[41] Id-If has an advantage that the choice of the auxiliary model (or parameters) is quite flexible. Id-If is asymptotically efficient as maximum likelihood as long as the auxiliary model is correctly described (Keane and Smith 2003).

[41] Smith (1993) and Gourieroux et al (1993) do not assume an auxiliary model but auxiliary parameters and a criterion function (e.g. a likelihood function) with some restrictions. I use the present the brief way to save words due to their complicated notations.

Suppose . Then auxiliary parameters can be estimated using OLS. The properties of imply .

Use and in the structural model to substitute out and from the two restrictions using those in the structural model yields two "binding functions" below which describe the relation between auxiliary parameters and structural parameters.

Then estimators on structural parameters could be achieved from and . More generally, however, analytical binding functions are too complicated to compute and then simulation method could be used. Define the structural parameter vector and the auxiliary parameter vector as and respectively. Following Gourieroux et al (1993)'s idea, firstly one can obtain the estimated auxiliary parameters using observations. Secondly, one could stimulate *S* samples using structural model with a given . Then replace observations by each sample of simulated data in the auxiliary model to estimate auxiliary parameters. Denote the achieved estimators based on simulated data as depending on the given . Lastly, the Id-If estimator of is the one such that and − are closest. Their simplest but time-consuming way to find optimal is calibration until and are convergent. Alternatively, one could solve the problem to yield the optimal estimator where is a self-chosen positive definite matrix and is one converges to .[42] More importantly, given true values of structural parameters as , it is proved that − asymptotically follows a normal distribution with zero mean.

[42] See the optimal choice of in Gourieroux et al (1993) Appendix 2

4.1 Wald Statistic

The asymptotic property of the early Id-If estimators provides more widespread applications of Id-If, like model testing. Le et al (2011) make an early contribution to test a DSGE model using Id-If and we also follow their way. A DSGE model (usually a linearised one) can be represented as a VARMA or VAR( ) with structural parameter restrictions, although some issues might possibly occur, like linearisation errors, multi-solutions and "stochastic singularity" (i.e. too few shocks). Then a VAR( ) could be further represented to a finite order VAR(p) or even a VAR(1) if all the endogenous variables are observable or some conditions for state space are satisfied (Dave and De Jong 2007; Giacomini 2013; Wickens 2014). Thus we could use VAR as the auxiliary model. An unrestricted VAR instead of a restricted one will be used for two reasons. Firstly, a restricted VAR requires an analytical solution which are difficult to be solved for complicated models. Secondly, as long as structural parameters are well identified, unrestricted VAR does not have identification problem. Moreover, even if structural parameters are not identified, the effect of using an unrestricted VAR also depends on the auxiliary parameters chosen. Consider Canova and Sala (2009)'s 3-equations model below composed of IS curve, Phillips curve and monetary policy.

The analytical solution is where , and are column vectors respectively comprise endogenous variables, constant terms and shocks. *P* is a 3 by 3 coefficient matrix which does not contain , and . Constant vector in a restricted reduced form reflects three restriction equations on structural parameters. However, the internal relations disappear when consider an unrestricted reduced form. That is, the structural parameters are not well identified in the solution. But the fact is that Id-If does not use a reduced form but the original structural model to simulate data. Namely, although , and disappear in the reduced form (suppose as the auxiliary model), the constant differs when , and values change and thus if is also considered into the auxiliary parameters, the bias could be neglected. Nevertheless, under-identification may occur when a difference form of reduced solution (e.g. VECM) is used for non-stationary shocks. In fact, DSGE models are quite often over-identified instead of under-identified (Wickens 2012). It should be noticed that although an unrestricted VAR is chosen as the auxiliary model, the data simulated from the structural model actually follow a restricted VAR. Suppose now the auxiliary model is VAR(1). Given null hypothesis of the Id-If model testing (i.e. model is true), to generate a testing statistic, we firstly create a vector which contains coefficients in the estimated VAR coefficient matrices and variances of the VAR residuals. This is the auxiliary parameter vector . A sample of is obtained after *S* times simulations (procedure will be illustrated later). Given the asymptotical property of , Wald statistic

asymptotically follows an distribution with degree of freedom where is the variance-covariance matrix of and is the number of elements in . Note that a full reflects the interactive volatilities of variables implied by a structural model and thus a Wald test using a full is a joint-distribution test. Meenagh et al (2008) conduct an experiment to compare two

Wald statistics, one with a diagonal and the other with a full , using a bivariate VAR(1) as the auxiliary model includes interest rate and inflation. They find model hard to be rejected using Wald with a diagonal .[43] Therefore to keep an adequate testing power, we also calculate Wald with a full . Then a simple way to evaluate the model is to compare the calculated Wald with a certain critical. But this way is more affected by the simulated distribution. A more usual way is to check the allocation of this Wald, denoted by (when is obtained using the actual data) in the simulated distribution which is composed of *S* Wald values where each is calculated when is obtained using the s^{th} sample of simulated data. One can sort from lowest to highest and will be rejected in a 95% confidence interval, if where is the 95^{th} percentile value of the sorted . One issue of this way is that has the possibility to be greater than all . An alternative way is the Mahalanobis Distance (MD) which equals . As , asymptotically follows . Following Le and Meenagh (2013) that as is close to a *t* distribution, we define a "transformed MD" below and compare it with the critical value of *t* distribution with a large df on a chosen confidence interval.[44]

where is the critical value of a one-tail t distribution on the *c*% confidence interval

[43] See (Meenagh et al 2008) Section 2.2.1; Figure 1

[44] In fact, we all use an approximation because the limiting distribution of is a *t* distribution only if for a large degree of freedom, say greater than 30 (Green 2002 Appendix B).

and is the c^{th} percentile value of the sorted .

4.2 Auxiliary Model

One should note that VAR is not the unique alternative for the auxiliary model but a proper one because we are more interested in whether the model could generate typical behaviours or interactions of data instead of closest data to the observations. Some others like IRFs and moments could also be used. Another discussible point is what orders (lags) of VAR and which VAR coefficients we should use. Le et al (2011) define two types of Wald statistic based on different use of VAR coefficients, a "Full Wald" where a vector composing all the VAR(1) coefficients and variances as the auxiliary model is used to calculate the Wald statistic, and a "Direct Wald" where some coefficients and variances we are interested in are used for Wald statistic. They firstly test three models, the Smets-Wouters model, a New Classical model and a hybrid model which is an average of NC and NK using a Full Wald and all the three are significantly rejected. Then they test the hybrid model using a Direct Wald. Despite that the model is still rejected, the behaviours of inflation, output and interest rate perform well and the model could provide some beneficial policy implication. That is, although Full Wald test is more powerful, it may reject some models which are moderately false. Le et al (2015) summarise the result of Monte Carlo experiments on the testing power comparison among different auxiliary VAR(1)s when test the S-W model. Model has an 83.5% probability to be rejected if 3variables (thus 9 VAR coefficients) are used to calculate Wald while the rejection probability increases to 96.6% when 5 variables (25 VAR coefficients) are considered. As an appropriate number of variables included in a VAR with order 1 could sufficiently have a great power, we could start with a Direct Wald with order 1 in an auxiliary VAR or VARX (transferred from a VECM) and probably increase the order of VAR if the test power is low.

As this paper aims to study the dynamics between the capital inequality and determinants of the aggregate growth in our model are capital inequality, aggregate output and aggregate capital as shown in Appendix 2 lack of rigorousness, the auxiliary model should at least contain aggregate output, capital inequality and aggregate capital.[45] Furthermore, since the aggregate productivity captures an aggregate production function which is lack in the model equations and it could also affect the individual productivities, to evaluate whether the relation between aggregate growth and aggregate productivity reflected from the actual data and that generated by our model are consistent, we introduce this non-stationary exogenous variable in the auxiliary model. Due to the non-stationarity, a VECM is preferred to a VAR. Meenagh et al (2012) show that the solution of a DSGE model with non-stationary variables can be approximated as a VARX (VAR with non-stationary variables X). Suppose the structural model is where , and are respectively a vector of endogenous variables, a vector of non-stationary variables and a vector of i.i.d. shocks. Given both and are non-stationary (because depends on), there should be cointegrations which requires both are I(1) and then a short-run equilibrium (stationary shocks are all zeros) could be written as the form of (this is also the form of terminal conditions we use). Combined this with the I(1) process of , they show that the long-run solution for comprise a deterministic trend and a stochastic trend. Then the solution could be approximated as the form of , a VARX depending on the

[45] Aggregate growth is the difference in logarithm of aggregate output and thus in the auxiliary VARX, series of aggregate output could implicitly describe series of growth.

non-stationary , deterministic trend and stochastic . Hence we use VARX(1) as the auxiliary model and estimate coefficients by OLS. Then a 12 by 1 coefficient vector including 9 VARX coefficients of the lagged endogenous variables and 3 variances is used to calculate the Wald.

4.3 Simulation

A crucial procedure of our test is simulation which could be carried out by following steps.

4.3.1 Back out Residuals

For convenience, define equation residual as the difference between the LHS value and the RHS value of one single equation for a certain variable based on actual data. Given structural parameter values, residuals in the equations with no rational expectations (RE) are directly backed out by LHS-RHS. For equations with RE, firstly following McCallum (1976)'s way to predict expectations using instruments, we estimate a VAR of variables with expectations (i.e. choose lagged variables as instruments) and set the fitted values one-period ahead as their expectations. Next, residuals are also equal to LHS-RHS.

4.3.2 Obtain Structural Errors and Innovations

Most equation residuals achieved using actual data are non-stationary due to deterministic trends except the exogenous aggregate productivity (Solow residual of an implicit aggregate production function) whose non-stationarity comes from both unite root and deterministic trend.[46] More explicitly, the aggregate productivity

[46] It might be argued that most errors should be stationary if the model is a true endogenous growth

follows ARIMA(1,1,0) with a drift while all the others follow ARMA(1,0) with deterministic trends. We firstly take difference on aggregate productivity and regress on lagged difference and constant to get its OLS residuals which are used for structural innovations for and the estimated constant is set as the Balanced Growth Path (BGP) rate. The structural errors for are the backed out aggregate productivity minus the estimated time trend which follows process . Next for other equation residuals, we regress each on a time trend to get the fitted value, denoted as , which is our time-detrended equation residual. Then regress on to obtain the stochastic process and the residuals are structural innovations for the error .

4.3.3 Solve the model

Due to the non-stationary endogenous variables, and , the steady state in our model is driven by non-stationary variables and thus we use Fortran with some numerical iteration algorithms instead of Dynare with stationarisation around a certain steady state to solve the model. We follow Le et al (2012)'s numerical solution method which originates from the one to solve the Liverpool model, like introduced by Minford et al (1984). To my knowledge, the algorithm works as follows. Suppose we have observations in N periods and we need to set a number of forecasting periods (explain this later), say F. First of all, we solve the model using actual data and structural errors (without innovations). Starting with initial values of rational expectations (RE) from period 1 to F, we calculate the values of variables in next period among which the variables we want to forecast are denoted by . Now we use

---

model which could capture the growth without exogenous drivers. One should note that structural errors depend on structural parameters whose values might be mis-specified and thus we follow the described way to let actual data and the current parameter values decide stationarity of errors and whether model could generate sufficient growth driving endogenously.

the values of variables in next period as their predictions in current period and thus we can have F periods predictions for each RE variable, denoted as . Next, calculate the difference between our prediction value and the given RE value in each period and we need all the differences less than a tolerance level. If the gap is greater than the tolerance level, then use a certain algorithm to renew the given RE from initial values to predictions (like a weighted sum and the present algorithm used is "Powell-Hybrid" searching method which has a high convergence speed) until predictions are close enough to the newest RE values and then we can say the model is solved in the first period.[47] Additionally all the differences between the values of variables when model solved and the actual data are saved in all the F periods (this step is called "Type II Fix"). Afterwards, the algorithm uses RE values one period ahead to repeat the steps above to obtain predictions and solve the model in the second period. We need to solve the model in all the N periods which require use to collect data in totally periods. In practice, we set 50.

One important thing for numerical methods is the multi-solution. To find a unique one, we set a terminal condition which represents model will be in steady state after a terminal period T. Although T should be infinity in principle, it has to be a finite number in numerical solutions. Both Matthews and Marwaha (1979) and Minford et al (1979) find solutions are not sensitive to the choice of T. We set T equal to the number of last forecasting period in practice, i.e. N+T. In our model terminal condition is the steady state solution for RE variables with stationary errors all zeros. Since individual productivities are non-stationary, RE variables will depend on the last period individual productivities in steady state and terminal condition can vary,

[47] There are several options of algorithms to renew RE, like the Gauss-Seidel algorithm used in an old version.

given different structural parameter values.

4.3.4 Bootstrap

After base run, we bootstrap innovations. Firstly define a time vector to ensure all the errors in each period will be randomly chosen together due to the interactive volatilities of errors. The bootstrapped samples in general do not follow same distributions as the original samples if these innovations are not i.i.d. However, we do not release the i.i.d. assumption. What we do is just to consider some occasional co-movements occur in certain unobservable environments. When bootstrap, we should not ignore those environments and thus randomly draw all innovations in same period together. The method could also work if this is released. Then we use a random generator to randomly draw elements from the time vector to yield a bootstrapped time vector with same dimension, also resulting in a bootstrapped sample of structural errors. We bootstrap structural innovations instead of randomly drawing innovations form assumed distributions because firstly the interactive volatilities as described in the previous note-foot can hardly be realised using single distributions. Secondly, the achieved innovations might imply these values have higher realisation probabilities (same idea as the likelihood), which cannot be reflected by random draws from given distributions. Furthermore, what distributions should be used if use them. It is true that bootstrap is based on the large sample theorem where the original sample of structural innovations has the same distribution with population structural innovations and the sample size in our model is insufficiently large. Le et al (2011) firstly conduct Monte Carlo experiments to evaluate the accuracy of innovation bootstrap by firstly setting a normal distribution for the original structural errors (equivalent to a distribution of original innovations), randomly drawing innovations and bootstrapping samples to calculate Wald for each sample.[48] It is found that small sample bootstrap is also accurate, although bootstrap is likely to "under-reject" and is more accurate in large samples. Additionally, they also verify that Id-If Wald based on bootstrapped simulations converges to an distribution as the sample size rises using a Monte Carlo experiment.

4.3.5 Simulate data and Calculate Wald

The bootstrapped innovations are used to renew structural errors following their stochastic processes achieved in 3.2. Then the model is solved following same procedures in 3.3 with two differences that innovations now are non-zero and the "Type II Fix" residuals are added to equations in each period to diminish unobservable uncertainties. The simulated outputs are variable values when model is solved. This simulation is called "Full simulation", different from a simulation for IRFs where innovations are non-zero only in the starting period (regardless of the lagged periods preparing for base run and simulation). Before calculate Wald, we add back BGP to simulated data except some stationary variables, like the real interest rate and individual labours. Since we use the estimated coefficient of deterministic trend of aggregate productivity as BGP (note that this is a steady state time trend coefficient, not the true BGP growth rate and the latter for actually is ——), the BGP growth rates of output, capital and consumption (both individual and aggregate) are all ———— while the BGP growth rates of individual productivities are both —— same as aggregate productivity.[49] Then the original simulated data could be

[48] See details in Le et al (2011) Section 6.1

[49] Steady-state process of the aggregate productivity is and thus its BGP growth rate

amended by multiplying a gross BGP growth rate in each period. Note that the net BGP growth rate generally is quite small for quarterly data, e.g. in our model around the range 0.002-0.004 and thus simulated data differ slightly. Moreover, BGP has almost no effects on Wald computation. However, we keep this step in order to compare simulated data with actual data which implicitly contains a time trend. Lastly, Wald and the Transformed MD are calculated after we collect S (1100 in practice) bootstrapped simulations.

4.4 Indirect Inference Estimation

Given parameter values, previous Id-If Wald test could evaluate whether the model is true, but in fact it could not tell us whether a rejection is due to the false of functional forms or due to inappropriate parameter values. Fortunately, we could firstly assume functional forms are correct and test the model given continuously adjusted values of parameters until the model cannot be rejected (we call the different models generally if parameter values are different). Our target is to find the parameter values such the transformed MD (TMD) or Wald is minimised and this procedure of searching for optimal parameters is called Id-If Estimation.

The Id-If estimation involves a searching algorithm which selects parameter values from one group to another. To avoid the trap of local optimum, we use Simulated-Annealing method (SA) with procedures given by (Johnson et al., 1990).

1. We start from initial parameter values and conduct a full Id-If Wald test with all 5 steps described in the previous section to obtain (or use Wald, so

---

is —— ——. Steady-state implicit aggregate production is where Y and K have same BGP growth rate and N has zero growth rate. Hence the BGP growth rate is ———— for both Y and K.

yield).

2. Then we randomly generate a neighbor parameter vector and follow same steps to obtain (or).
3. If , is chosen starting parameters and we repeat the two steps above to search for a better one. If , we do not surely reject to move towards , but calculate an "acceptance probability" (ACP) which provides a probability to move. Namely, we always surely move to a better new choice, but we still have a certain chance to move towards a worse choice, which is our mechanism to escape from a local optimum. The ACP generally takes the form of where *T* is called "temperature" which usually starts from unity and decreases after each estimation iteration. ACP could follows Gibbs equation in practice where . We randomly draw a value *u* from a uniform distribution in each estimation iteration. If the calculated , we move to ; if else, we stay at . Note that the process of T implies that it would be more difficult to move towards a worse choice as more iterations done.
4. Repeat steps above until we find an acceptable solution for or reach a maximum number of iterations.

It is an advantage of Id-If that estimation and test are not independent. Id-If estimation tells us the optimal parameters such that the "difference" (key characteristics) between simulated data and the actual data is minimised while Id-If test tells us how good this minimised difference is. Suppose the minimised difference can still be strongly rejected. One could believe that the functional forms of the model are problematic.

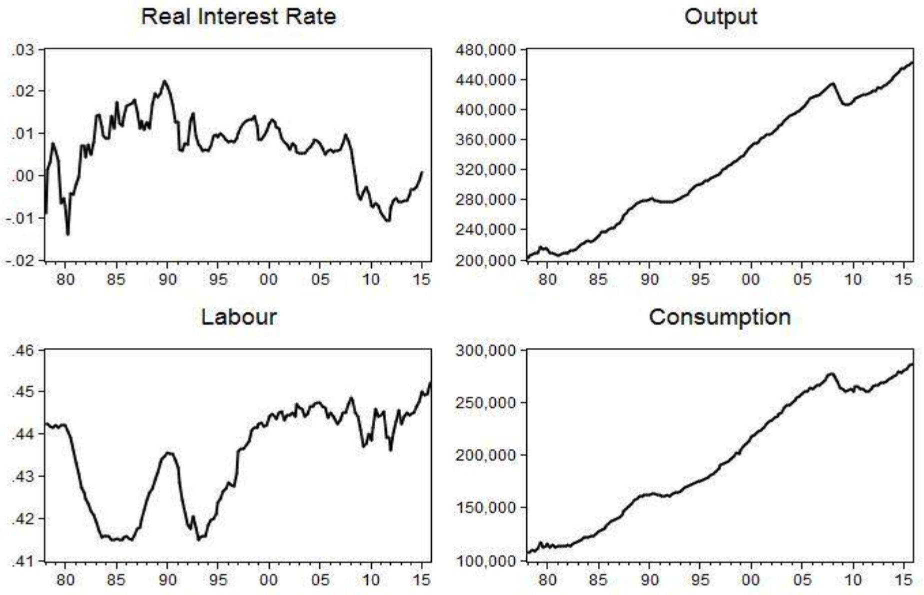
5 Model Data

We collect seasonally adjusted quarterly data in UK from 1978Q1 to 2015Q4 because many micro statistics in UK are only available after 1977. Raw data will be used without stationarisation primarily because we use VARX as the auxiliary model in which non-stationary variables could be directly considered. Besides, residuals backed out from model are observable even if non-stationary and thus we can separate deterministic trends from residuals to keep regression is used on stationary stochastic processes. Furthermore, stationarisation method could do harm to shocks. For example, HP filter will affect volatilities of shocks and introduce extra cyclical properties. Now consider aggregate and individual data sequentially.

5.1 Aggregate Data

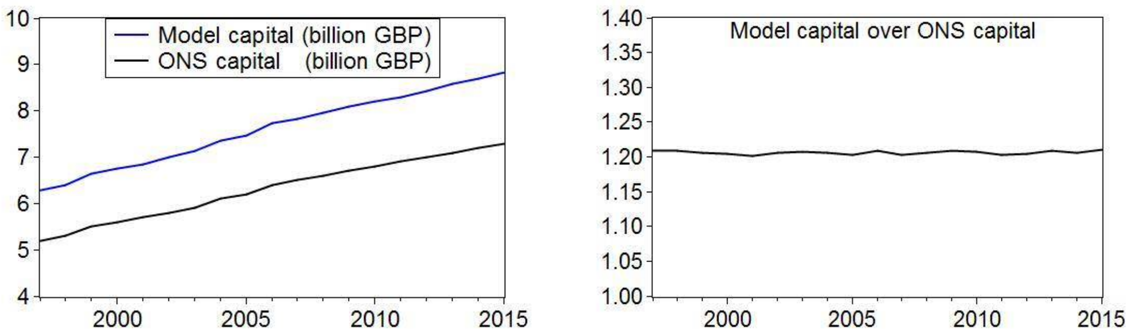
To exclude the inflationary effects, we collect the data measured by chain volume with base year 2012. Aggregate output and consumption are GDP and household final consumption expenditure respectively from UK national statistics. Real interest rate is one quarter of the difference between annual nominal interest and annual next-period inflation rate where nominal interest rate is measured by 90-day Yields Rate reported by Bank of England and inflation rate is CPI annual growth rate of all indexes. The aggregate labour is calculated the number of employment in UK aged 16 and over (MGRZ) divided by the sum of the total claimant count (BCJD) and the number of UK workforce jobs (DYDC) and then divided by two, sourced by Office for National Statistics (ONS).

Figure 2: Data on four aggregate variables (level unit: thousand GBP)



As the earliest data on capital stock available published by ONS starts from 1997, we estimate full observations by perpetual inventory method. Firstly we choose a constant annual capital-income (K-Y) ratio, 4.715 in practice which is the average of 4.16 in 1980 and 5.27 in 2010 reported by Piketty and Zucman (2014). The choice of a constant K-Y ratio aims to estimate an average capital-investment (K-I) ratio later for further iterations and does not mean that the actual volatilities of this ratio are ignored. Next, we collect data on aggregate investment by summing up "changes in inventories including alignment adjustment" (CAFU) and "total gross fixed capital formation" (NPQT), sourced by ONS. An average quarterly K-I ratio equals K-Y ratio divided by average investment and then multiplied by average income and 4 respectively. Suppose the equilibrium net growth rate of capital — is equal to that of income which can be easily calculated. Then a constant capital depreciation rate can be backed out from the equilibrium investment equation and the average K-I ratio. Lastly, we set the first period quarterly capital equal to the initial quarterly income multiplied by K-Y ratio and 4 respectively and use the investment equation   to iterate sequential capital stocks. The annually averages of capital stocks generated by us (Model capital) are compared with the annual capital published by ONS from 1997 to 2015. Their ratio is almost constant as shown below which implies that the almost constant gap after logarithm have no effect on model simulation.

Figure 3: Compare generated capital with ONS capital



measures the entrepreneurship cost in which a decrease could stimulate entrepreneurship incentives, reflected by time on entrepreneurship , could improve productivity regardless of shocks. Following Minford (2016)'s idea that costs of incentives could be represented by two factors, the Labour Market Regulation (LMR) and tax rate. LMR describes the degree of supervision in the labour market, protection to labour welfare and employment costs. For example, Fraser Institute reports the LMR indicator comprising 6 items among which we only use the third item "Centralized Collective Bargaining" (CCB) and the fifth item "Mandated cost of worker dismissal" (MCD) because others are most related to wages already implicitly included in agent income.[50] CCB describes the whole procedure for both employers

[50] The 6 items are "hiring regulations and minimum wage", "hiring and firing regulations", CCB, "hours regulations", MCD and "conscription". See more details in measurement of "Economic freedom"

and employees to make a collective agreement on working contracts (Gernigon et al 2000) and it is measured by a value from 0 (hardest to agreements) to 10 (easiest), so is MCD. Some preparations are needed before we use them. First of all, data on both CCB and MCD are not available in all time periods. Moreover they are annual data instead of quarterly data and lack of volatilities for model simulation. Therefore we begin with supplementing the omitted annual observations using 3-points quadratic estimating interpolation. Then to generate quarterly series with sufficient volatilities, we follow Minford (2016) to use a Denton method. We firstly choose an instrument with relatively high frequency, e.g. "trade union membership" (TUM) rate in practice. We calculate the TUM rate by dividing TUM which is the number of employees who are trade union members), collected from "UK National Statistics", by total number of employees. Next, use same quadratic interpolation method to achieve quarterly data of the TUM rate.[51] Since both CCB and MCD describe how lax the regulation is, we then define the final instrument by IV inverted ( ). Afterwards, we can yield quarterly data of CCB and MCD using Denton method.[52] Lastly, these quarterly data are inverted (in order to describe costs instead of benefits) and scaled to [0, 1] (scale them to keep consistent with tax rate) using ———. We take the average of these transformed CCB and MCD as transformed LMR (TLMR). For the other incentive determinant, different from Minford (2016) we choose corporation tax rate (CTR) rather than marginal income tax because entrepreneurship is more sensitive to the former. We also expand annual CTR to quarterly using quadratic interpolation and calculate quarterly by the average between TLMR and CTR. The relevant data are

---

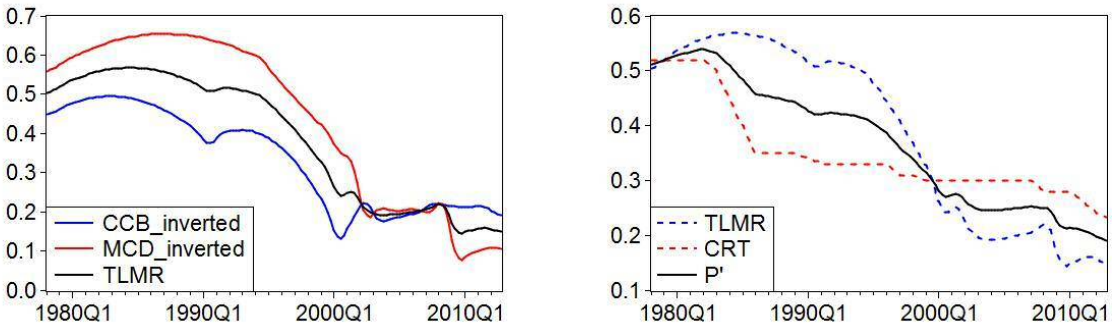
from Fraser Institute on Website:
https://www.fraserinstitute.org/economic-freedom/approach

[51] Source of TUM is https://www.gov.uk/government/statistics/trade-union-statistics-2016

[52] To obtain quarterly data of x denoted by y, run "y denton(IV,x,4,1)" in MatLab where x is annual data and IV is quarterly data

shown below where gradually decreases during the chosen time horizon.

Figure 4: Generated Data on



5.2 Individual Data

Theoretically, we should collect the data on one income group with fixed members. In practice, however, there are no statistics tracking same agents for decades. Hence, we make an assumption that although agents might move across income groups in different time periods, the average behaviour of one group is hardly to be affected. That is, we assume that the represent agent in one group is always "him" and this assumption is feasible if a large number of agents are included in one group so that we do not use too many groups but only two groups which are classified by the income deciles, top 10% and bottom 90% in UK.

Data on income distributions and wealth distributions could be found from "World Wealth & Income Database" (WID) website and UK ONS statistics like "average incomes, taxes and benefits by decile groups of all households" (ITB), "identified personal wealth" (IPW) and "distribution among the adult population of marketable wealth" (AMW).[53] For the 10%-90% segmentation on income and wealth, we use WID data. As data on consumption distribution is not available, we use "taxes

[53] See more details on these sources in Appendix 3

on final goods and services" density included in the ITB statistics as an appropriate proxy because all the tax rates (like VAT, duty on tobacco and Vehicle excise duty, etc.) within this category are directly relevant to consumption and are identical across agents unlike income tax rates which are different across agents according to base income values. ITB provides amounts of tax payment by income deciles, given a population sample and we calculate fractions of tax payment across groups using group payment over total payment. Since consumption tax rates are same across agents, individual consumption equals the total amount multiplied by his consumption tax payment fraction. See Figure 5 below where although both groups have significantly increasing trend as aggregate data, the rich (group one) have lower volatilities, particularly during 2008 financial crisis. This difference across agents is unobservable when representative-agent models and solely aggregate data are applied. To estimate individual labours, we firstly estimate individual labour ratio using the implicit wage which is equal to marginal productivity of labour. Then individual labour ratio can be approximated as individual income ratio over individual wage ratio where the latter could be calculated from the item "wages and salaries" within ITB. We find individual labour ratio is quite close to the group population ratio so that one could also simply assume both inner-group representative agents have same labour. All the individual data above are transformed with group population weights, denoted by (e.g.), to data of inner-group representative agents. To my knowledge, data on individual entrepreneurship time is unavailable to collect and that is why we substitute out from model equations. Individual can be estimated from aggregate and individual wage distribution. Recall the definition — and — where is the real wage of inner-group representative agent *i* and is aggregate wage. We estimate as follows.

where is the wage share of group *i* which can be calculated from ITB statistics. Lastly, individual productivities are individual Solow residuals.

Now we have gathered all the data we need and some annual data are transformed to quarterly by quadratic interpolation again. For convenient application in model simulation, we scale values of most variable except interest rate, labor, and productivities by firstly dividing each value by a same constant (e.g. sample average of aggregate output) and then taking logarithm. For labours, we dividing each value by sample average of aggregate labour and then taking logarithm. Last of all, productivities are multiplied by the sample average of aggregate labour with power , divided by average of aggregate output also with power and then taken logarithm.

Figure 5: Data on individual output, capital and consumption

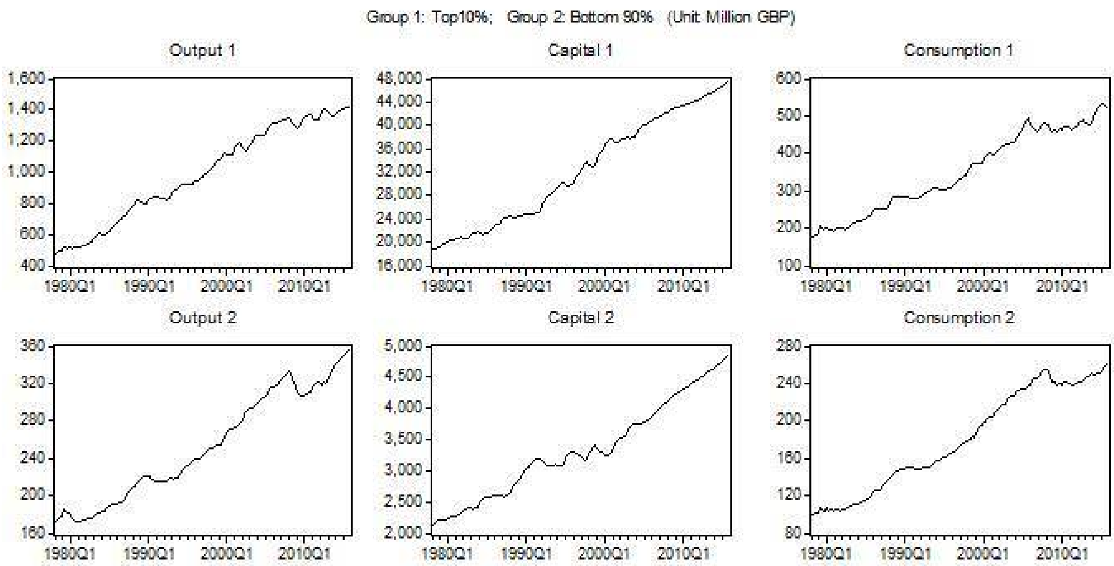
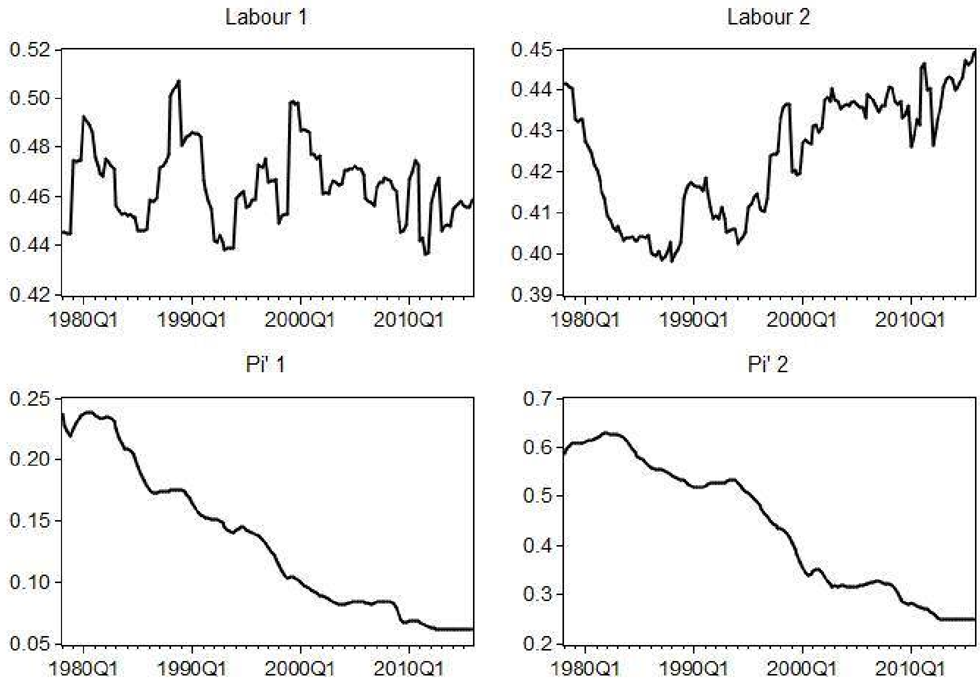
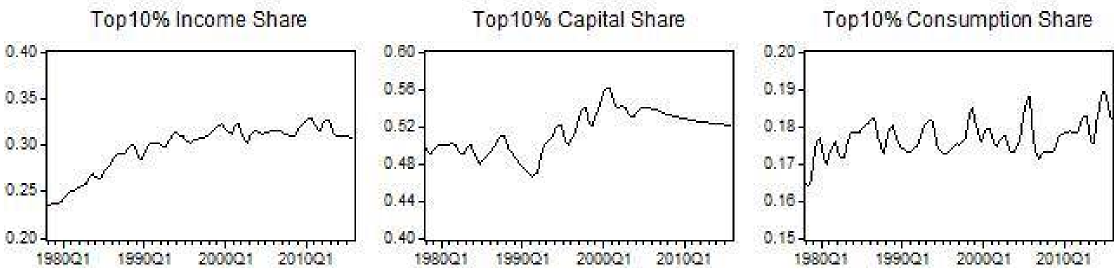


Figure 6: Estimates of individual labour and entrepreneurship penalty rate



Gini-coefficient is more often used for inequality measurement in reality which is calculated according to the Lorenz curve (cumulative shares distribution) and used to describe the ratio of unequally allocated resources (e.g. income) over all so that 0 represents completely equality and 1 represents completely inequality. However, Gini-coefficient has a shortcoming that it cannot tell where the inequality comes from. Piketty (2014) compares different inequality measurements and suggests using shares table, e.g. shares of each quintile, decile or percentile. We follow Piketty (2014)'s way to measure inequality by shares, although we have only two groups. Our model could work on more than two groups actually. As shown in the model-setting section, our inequality indicator is the top 10% decile average capital over aggregate capital. To keep consistent form with the prevailing expression, we can transform it to Top 10% shares over total. One clarification should be cleared that we simply approximate wealth distribution to capital distribution, different from Aiyagari and some others who distinguish these two, primarily because capital and income show close tendency and income generally has small values relative to capital so that whether income is cut off from capital does not have significant effects. 10%-90% segmentation individual data and some inequality indicators are shown in Figure 7 where income inequality has a remarkable rise, especially in 1980s while capital inequality experiences a sharp decline at the beginning of 1990s and turns to rise afterwards. Consumption inequality does not perform a clear tendency in last decades.

Figure 7: Inequality Indicators



Before we close the data illustration, one more point should be noted that we need to increase sample size because the simulation algorithm requires original N periods plus the number of forecasting periods F, e.g. 50 in practice. Therefore, we extend non-stationary data by their trend while extend stationary data using the last period observation.

6 Empirical Results

6.1 Detection on tendency

This section aims to examine whether the model simulated relation between inequality and aggregate growth has a stable tendency. To check whether inequality and growth could be generated simultaneously, we assume that two groups are innitally identical with equal population weights, half-half and identical resource allocations. Inequality will be generated solely by innovations. As we expect the tendency is not occasional (parameter-independent) if it exists, parameter values are calibrated and freely set. Identical groups imply that data on each one is same as the aggregate representative-agent and thus we calibrate most parameters based on UK aggregate data from 1978Q1. Details of calibration are same as those shown in the empirical section where estimation basedon actual heterogeneous-agent data is also adopted and are omitted here.

Table 1: Freely set values of key parameters

Marginal effect of entrepreneurship time on individual productivity growth		0.500
(Negative) Marginal effect of capital term Q on individual oppurtunity cost of entrepreneurship		0.005
(Negative) Marginal effect of entrepreneurship penalty rate on individual productivity growth for group		0.850

To ensure the time period is sufficiently long to observe tendency, data are extended to 250 periods from 1978Q1 by same way illustrated in the data section.

6.1.1 Benchmark tendency

First of all, to generate artificial inequalities, we randomly draw individual innovations from chosen normal distributions where any one group can be selected to be the rich by shocks rather than bootstrap innovations backed out from model equations because individual innovations now are same across identical agents. For convenience, we set the distribution mean of innovations slightly different across groups to make one group, say group one in practice, always be chosen as the rich;

otherwise inequality indicator is difficult to compute by programming.[54]

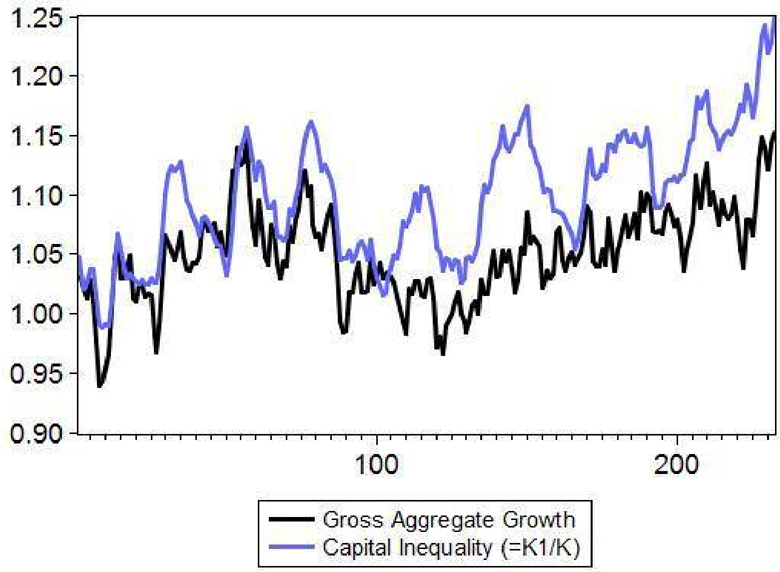
One typical simulated sample is shown in Figure 8 where capital inequality is calculated by capital per capita in group one (now selected as the rich) over the aggregate capital per capita which implies that this indicator equal to unity represents a perfect capital equality while a value either greater than or less than unity stands for unequal capital distribution.

In consideration of the non-negligible effects of actual data on simulated growth (for example, suppose the actual data experiences a recession at a certain period. The economy is hard to be pushed to boom at that period even if simulated data affected by model behaviours and shocks could vary differently from actual data), when think about aggregate growth, one can either directly compute quarterly growth rate of simulated output or observe the deviation of simulated data from the actual data. The latter measurement which is convenient for intuitive comparison with inequality due to relatively lower volatilities is used in Figure 8. First of all, capital inequality and aggregate growth are simultaneously generated in despite of some fluctuations due to random shocks (both are almost always greater than one) which fit most empirical studies that there might be a relation between inequality and growth. Second, aggregate growth has almost same tendency with inequality and inequality is likely to lead the volatilities of growth. These two findings support our point of view to some extent that capital inequality might stimulate growth, especially in long term. The short-run fluctuation also fits our theoretical prediction that labour input might be partly transferred to entrepreneurship time which has a temporary negative shock to

[54] One could solely randomly draw individual shocks as they are sufficient for inequality generation. Specifically, individual labour supply shocks of group one are drawn using a slightly higher mean than those of group two in practice in order to always choose group one as the rich because our inequality indicator is capital per capita of the rich over aggregate per capita. Of course, there is nothing different if group two is selected to be the rich. This setting is only for programming convenience.

growth. Specifically, the increasing wealth concentration is unlikely to die out without policy intervention which implies that redistribution policy is necessary if policy makers concern social equality and stability.

Figure 8: Tendency of the relation between aggregate growth and capital inequality



To exclude the occasionality, we also conduct a Monte Carlo experiment by simulating 1,000 samples to examine the average relation where the aggregate growth rate is accurately quarterly growth rate of simulated aggregate output instead of deviation from actual data. For convenience, we calculate the average correlation coefficient between two indicators across 1,000 samples in practice and yield the coefficient value as 0.2726, implying a positive correlation. This qualitative finding is consistent with the reality where quarterly aggregate growth rate in UK is positive correlated with the inequality indicator measured by capital share of top 10% from 1978Q1 to 2015Q4 and the correlation coefficient is 0.2291. Therefore, we can believe the stimulating effect of capital inequality on growth explained by our model.

6.1.2 Redistribution tendency with tax

Equality concerning both social welfare and stability is always an important consideration of policy makers. Although wealth inequality in UK experiences a remarkable decline over two centuries, it has a significant rising tendency in recent decades and the tendency was not interrupted during the 2008 financial crisis. For example, wealth share held by top 10% in UK was 51.98% in 2006 before crisis and rose to 54.01% in 2009. It is noteworthy that this indicator was only 45.99% in 1990. Coincidentally income inequality is also exhibiting an increasing tendency. Hence, there is necessity to consider the redistribution problem.

This section aims to examine the redistribution tendency by introducing an income taxation regime. Constant proportional tax rate will be considered for simplification but with various utilisations of tax revenue. For instance, tax revenue could be used as government spending funding, identical per capita subsidy across groups or even income transfer from the rich to the poor. Note that identical per capita subsidy is unlikely to be offset by tax payment in HAMs, quite different from that in representative-agent models which is the lump-sum tax. The three ways of taxation have different issues. Taxation with no subsidy (funding government spending) is simplest. However it might lead to a self-fulfilling result that inequality will be reduced and meanwhile economic growth will undoubtedly fall due to the decrease in aggregate capital used. The other two need approximation; otherwise subsidy or tax transfer will have no effect on individual capital accumulation and consumption because of the lack of individual budget constraints themselves as model equations. To understand this issue, consider the poor who benefit from subsidy or income transfer. If individual budget constraints are used as a model equation, subsidy or transfer has a positive income effect on the poor like that the poor have more endowment to consume or accumulate capital. Nevertheless, individual budget constraints themselves are kicked out from model equations to omit individual bonds so that subsidy or transfer has no effect when the poor make the optimal decision because the first order derivative w.r.t. the predetermined subsidy or transfer is zero and is used nowhere else. We can use the following approximation to avoid this trap. Suppose an income tax transfer is adopted where a constant income tax rate is enforced on the rich (group one for example). Then the tax revenue per capita across whole population is which is transferred from the rich to the poor with population weight and thus aggregate output is unchanged, i.e. where is actually a subsidy rate for the poor. It can be approximated as . Then individual income after taxation transfer of the rich is and the income transfer to the poor is . The imperfection of this way is using approximation , but similar approximation is also used when linearise aggregation equations.

In this section, the simplest tax regime is used where both groups are enforced a same constant income tax rate, say , with no subsidy because to groups are initially identical and different due to shocks which make differences insignificant so that an income transfer may invert the rich to the poor. Our main target here is to examine a stable tendency given a policy intervention, like tax no matter what type used and some various tax regimes will be compared in the "Redistribution and policy measurement" section later.

The individual budget constraint now is

The nonlinear optimal rule for capital accumulation (6) now is

The nonlinear optimal rule for labour supply (7) now is

The nonlinear optimal rule for individual productivity growth (12) now is

All the other equations are same as those in the benchmark model. The linearised equations of (37), (38) and (39) take same forms as (26)-(27), (30)-(31) and (32)-(33) respectively except new constant terms containing tax rate which can be captured by “Type II Fix” of our algorithm. Coefficient now becomes which differs from that in the benchmark model and thus all the parameter values are set same as those in previous benchmark tendency section except 0.648. Following same simulation procedures, we draw same random shocks across models by setting same “time seed” in order to efficiently compare simulated behaviours with and without tax.

Figure 9 shows one simulated sample using same random shocks as the sample in previous section where growth and inequality have almost same tendency, but lower levels. The differences in growth rates and also capital inequality (the value simulated by benchmark model minus the value simulated by the model with) are shown in Figure 10. Obviously, income tax reduces inequality and aggregate growth simultaneously, but the change in inequality is much more remarkable. Moreover, the gap due to tax is gradually enlarged. Similar findings are found after repeating this experiment many times and the significant tendency makes it feasible to carry on further analysis on actual distribution.

Figure 9: Tendency of the relation between growth and capital inequality with tax

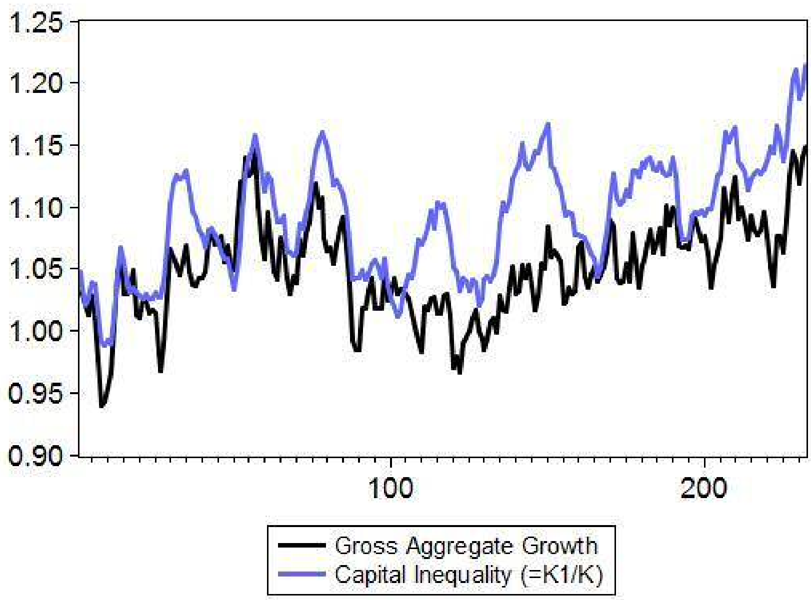
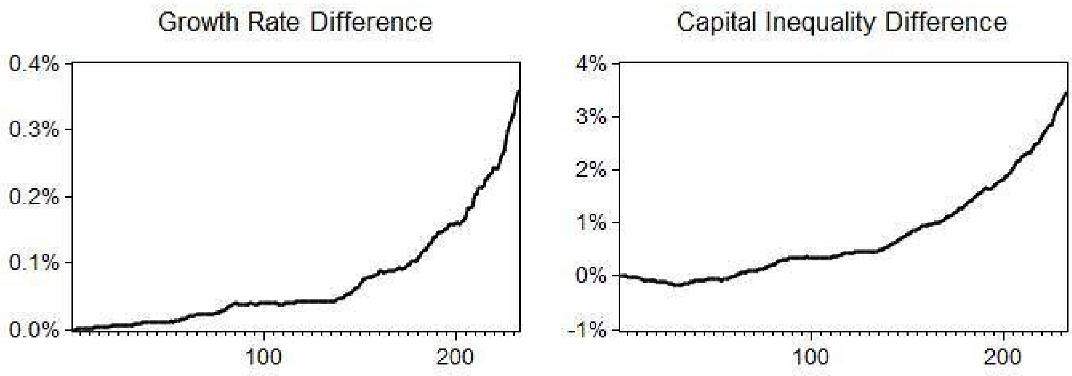


Figure 10: Experimental comparison between the tax model and benchmark model



6.2 Empirical study on actual data

The primary purpose of the empirical research is to examine whether present model could minic main characteristics and tendency of economic growth and distributions of capital and income in recent decades in UK. Then the final aim is to search for appropriate policy intervention instriuments and provide practical suggestions by analysing dynamic effects of changes in exogenous factors on growth and inequality. This chapter mainly consists of following sections, illustration on calibration and estimation, Impulse response and variance decomposation analyses, and redistribution.

6.2.1 Calibration

The capital share in production   is set to 0.3 according to the average of annual labour share in UK from 1978 to 2015 reported by FRED economic data. Utility discount factor   is set to 0.97 to match UK average quarterly risk-free rate 0.031. As Oulton and Srinivasan (2003) report the aggregate annual depreciation rate of capital stock in UK is 6% measured by ONS method in 2000 and has no trend to increase, we set the quarterly .[55] Steady-state aggregate output/consumption ratio and capital/consumption ratio are calculated from the average ratio from 1978 to 2015. The steady-state quarterly growth rate of output, capital and consumption are all 0.0057. The shares of group one's holdings over total amount in three aggregation equations are calibrated to , and respectively using average shares during the empirical periods while shares of group two are equal to one minus shares of group one. The steady-state individual labours are approximated to 0.5 as used to linearise equations of individual labours. The weight of consumption preference in utility is set to 0.5. We do not estimate this parameter because any change in the constant terms will be corrected by the "Type II Fix" procedure in our algorithm and only occurs as a constant term in the linearised labour equation so that the change on this parameter does not affect the simulated behaviours significantly (but could be estimated). In fact, there is an under-identified problem as I mentioned in Id-If section when parameters only come up in the constant terms after linearisation. Minford (2016) estimates states the estimated value on it is 0.5276, close to our setting here. Similarly some other parameters only involved in the constant terms after linearisation, like drifts in the linearised individual productivity equations and the drift in productivity as a function of entrepreneurship time , are all set to the steady-state productivity growth rate 0.004. The elasticity of consumption in utility is set to unity following the value used in the proof in Appendix 1. Labour elasticity is also set to unity for simplification and one could estimate this parameter. The constant term in individual entrepreneurship penalty

[55] ONS indicates that the life of capital like plant and machinery is from 25 to 30 years, longer than that in US and thus the relevant depreciation rate is lower than that in US.

equations is set to the sample average of the aggregate penalty rate.[56]

Table 2: Calibrated parameter values

Share of capital in Cobb-Douglas production		0.3
Utility discount rate		0.97
Capital depreciation rate		0.015
Share of conumption preference in CRRA utility		0.5
Elasticity of consumption in CRRA utility		1
Elasticity of labourr in CRRA utility		1
Steady-state net growth rate of productivity		0.004
drift in linearised productivity equation for group 1		0.004
drift in linearised productivity equation for group 2		0.004
drift in individual entrepreneurship penalty equations		0.369
Steady-state output share by top 10% income decile		0.3
Steady-state capital share by top 10% income decile		0.5
Steady-state consumption share by top 10% income decile		0.3
Steady-state aggregate output/consumption ratio	Y/C	1.693
Steady-state aggregate capital/consumption ratio	K/C	18.43

6.2.2 Estimation

We emphasise estimation on parameters and where the first one tells us how capital distribution affects policy attitude on entrepreneurship penalty rate and the second tells us how this intermediate factor drives productivity growth and even the aggregate growth. To estimate , we firstly estimate . Since the ratio could be calibrated using individual observations, the determinant of needs to estimate is . Although is the steady-state of individual penalty rate, the way to collect data on it as shown in the data section leads to possible errors so that we still need estimation. That is, we indirectly estimate by the parameter restriction. Id-If estimators are shown in Table 3.

[56] The simulated annealing method to search global optimum is time-cost with too many parameters to estimate and thus we focus on key parameter estimation.

Table 3: Id-If estimators

Marginal effect of entrepreneurship time on individual productivity growth		0.7016
(Negative) Marginal effect of capital term Q on individual oppurtunity cost of entrepreneurship		0.001
(Negative) Marginal effect of entrepreneurship penalty rate on individual productivity growth for group 1		0.7822
(Negative) Marginal effect of entrepreneurship penalty rate on individual productivity growth for group 2		0.5534

The small estimated value of , 0.001, does not mean that entrepreneurship penalty rate is insensitive to capital distribution and it is because penalty rate has a lower magnitude order than capital. The estimated greater than implies that individual productivity is more sensitive to changes in penalty rate for the rich than that for the poor.

As illustrated in the Id-If section, our auxiliary model VARX takes the form of

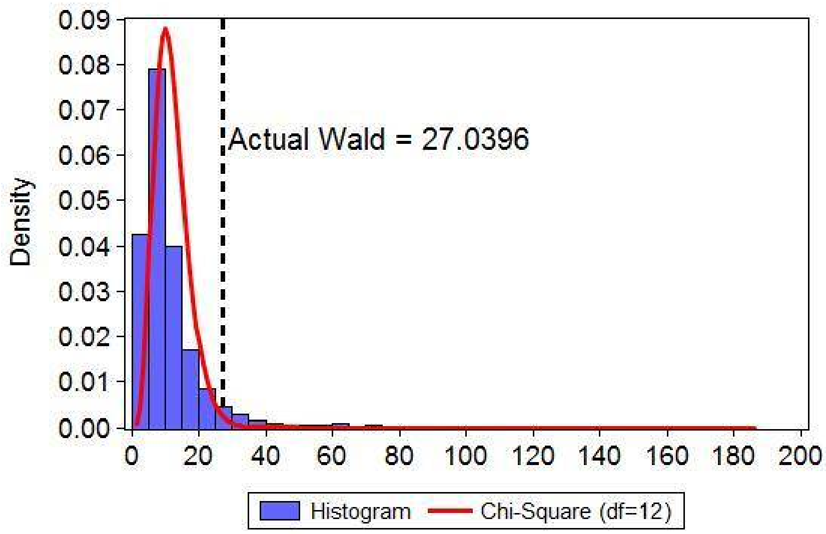
where is a 3 by 3 coefficient matrix and is a 3 by 2 coefficient matrix. The exogenous non-stationary variable vector and the error vector is . Our auxiliary parameter vector used to compute the Wald statistic is which contains the 9 elements in and also the 3 variances of the VARX regression residuals . Given the actual data, the estimated and are shown below. The VAR coefficient 0.1605 reflects a positive effect of capital inequality on the aggregate growth.

Details of our Id-If Wald test on the null hypothesis that structural parameters are equal to our estimators above are shown by Table 4 and Figure 11. Obviously, the null hypothesis cannot be rejected under a 5% confidence interval (The actual Wald is greater than 94.35% of all the 1,008 simulated Wald statistics).[57]

Table 4: Id-If Wald test result

Wald Statistic	Transformed MD	Wald Percentile
27.0396	1.5219	94.3452 %

Figure 11: Distribution of simulated Wald statistics



6.2.3 Structural Errors and Innovations

Given estimators, AR coefficients of the structural errors are summarised below.

Table 5: AR coefficients of structural errors

0.9651	0.9453	0.8312	0.0026	0.8892	0.8773	0.9390	0.9388	0.9705

57 Simulations are run 1,200 times among which some simulations do not meet format setting and thus no output.

It is introduced in the Id-If section that aggregate productivity is assumed to have an I(1) process. Given estimated parameters, aggregate productivity now is tested by ADF test and Phillips-Perron test and the process is found to apparently follow an I(1) shown below.

Table 6: Unit root test on aggregate productivity

Unit Root Test	Level	Trend	1st Difference
ADF test Statistic	-0.0527 (95.10%)	-1.8443 (67.67%)	-11.3836 (0.00%)
Phillips-Perron Test Statistic	-0.0388 (95.24%)	-2.1592 (50.74%)	-11.3697 (0.00%)

(Note: values in the parentheses are p-values)

The structural errors and structural innovations used for bootstrapping are shown in Figure 12 and 13 where some errors are separated from a deterministic trend following the way introduced in the Id-If section.

Figure 12: Structural Errors

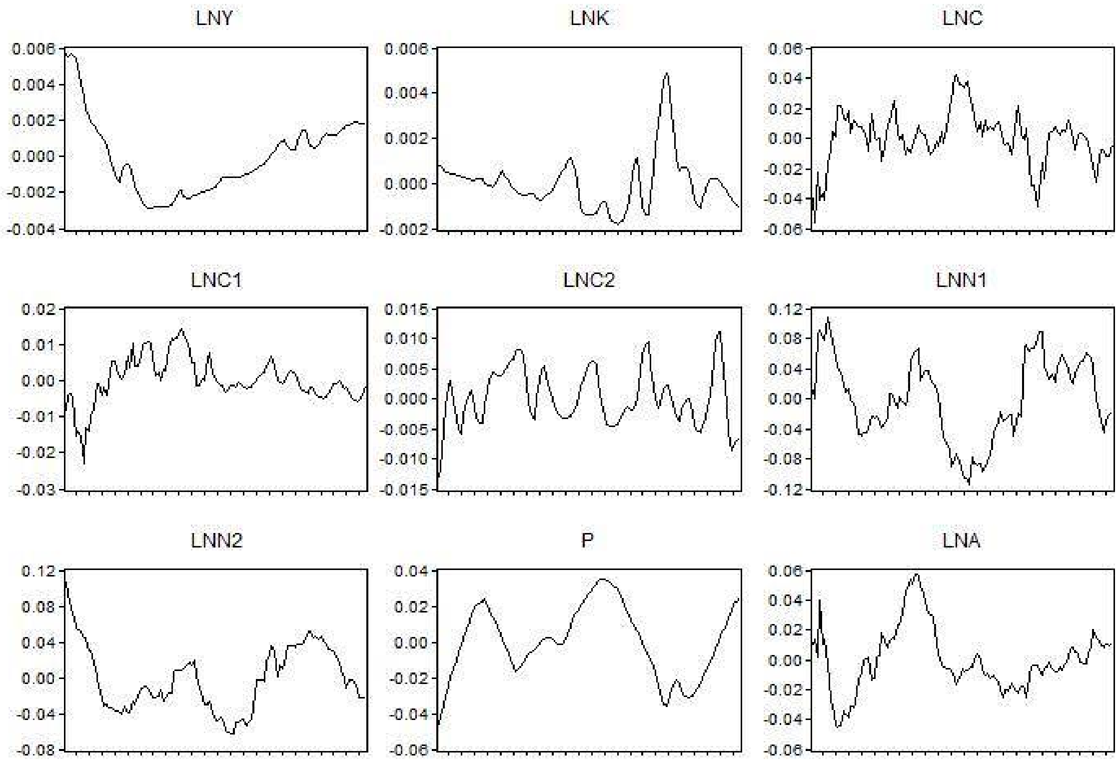
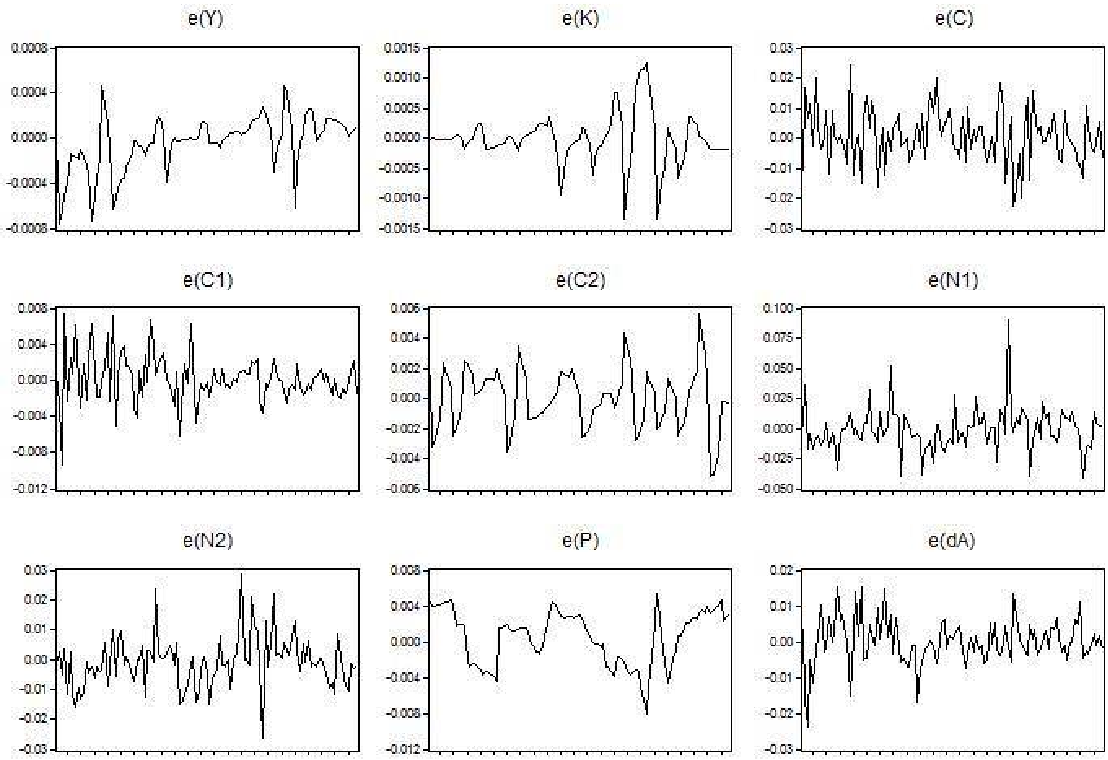


Figure 13: Structural Innovations



6.2.4 Id-If Power Test

Le et al (2012) and Le et al (2015) conduct Monte Carlo power test on three testing method Id-If, Likelihood ratio test and "unrestricted Wald" test on different models. Id-If is found more powerful than some other classical testing methods. To evaluate the power of Id-If on our model, we can provide a powerful Monte Carlo statistical test both against parameter mis-estimation and more importantly against model mis-specification, including models with different causal sequencing and capable apparently of providing 'observationally equivalent' data. We firstly generate 500 samples from the true model and the actual data. Then given each simulated sample from the true model as the observation, we test the false model by Id-If and calculate the rejection rate out of the 500 Monte Carlo experiments. Table 7 shows the result of our power test against the false models with mis-estimation where both structural parameters and the AR coefficients of the errors are steadily falsified by a percentage degree +/- x each time. The probability of rejecting the false models is getting higher associated with an increase in the falsity of parameters and the power is considerably high given a significant falseness.

Table 7: Power test against numerical falsity of parameters

Parameter Falseness	True	1%	5%	7%	10%
Rejection Rate with 95% Confidence	5%	13.4%	67.2%	82.6%	100%

We then test the power of Id-If against a mis-specified model in which the basic mechanism of wealth inequality on entrepreneurship is turned off. Namely, the equations of penalty rate (34)-(35) are replaced by a simple AR(1) process. Wealth inequality in this false model is generated by randomness. We keep parameters same as the full-estimated values in the benchmark model. As the rich are still more sensitive to a reduction in the penalty rate ( ), this mis-specified model can also generate both growth and rising wealth inequality but they will not be correlated. We still consider 500 samples and the rejection rate of this mis-specified model with 95% confidence is as high as 99.4%. In fact, out Monte Carlo test has a limitation that we have to test the true model against any mis-specified model which could generate similar observed data. However, in practice, we selected a certain mis-specified model which is much more likely to generate both growth and rising capital inequality simultaneously than many other mis-specified models. And hence we can conclude that Id-If provides huge power against the mis-specified models which attempt to mimic the results from the true model.

6.2.5 Impulse Response

In this section, we analyse sequential effects of temporary shocks on the aggregate economy, individual variables and also inequality indicators within forecasting period. The analysis is focused on some typical results and more impulse response (IR) figures could be found in Appendix 4. Some stationary aggregate shocks have same effects as in basic RBC models. For example, a positive aggregate production shock stimulates aggregate output, aggregate capital and consumption, which lead to a decrease in real interest rate. A lower interest rate induces more individual consumptions and more capital holdings (fewer bonds). Individual labours begin with a decrease because marginal utility of leisure is greater than marginal production of labour. Individual outputs do not catch up the sudden increase in aggregate output immediately, but gradually rise after a sudden fall due to the increase in individual capital input and individual labour recovery. Overall, capital distribution does not change corresponding to stationary aggregate shocks and neither do individual productivity growth rates. Therefore the sequential effects finally die out. It should be clarified that some IRs do not behave hump shapes mainly because of the individual capital equation where capital is too sensitive to the real interest rate determined by the coefficient of on the RHS. If one insists on performing the hump-shaped (Impulse responses) IRs, the "Capital Adjustment Cost (AJ)" originating from Hayashi (1982) can be used in individual's budget constraint taking the following form.

where is a convex in terms of capital and could take the form – —— used by Hayashi (1982) or the form – used by Meenagh et al (2005), Davidson et al (2010) and Minford (2016). The linearised capital equation generally takes the form below. For example, the estimated marginal effect of interest rate on capital by Minford (2016) is 0.2365.

In this model, we ignore the analysis on sequential effects of an aggregate productivity shock because the logarithm of individual productivity has a unit root process and thus one-period aggregate productivity shock will dominate all the other one-period shocks. Our focus is on individual shocks and the entrepreneurship penalty shock since they are vital to the dynamics of inequality. The following two figures show the IRs to one-period individual shock on consumption preference and labour supply preference respectively for group 1 (the rich). A sudden rise in consumption preference of the rich firstly reduces their individual capital holdings. Meanwhile, their individual labour also has a sudden fall because more entrepreneurship time is required to enhance future productivity in order to support their higher consumption demand in the future; although this is harmful to current individual output. Hence their individual output has an initial decline but gradually recovers afterwards. The rise in individual consumption of the rich also raises the aggregate consumption up, leading to a lower real interest rate which in turn stimulates individual consumption of the poor but with a smaller increase than the rich. Corresponding to the decrease in interest rate, the poor input more capital in input (transfer from bonds) and reduce labour input slightly to increase entrepreneurship time while the rich reduce capital input to satisfy the consumption requirement. Although the poor's output also has a decline due to the decrease in labour input at the beginning, it recovers soon. As expected, finally both income inequality and capital inequality (share of holdings by the rich) decline. Similarly as shown in Figure 15, a sudden rise in the labour supply preference of the rich finally results in more unequal distributions of both capital and income.

Figure 14: IRs to a +5% one-period individual consumption shock on the rich

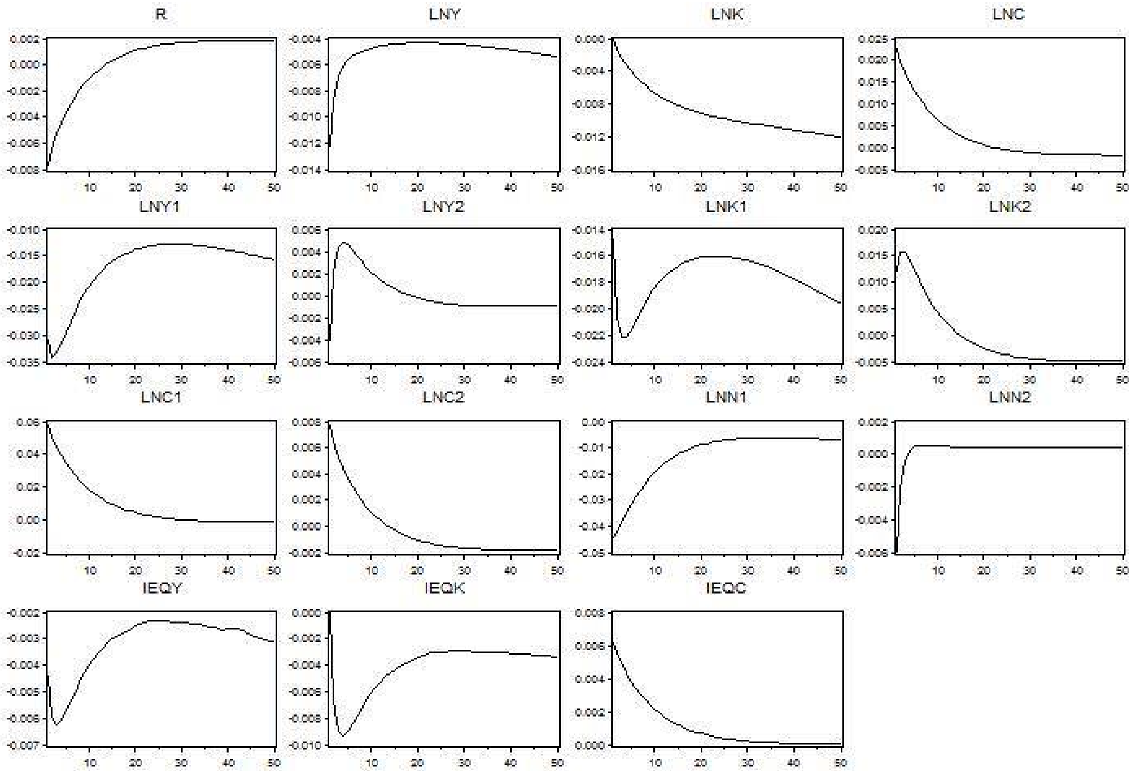
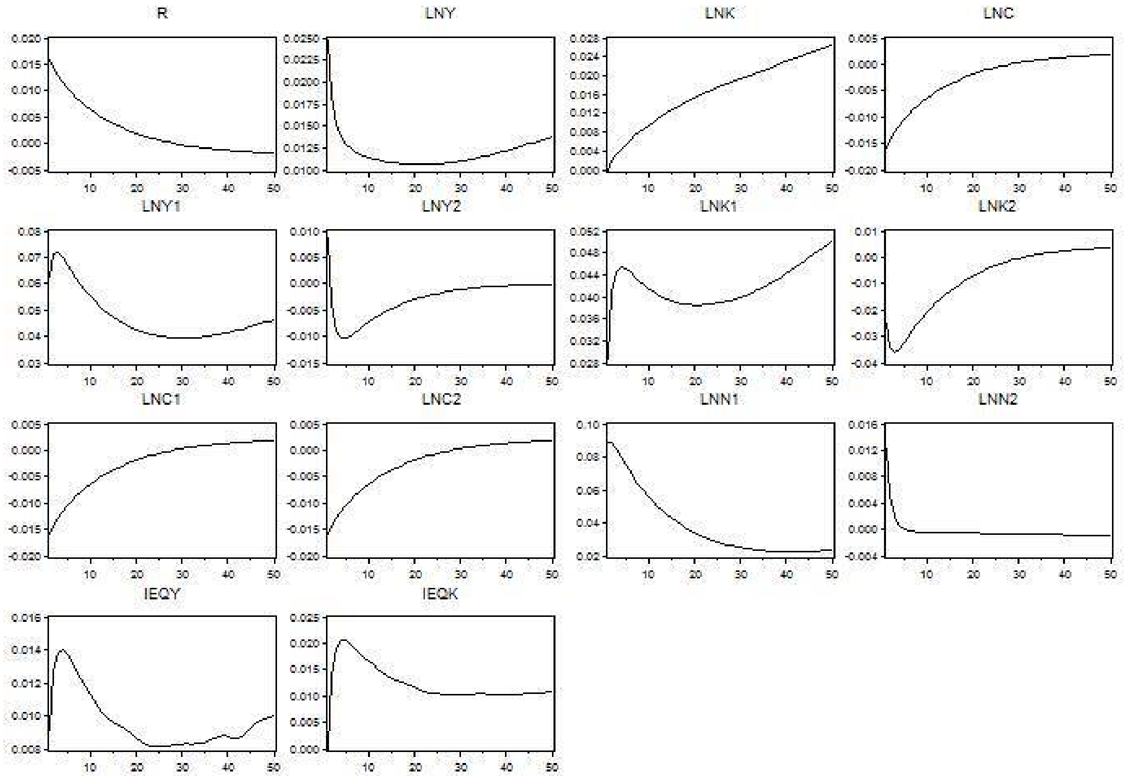
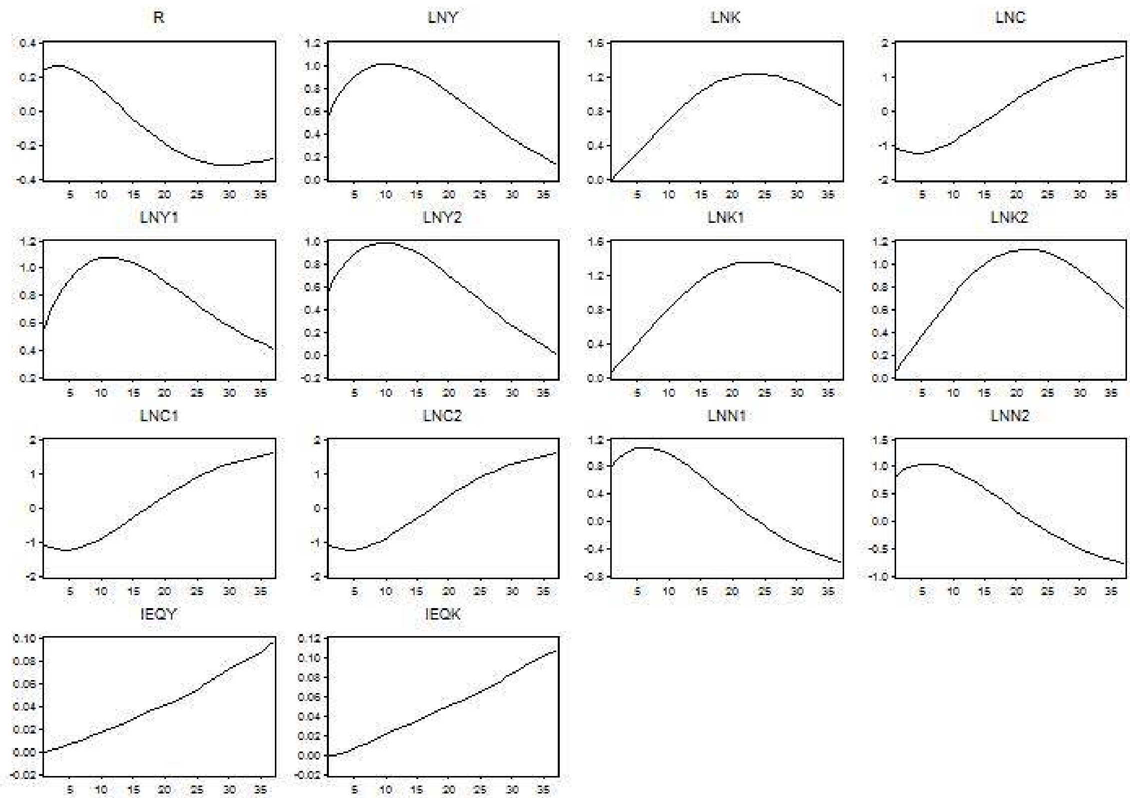


Figure 15: IRs to a +5% one-period individual labour supply shock on the rich



The effects of entrepreneurship penalty rate shock are of our most interest because the previous individual shocks are generally non-controllable for policy makers. Conversely, , like a tax rate in some sense, could be intervened by government. As Figure 16 shows, when due to an identical negative shock to both groups, first of all, both individuals have incentives to spend more time on entrepreneurship and this extra time is more likely to be squeezed from leisure instead of labour because both marginal productivity of labour and capital will rise after an increase in individual productivity due to the rising entrepreneurship time and more labour and capital will be input into production. This is shown by the figure that both individual labours and capital holdings go up and so do individual outputs. During this process, consumption starts with a decline (consume less and input more) and interest rate starts with an increase. One important finding is that although both groups have an increase in individual income as well as capital, distributions of income and capital gradually become more unequal. The reason is that in despite of the same lower , the individual productivity of the rich has a higher growth rate to make the gap wider probably because firstly the rich have higher capital holdings per capita initially and secondly individual productivity is more sensitive to changes in penalty rate for the rich ( ). This tells a policy implication that social inequality could be aggravated if governors attempt to reduce entrepreneurship penalty rate like the corporation tax rate to stimulate economic growth.

Figure 16: IRs to a -3% one-period entrepreneurship penalty shock on both groups



6.2.6 Variance Decompositions

In order to analyse how various shocks could influent economic volatilities and further examine the feasibility of policy interventions (for example, if the change in economic behaviour by a certain intervention is neglectable, this intervention instrument is infeasible), we implement variance decomposition analysis by two different ways.

The first is a usual way (like IRs) to consider the dynamics of volatilities caused by a certain temporary shock across a period of time. Given 1,000 initial one-period innovations to a certain structural error (randomly draw the innovations from a normal distribution with mean zero and variance 0.05 in practice[58]), we yield 1,000 simulated samples and calculate the variance of each endogenous variable across the samples in each period. We repeat this experiment on each structural error and calculate the proportion of variance on a certain endogenous variable in one period caused by each shock out of the total variance (sum of all variances for a certain variable in one period). Given simulations, we can yield 16 (number of endogenous variables) groups of variance decompositions among which some typical analyses are shown in Appendix 5. For example, an aggregate capital shock initially plays an important role in the volatilities of real interest rate, aggregate capital and aggregate consumption, but the effect diminishes quite soon after several periods. Conversely, both the entrepreneurship penalty rate shock and the aggregate productivity shock have an increasing influence on volatilities of all the aggregate variables. For individual variables, their volatilities are all dominated by the entrepreneurship penalty rate shock as time goes by. Our focus now is on the time-variant volatility effects on aggregate output growth and inequalities since this analysis could provide a reasonable reference on the redistribution policy which we will discuss about in next section.

[58] The distribution variance is selected to be small in case of the divergent iteration due to large magnitude shock.

Figure 17 shows the variance decomposition for aggregate growth within 25 forecasting periods where 4 shocks (respectively on entrepreneurship penalty rate, aggregate productivity, aggregate capital and aggregate output) capture almost all fluctuations of aggregate growth and thus proportions of the other 5 shocks are omitted. Figure 18 below shows 5 most important shocks respectively on entrepreneurship penalty rate (almost the blank area), aggregate capital, individual labours and individual consumptions to determine volatilities of capital inequality. These findings suggest that to affect growth and inequality together, policy interventions should be levied on aggregate capital or entrepreneurship penalty rate because shocks on them are both determinants on volatilities of growth and inequality. Moreover, as stated before, adjustment on entrepreneurship penalty rate results in a trade-off between growth and inequality and this implies that governors could consider a way to control aggregate wealth for appropriate interventions.

Figure 17: Proportions of key determinant shocks on volatilities of aggregate growth

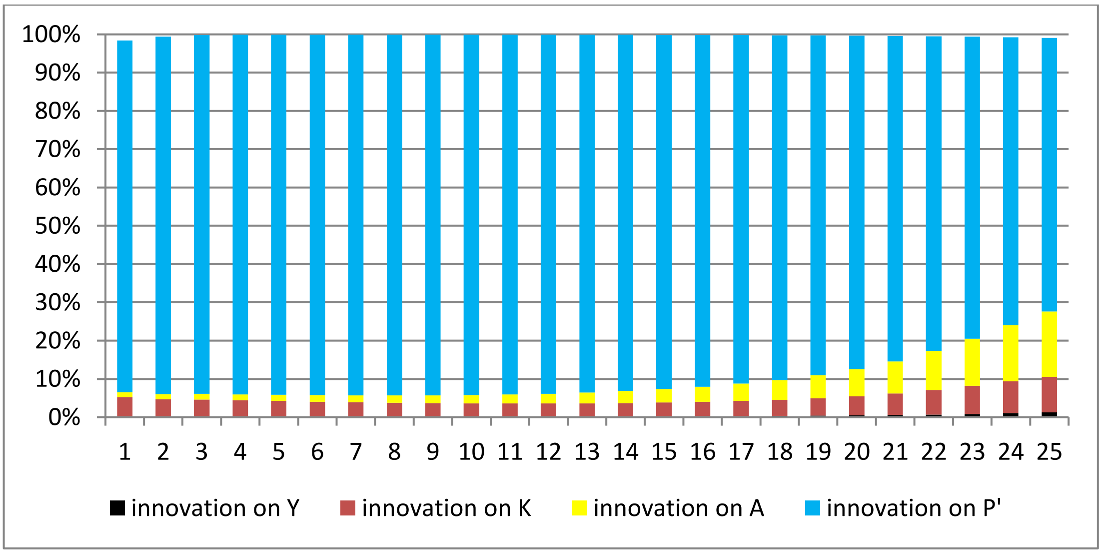
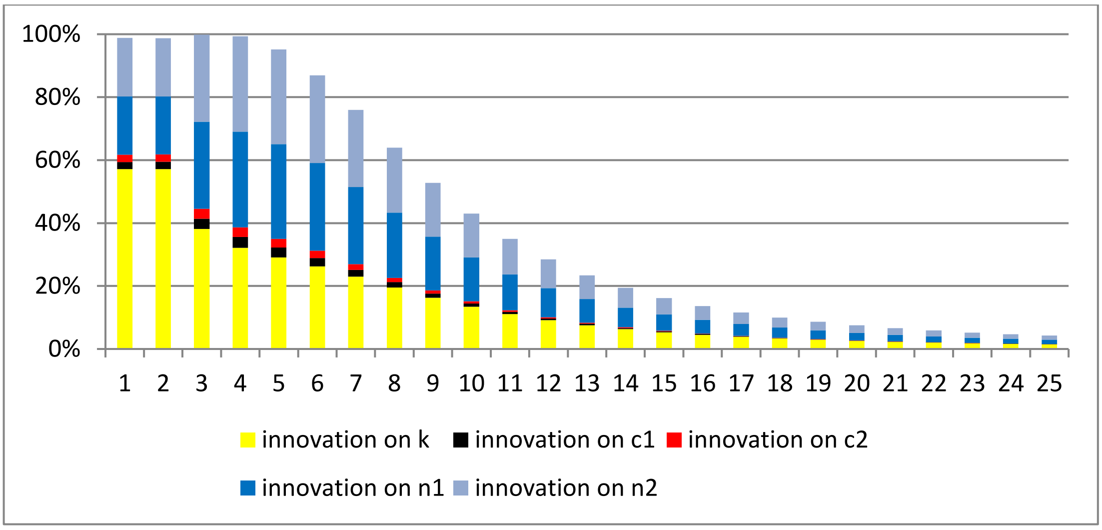


Figure 18: Proportions of key determinant shocks on volatilities of capital inequality



We also analyse volatilities of each endogenous variable corresponding to different shocks across time horizon following Minford (2016)'s way to reduce the errors caused by non-stationary shocks or shocks with high auto-correlation process. We firstly bootstrap innovations on one structural error individually to obtain 1,000 simulated samples. Then calculate the variance of each endogenous variable along the time horizon and average variances over the 1,000 samples for each variable. Repeat this experiment on each structural error and calculate the proportions corresponding to different errors for each variable.[59] Different from the usual way of decomposition which tells us whether variables are sensitive to a certain shock and whether the sensitivity could remain along periods, this way of decomposition tells us whether the change in a variable caused by a certain shock is stable across time. For example, proportions summarised in Table 8 below indicate that volatility of aggregate growth across time caused by aggregate capital shock only accounts for 0.6%, implying that relatively frequent adjustments on aggregate capital as well as on aggregate output, consumption, individual labours and individual consumptions (intervention) are

59 This decomposition gives us static proportions instead of time series of proportions.

possible because the effects on fluctuations of growth path are moderate. Contrarily, no shock has stable effects on the path of capital inequality and thus we have no preferences between different interventions suggested by the previous usual decomposition analysis. Overall, capital intervention is considerable to stabilise capital inequality and moderately affect the growth path meanwhile.

Table 8: Unusual way (static proportions) to decompose variances of key variables

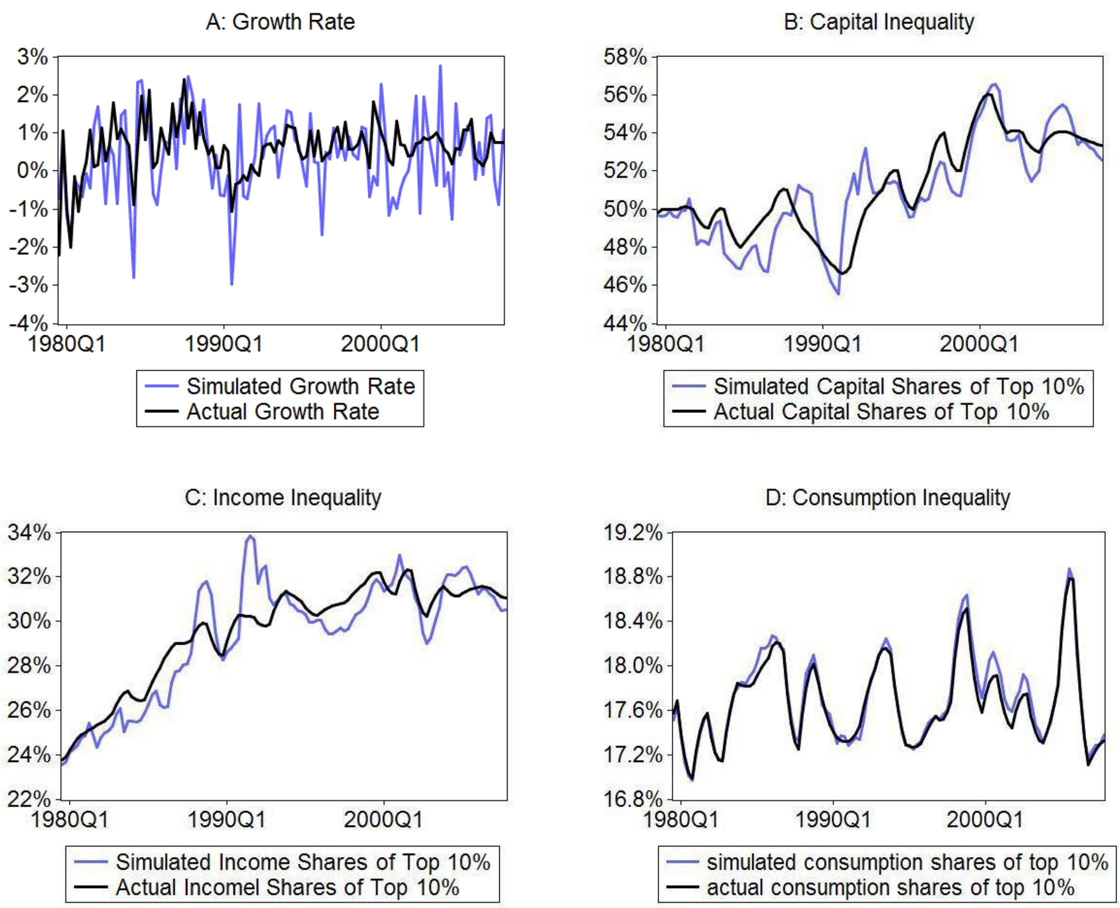
Variable									
R	0.17%	0.29%	0.22%	17.3%	0.26%	0.15%	13.6%	14.7%	53.4%
Y	9.94%	10.14%	10.6%	14.2%	10.1%	10.3%	8.37%	7.14%	19.2%
K	4.65%	4.99%	5.76%	18.7%	4.90%	5.26%	8.75%	4.58%	42.4%
C	9.85%	9.97%	10.3%	11.6%	9.96%	10.1%	8.66%	6.95%	22.7%
Y1	10.1%	10.3%	10.6%	12.4%	10.1%	10.4%	8.47%	9.46%	18.3%
Y2	9.71%	9.91%	10.5%	15.4%	9.88%	10.1%	8.67%	6.14%	19.7%
K1	5.41%	5.75%	6.50%	17.1%	5.62%	6.02%	8.85%	4.03%	40.7%
K2	3.75%	4.08%	4.81%	20.6%	4.04%	4.32%	9.20%	6.79%	42.4%
C1	9.88%	10.0%	10.3%	11.6%	10.0%	10.1%	8.71%	7.02%	22.4%
C2	9.85%	9.96%	10.3%	11.6%	9.95%	10.1%	8.65%	6.94%	22.7%
N1	3.40%	3.40%	3.40%	6.08%	3.47%	3.40%	19.9%	7.17%	49.8%
N2	2.95%	2.96%	2.99%	6.51%	2.96%	2.97%	4.71%	9.58%	64.4%
Growth	0.57%	0.57%	0.58%	2.03%	0.57%	0.57%	1.29%	1.66%	92.2%
InEq.K	10.9%	10.9%	10.9%	10.9%	10.9%	10.9%	11.44	11.3%	11.6%
InEq.Y	8.93%	8.91%	8.85%	8.85%	8.66%	8.85%	10.4%	17.7%	18.9%
InEq.C	11.0%	11.0%	11.0%	11.0%	11.6%	11.3%	11.0%	11.0%	10.9%

6.2.7 Comparisons between Actual and Simulated Data

We pick up one typical simulated sample to compare with actual data on 4 indicators, aggregate growth rate, capital inequality, income inequality and consumption inequality as shown in Figure 19. Here we focus on the UK economy before 2008 financial crisis. The present model could approximately generate the economic boom in 1980s, a gradual fall in 1990s and a moderate rise afterwards, but the simulated fluctuations are more drastic that the actual. Piketty (2014) summarise that inequality in developed countries across 21th century exhibit three properties, wealth inequality higher than income inequality, a significant increase in income inequality and a sharp reduction in wealth inequality before the last decades of 21th century.

Our model simulations could meet all the three properties and approximately perform similar tendencies of capital inequality and income inequality. Especially, the consumption inequality is outstandingly performed. For instance, the actual capital inequality from 1979Q2 to 2007Q4 has mean 51.22% and standard deviation (s.d) 2.43% while actual income inequality during that time has mean 29.55% and s.d 2.28%. Given the chosen simulation, capital inequality has mean 50.97% and s.d 2.63% while income inequality has mean 29.37% and s.d 2.69%. However, model behaviours cannot mimic all the cycles, like that simulated income inequality is obviously higher at the beginning of 1990s and much lower at the end of 1980s and at the second half of 1990s while simulated capital inequality has close performance. Overall, volatilities of simulated inequality indicators are found to be somehow higher by observing many simulations. Secondly, model generated growth rate is more fluctuated than the actual growth like that the s.d of growth rate in the selected sample is 1.08% while s.d of actual growth during that time is only 0.69%. This drastic fluctuation is probably caused by the bootstrapped innovations of the penalty rate. Recall the dominant effects of the penalty rate shocks on the variances across time of the aggregate growth analysed in the variance decomposition section. However, this fluctuation has a consistent trend with that of the actual growth rate and also has little effect on our Indirect Inference as aggregate growth is only one element out of the total 12 elements in our auxiliary model. But model simulation could basically mimic the cycles of aggregate growth such as the peak in the middle of 1980s and the bottom at the beginning of 1990s.

Figure 19: Comparisons between Actual data with one simulation



We can probably explain the sharp rise in income equality before 1990s based on the model growth mechanism as a reduction in entrepreneurship penalty rate. Recall the data corporation tax rate (CCB) shown in data section where a significant decline in CCB results in a fall in entrepreneurship penalty rate and then leads to an increase in individual productivity growth and income equality rises because given an aggregate reduction in entrepreneurship penalty rate, the rich benefit more and could gain higher growth. However, we can hardly explain why capital inequality has a decline during that time. This question will be reconsidered in next section.

With regard to the relation between capital inequality and growth, same as the benchmark tendency experiment, our model simulations based on actual heterogeneity also show an approximately stimulating effect of capital inequality on aggregate growth in UK from 1980s to 2000s (average of correlation coefficients between these two items is 0.2319 across 1,000 simulations).

6.3 Redistribution and Policy Measurement

6.3.1 Income transfer from the rich to the poor

In this section, we assess political interventions to redistribute capital or income allocations across agents. We have concluded in the tendency section that income tax interventions firstly do not affect the stability of behaviour tendency and secondly inequality might be improved to some extent while aggregate growth might be reduced. For the instruments of interventions, our empirical study on the benchmark model provides sufficient reason for the use of individual income tax. In previous section, it is shown that intervention on capital is a direct way to stabilise inequality and it could moderately affect. Conversely, other shocks may result in either too sever changes like the entrepreneurship penalty rate shock or too negligible influences like the output aggregation shock. Hence individual income tax regime which is also a usual instrument in practice is introduced to adjust both individual capital and aggregate one.

In the tendency section, three options for individual constant income tax regime are suggested, tax with no subsidy, tax with same per capita subsidy and an income transfer tax. The first one is employed in that section because subsidy or income transfer might switch the rich and the poor, given initially identical groups. However, this section studies actual 10%-90% income groups. As the rich is difficult to switch to the poor due to interventions, the latter two could be considered. We start with the income transfer tax. We follow the approximation introduced in that section in order to avoid the self-fulfilling trap where the rich (group one) is charged a constant income tax rate and the transfer to the poor per capita relative to the poor's income can be approximated by rate — . The linearised individual capital equations for the rich and the poor (26) and (27) have extra constant terms and — respectively on RHS. The linearised individual labour equations for the rich and the poor (30) and (31) now have extra constant terms and — respectively. The marginal effect of penalty rate on individual productivity growth for the rich and the poor becomes and respectively. The new constant terms could also be computed by "Type II Fix" step of our algorithm.

Theoretically, various tax rates represent different models and model parameters should have been re-estimated. However, that is against our intentions because we are interested in the effects of tax rates on present model simulation (i.e. same model with same parameter values). Hence, we keep all the parameters same as the benchmark model except which are calculated by same estimators of other parameters like but with tax rate used.

Table 9: Values of marginal penalty effect on productivities given various

Marginal penalty effect				
	0.7822	0.8144	0.8470	0.8790
	0.5534	0.5713	0.5884	0.6036

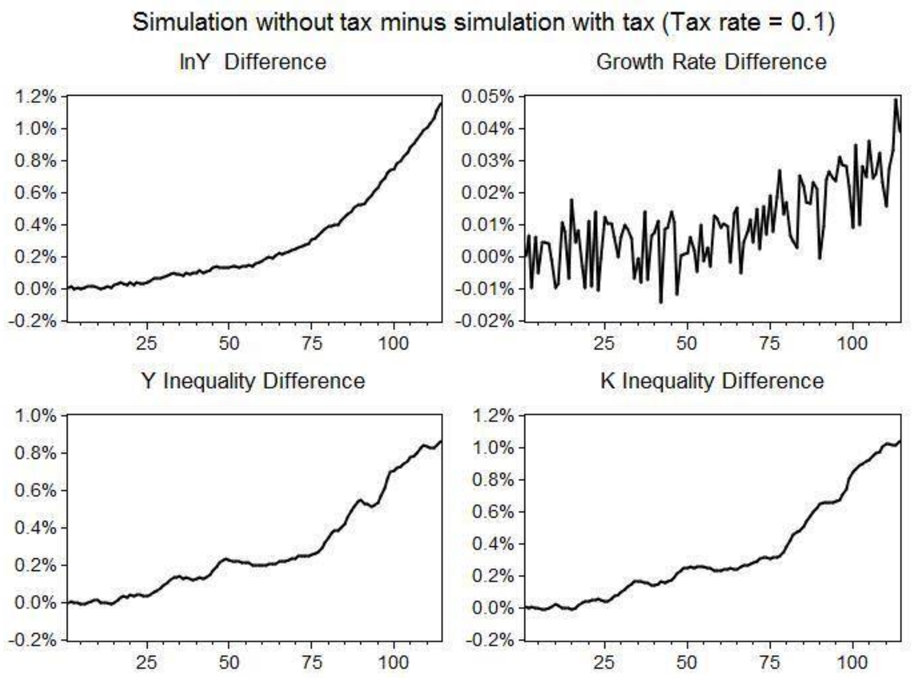
Practically, we apply Id-If Wald test on the three models with various tax transfer rates using same auxiliary model where we bootstrap structural innovations for simulation rather than randomly draw from certain distributions in the tendency section. Test results are shown in Table 10. Although strictly speaking three models are rejected on 95% confidence interval, the rejection percentile is not extremely high for models with tax rate equal to 0.1 and 0.2 respectively. On the other hand, if one model were not rejected, that would be problematic because that means the benchmark model might not be well identified or the auxiliary model may be inappropriately chosen so that more than one model pass the test. In fact, rejection of these models does not affect our analysis since we are interested in the ceteris paribus effects on changes in tax rates on economic behaviours. Conversely, renewed estimation may interfere redistribution analysis because one can hardly conclude whether should attribute behaviour change to tax change or to parameter change. Rejection on the model with positive tax rate given current parameter values does not mean that the structural model cannot be the one with tax. One could estimate all the key parameters and probably could find a structural model with positive tax rate because non-zero tax in fact is closer to reality.

Table 10: Id-If Wald test with various tax rates

Tax Rate	Wald Statistic	Transformed MD	Wald Percentile
0.1	28.669	1.664	95.165%
0.2	28.858	1.691	95.478%
0.3	31.547	1.862	96.744%

Now we use actual 10%-90% data, starting with the tax rate . Structural innovations are bootstrapped to simulate samples and innovations in same periods are chosen for both models with and without tax transfer in order to exclude effects caused by random shocks in different time periods. One group of simulations is shown by Figure 20 below. Both capital inequality and income inequality (the indicator is top 10% share over all) are reduced due to tax transfer and capital inequality changes more. Additionally, both aggregate output and growth rate have gradual declines in despite of fluctuations of growth rate. This seems consistent with the fact that the rich contribute more to entrepreneurship so that capital transfer to the poor lower their entrepreneurship incentives, resulting in lower aggregate growth. Our empirical finding on the role of tax in capital redistribution is also raised by Piketty (2014), although different types of tax are considered. He states that the primary reason for a high degree of wealth inequality is rate of return on capital greater economic growth rate during a long period. The third property of inequality introduced in previous section that wealth inequality experiences a reduction during the first half of the twentieth century now can be explained as a result of changes in tax policies such that return rates became lower.

Figure 20: Redistribution comparison when



What we want to know further is to what extent aggregate growth declines and two what extent social inequality could be improved. To answer these, we consider different tax rates, say and for example. Comparisons are summarised by Figure 21 and Figure 22. It is obvious that inequalities are improved better while both aggregate growth and aggregate output level fall further as tax rate increases. To make it clear, we select several time nodes with forecasting periods to compare deviations due to three tax rates in Table 11. Intuitively all the changes due to tax including reduction in growth and improvement in inequality exhibit a gradient growth as tax rate uniformly rises.

Figure 21: Redistribution comparison when

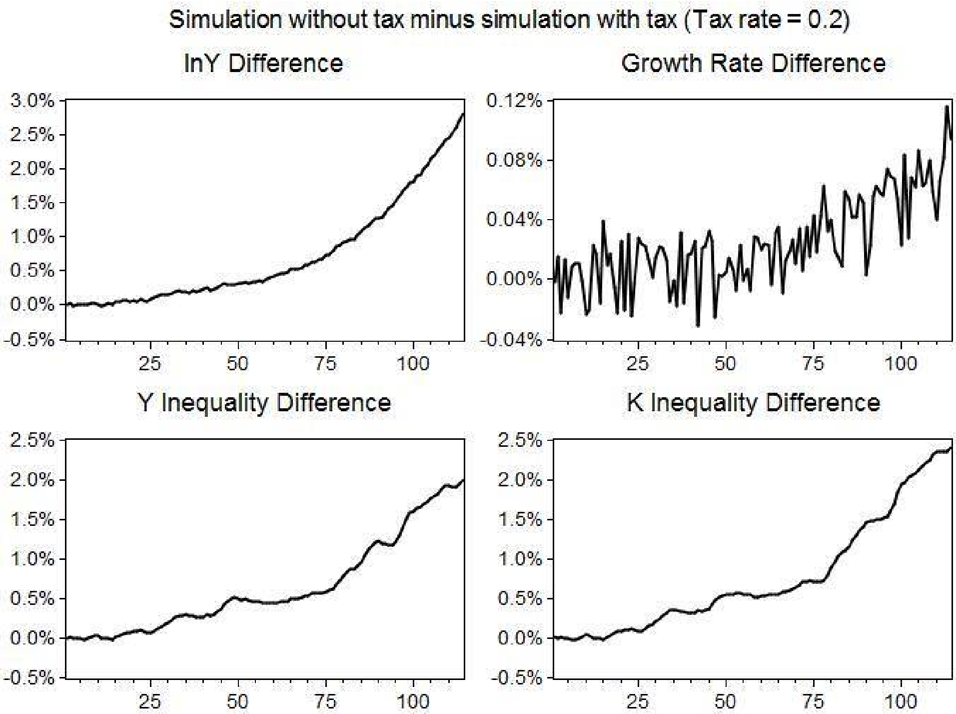


Figure 22: Redistribution comparison when

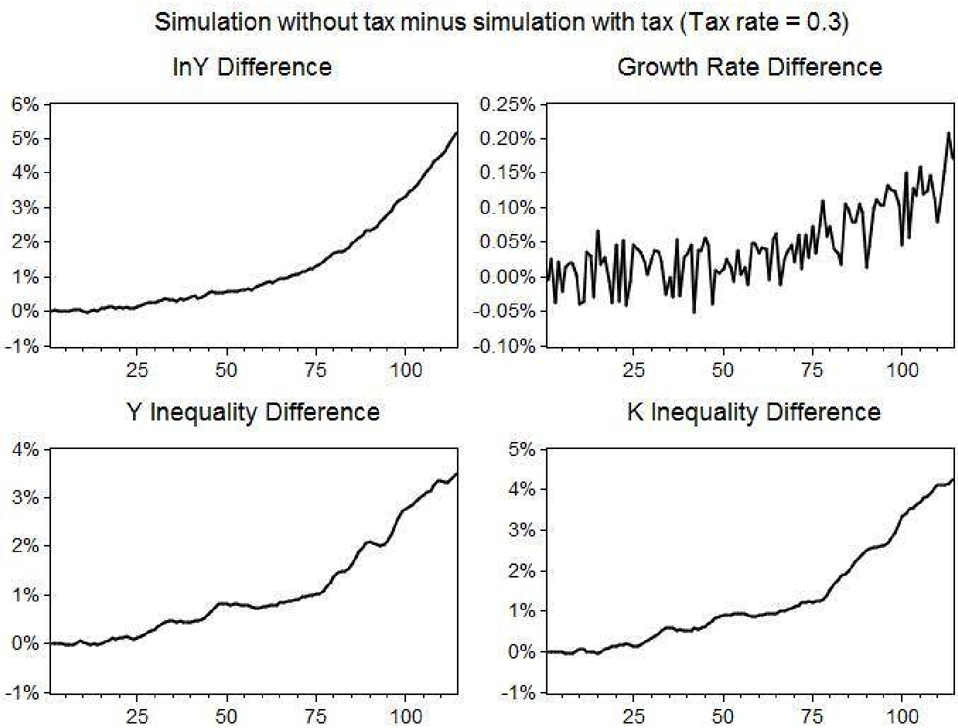


Table 11: Redistribution comparison using various tax rates

Deviations (no tax-tax)	Period / Tax Rate	25	50	75	100
Aggregate Output	0.1	0.03894%	0.12794%	0.30069%	0.74750%
	0.2	0.08383%	0.30811%	0.72281%	1.81187%
	0.3	0.13623%	0.55564%	1.31303%	3.32699%
Aggregate Growth	0.1	0.01235%	0.00119%	0.01905%	0.00882%
	0.2	0.02745%	0.00463%	0.04264%	0.02300%
	0.3	0.04597%	0.01120%	0.07280%	0.04622%
Income Inequality	0.1	0.03067%	0.22773%	0.25473%	0.70309%
	0.2	0.06651%	0.49778%	0.58279%	1.59609%
	0.3	0.10878%	0.82428%	1.01554%	2.75525%
Capital Inequality	0.1	0.03694%	0.24720%	0.30729%	0.84858%
	0.2	0.07984%	0.53942%	0.70329%	1.93011%
	0.3	0.13014%	0.89165%	1.22645%	3.34284%

Since it has been shown that capital inequality might has a stimulating effect on growth, we could loosely regress the deviation of growth rate in a tax model simulation away from a benchmark model simulation on the deviation of capital inequality away from benchmark simulation and constant marginal effects are found not rejectable.[60] We then simulate 1,000 samples for each model with same series of "time seed" and compute the average marginal effects of capital inequality deviations as 3.04% ( ), 3.27% ( ) and 3.50% ( ) when tax rate equals 0.1, 0.2, 0.3 respectively. For example, the estimated 3.04% when tax transfer rate is 0.1 tells that as each 1% share of capital held by the rich is reduced compared with the benchmark model, the economy will lose 0.0304% growth rate compared with the benchmark model.[61] Apparently the marginal cost of improvement on inequality is increasingly higher as tax rate rises. Moreover,

[60] Same "time seed" is used for both simulation and this marginal effect has nothing about the simple correlation coefficient between growth rate and inequality.

[61] One unit change in the inequality deviation results in 3.04% change in growth deviation and thus 1% change in inequality deviation results in 0.0304% change in growth deviation.

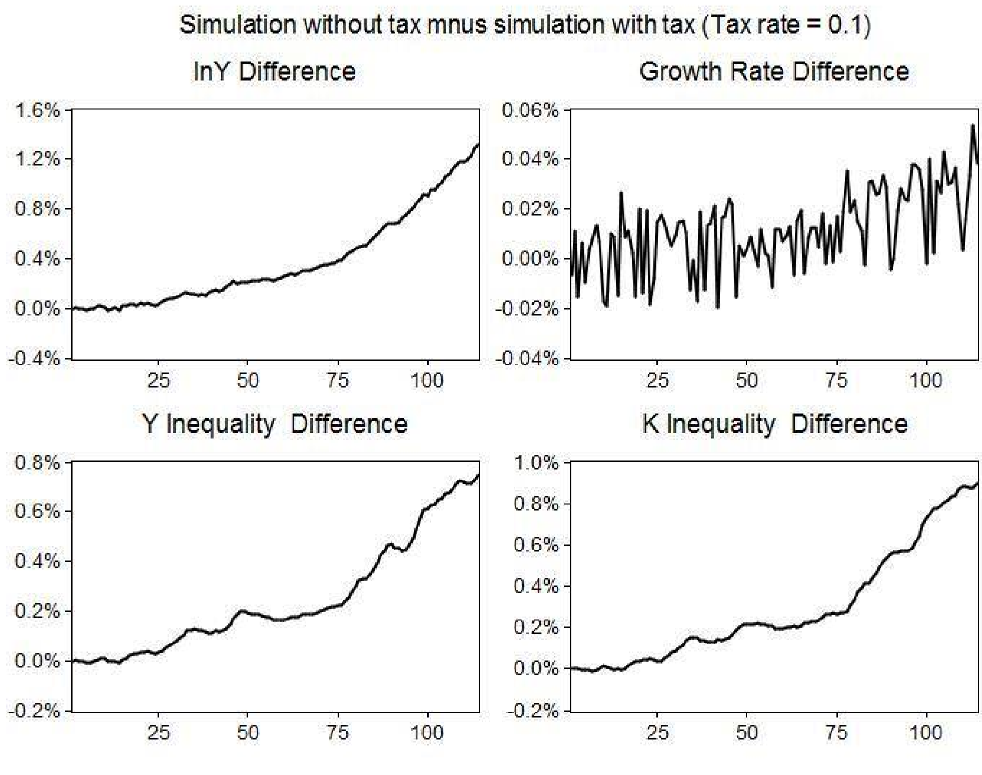
fluctuations of economic growth rates also become fiercer when tax rate increases (average s.d of growth rates are 0.000123, 0.000292 and 0.000527 at tax rate 0.1, 0.2 and 0.3 respectively).

The policy implication behind these findings is that policy makers need to be cautious if they attempt to improve inequality status by tax transfer. Although higher tax rate levied on the rich could rapidly narrow the wealth gap, the resulted risks of economic recession and economic instability are quite high. As tax rate increases, the cost of reducing inequality, reduction in growth, has a gradiently increasing trend rather than a uniformly increasing trend. This nonlinear tax effect is also found by Jaimovich and Rebelo (2012) who indicate that the negative tax effect on growth becomes more significant with an increase in tax rate. Hence the attempt to immediately equalise allocations across individuals by a considerably high tax rate is inappropriate. Although a lower long-run tax rate still has negative effects on economic growth, an appropriate tax transfer rate could gain a gradual and moderate wealth gap narrowing with a relatively low cost of growth. This trade-off between economic growth and social inequality needs deeper considerations and studies in the future.

6.3.2 Robustness of redistribution

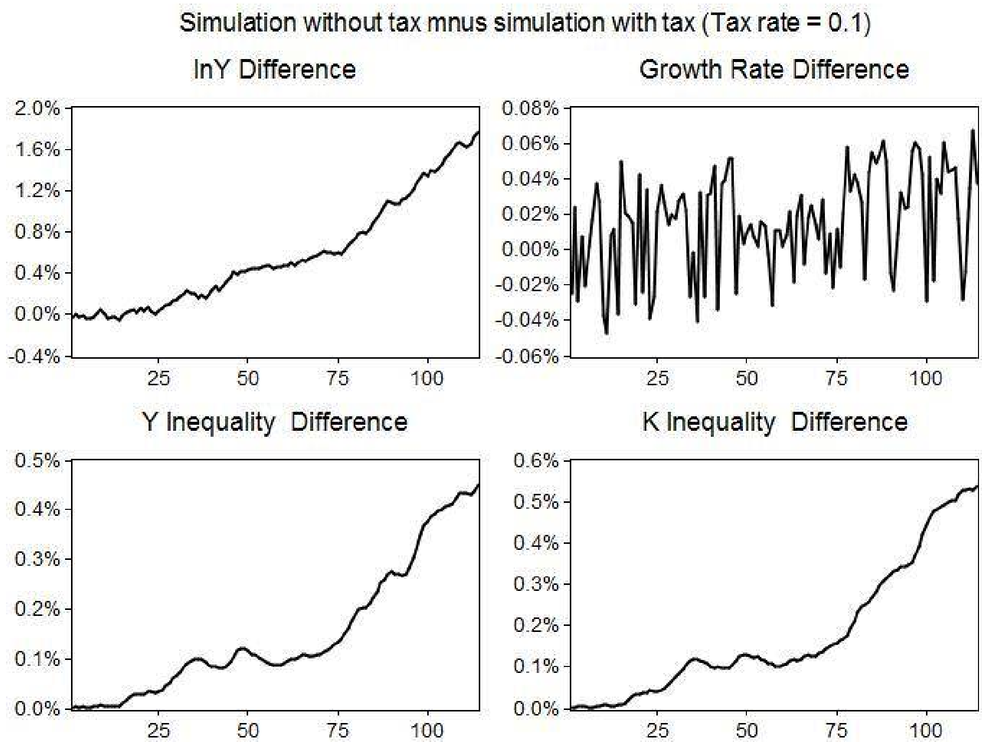
In this section, we discuss about redistribution effects of various tax regimes. As suggested in the tendency section, one could also levy proportional tax rate on both groups either with subsidy or not. Now consider the regime where a constant income tax rate, say for example, is still solely enforced on the rich but with no transfer to the poor. Model equations as well as coefficients now change to the ones such that the rich have same in the previous section while the poor have same in the tendency section. For convenient comparison, same "time seed" used for the selected simulations by Figures 20-22 in previous section is chosen for current bootstrapping. Figure 23 also indicates that the cost of reducing inequality is still loss of growth. Additionally, this tax instrument is less efficient than the income transfer tax in previous section because inequality is reduced to less extent while growth falls more. Quantitatively, the average marginal effect of capital inequality deviation on growth rate deviation across 1,000 samples is 3.77% ( ) much higher than 3.04% with same tax rate but with income transfer to the poor. This finding makes sense because no transfer implies total loss of aggregate capital accumulation used for future production in this model and therefore one can consider the higher marginal loss as the sum of both cost of inequality reduction and aggregate capital loss.

Figure 23: Compare income tax on the rich but no transfer with benchmark model



Now we consider the tax regime where both the rich and the poor are charged a constant income tax rate, say but with no individual subsidy (tax revenue all funds government spending). Model equations now take completely same forms as those in tendency section. Figure 24 shows results as what we expect that inequality is reduced to further less extent while loss of growth is much higher and also quite fluctuated.[62] This is also one of the reasons why we do not start with this tax regime in previous section. Tax without any subsidy is also harmful to the economy by reducing aggregate capital accumulation.

Figure 24: Compare income tax on both but no transfer with benchmark model



A constant proportional income tax with any other subsidy scheme can be considered as a combination of the tax with no subsidy and the transfer tax. Hence the

[62] Marginal effect of inequality deviation of growth deviation is not used here for illustration because the regression is lack of good fitness, although the estimated marginal effect is 7.50% much higher as expected.

empirical experiment on this tax is skipped here and undoubtedly it will have a less efficient redistribution effect than the transfer tax. To conclude, income tax rate regime regardless of the utilisation of tax revenue generally has a redistribution effect that inequality could be reduced with a certain loss of aggregate growth. However, those tax regimes with tax revenue returned to individuals have less harm to economic growth and those regimes with tax revenue more transferred to the poor are more efficient to reduce inequality. These findings are all consistent with our standpoint that there is a trade-off between inequality and growth.

7 Conclusion

This paper establishes a theoretical framework to consider the effect of wealth (capital) inequality on economic growth by introducing an endogenous mechanism of entrepreneurship incentive stimulating individual productivity growth and an effect of capital distribution on entrepreneurship penalty. First of all, our model could generate a stable tendency of the relation between capital inequality and economic growth and this tendency is almost independent of parameter values and population shares of individual groups, which make it also applicable to different countries and different population group-segmentations. When consider the 10%-90% income segmentation, model well fits main characteristics of actual data in UK from 1978Q1 to 2007Q4, such as wealth inequality higher than income inequality, a significant increase in income inequality and a sharp reduction in wealth inequality at the beginning of 1990s. By analyzing the influence of external shocks on deviations and volatilities of endogenous variables, we find that some shocks have fairly large impact on both aggregate growth and inequality such as the shock on entrepreneurship penalty rate while some shocks have almost no effect on one aspect or on both, like aggregate consumption shock and output aggregation shock. And some shocks could moderately affect both, like capital aggregation shock. This finding implies that if policy makers attempt to control inequality or stimulate economic growth by adjusting entrepreneurship penalty rate like the corporation tax rate, economy may have considerable fluctuations. Conversely, if the intervention is conducted on capital (aggregate or individual) like income tax rate, there might be acceptable and moderate effect on both inequality and growth. More importantly, we find capital inequality is most likely to stimulate economic growth and thus policy makers have to face a trade-off between equalisation of wealth allocation and economic stabilisation. Although our model only considers the mechanism of how inequality affects growth, we can hardly deny that growth might also have a positive effect on inequality because growth and inequality always exhibit a positive relation in our benchmark model, tendency experiment and redistribution experiment with no exception. Lastly, according to the comparisons on tax or not, and different income tax regimes, we find the tax regime which transfers income from the rich to the poor is more efficient to balance inequality and growth. Moreover, as tax rate rises, the cost of reducing inequality, reduction in growth, has a gradiently increasing trend instead of a uniformly increasing trend and this suggests that an appropriately low tax rate instead of a high rate is prior for policy makers.

Present study also has some limitations and shortcomings. In terms of model setting, we assume that only a few agents within one group may switch group in each period and this ignores cross-group behaviors and their sequential effects; although this assumption is reasonable for an infinite number of agents in each group. Individual bonds are eliminated in linearised model equations due to the unavailability of individual credit data which finally results in an inevitable approximation when consider tax subsidy or income transfer. Some parameters are set same across individuals for simplification such as the marginal effect of entrepreneurship incentives on individual productivity growth rate which in reality for the rich is generally higher than the poor, implying that some empirical results might be more significant in reality. In the aspect of data, this paper simplifies wealth into capital, probably resulting in measurement errors in despite of small differences because wealth also contains income and individual credit in each period. Regarding model fitness, although model could fit behaviours of most variables and especially their interactions, the fitness of interest rate behaviour is not good enough and this also causes behaviours of individual capital not to be well fitted (despite of the good fitness of the relative relation between individual capitals). Additionally, both model simulated inequality and growth are somehow more volatile than actual ones. In the future on the basis of sufficient micro-data, current study could be developed by employing individual credit equations in the list of model equations in order to make different behaviours of real interest rate and by distinguishing individual capital and wealth, etc. furthermore, we plan to apply time-variant resource shares in linearised equations to consider cross-group behaviours. In terms of redistribution, more instruments besides income tax such as asset return tax can be considered. Last, the mechanism of the growth effect on inequality can be modeled in the future.

Bibliography

[1] Acemoglu, D. and Robinson, J.A. (2002). The Political Economy of the Kuznets Curve. *Review of Development Economics,* 6(2), pp.183-203.

[2] Aghion, P. and Bolton, P. (1997). A Theory of Trickle-down Growth and Development. *The Review of Economic Studies,* 64(2), pp.151-172.

[3] Aghion, P., Caroli, E. and García-Peñalosa. C. (1999). Inequality and Economic Growth: The Perspective of the New Growth Theories. *Journal of Economic Literature,* 37(4), pp.1615-1660.

[4] Alesina, A. and Rodrik, D. (1994). Distributive Politics and Economic Growth. *Quarterly Journal of Economics,* 109(2), pp.465-490.

[5] Algan, Y., Allais, O. and Den Haan, W. J. (2008). Solving Heterogeneous-agent Models with Parameterized Cross-sectional Distributions. *Journal of Economic Dynamics and Control,* 32(3), pp.875-908.

[6] Algan, Y., Allais, O. and Den Haan, W. J. and Rendahl, P. (2014). Solving and Simulating Models with Heterogeneous Agents and Aggregate Uncertainty. In: Schmedders, K. and Judd, K. L., *Handbook of Computational Economics,* North Holland, vol. 3, pp.277-324.

[7] Aiyagari, S. and Gertler, M. (1991). Asset Returns with Transactions Costs and Uninsured Individual Risk. *Journal of Monetary Economics,* 27(3), pp.311-331.

[8] Aiyagari, S. R. (1994). Uninsured Idiosyncratic Risks and Aggregate Saving. *Quarterly Journal of Economics,* 109, pp.659-684.

[9] Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies,* 58, pp.277-297.

[10] Barro, R.J. (2000). Inequality and Growth in a Panel of Countries. *Journal of Economic Growth,* 5(1), pp.5-32.

[11] Bagchi, S. and Svejnar, J. (2015). Does Wealth Inequality Matter for Growth? The Effect of Billionaire Wealth, Income Distribution, and Poverty. *Journal of Comparative Economics,* 43(3), pp.505-530.

[12] Banerjee, A. V. and Duflo, E. (2003). Inequality and Growth: What Can the Data Say? *Journal of Economic Growth,* 8(3), pp.267–99.

[13] Bénabou, R. (1996). Inequality and Growth. *NBER Working Paper,* 5658.

[14] Benhabib, J. (2003). The Tradeoff between Inequality and Growth. *Annals of Economics and Finance,* 4(2), pp.329-345.

[15] Benhabib, J. and Rustichini, A. (1996). Social Conflict and Growth. *Journal of Economic Growth,* 1(1), 125-142.

[16] Bewley, T. F. (1977). The Permanent Income Hypothesis: A Theoretical Formulation. *Journal of Economic Theory,* 16(2), pp.252-295.

[17] Bewley, T. F. (1980). Stationary Equilibrium. *Journal of Economic Theory,* 24(2),

pp.265-295.

[18] Bewley, T. F. (1983). A Difficulty with the Optimum Quantity of Money. *Econometrica,* 51(5), pp.1485-1504.

[19] Bhattacharya, J. (1998). Credit Market Imperfections, Income Distribution, and Capital Accumulation. *Economic Theory,* 11(1), pp.171-200.

[20] Bohacek, R. and Kejak, M. (2005). Projection Methods for Economies with Heterogeneous Agents. *CERGE-EI Working Papers,* 258.

[21] Burdett, K., Lagos, R. and Wright, R. (2003). Crime, Inequality, and Unemployment. *The American Economic Review,* 93(5), pp.1764–1777.

[22] Caballero, R. (1990). Consumption Puzzles and Precautionary Savings. *Journal of Monetary Economics,* 25(1), pp.113-136.

[23] Canova, F. and Sala, L. (2009). Back to Square One: Identification Issues in DSGE Models. *Journal of Monetary Economics,* 56(4), pp.431-449.

[24] Castelló-Climen, A. (2010). Inequality and Growth in Advanced Economies: An Empirical Investigation. *The Journal of Economic Inequality,* 8(3), pp.293-321.

[25] Chang, Y. and Kim, S. (2007). Heterogeneity and Aggregation: Implications for Labor Market Fluctuations. *The American Economic Review,* 97(5), pp. 1939-1956.

[26] Dahan, M and Tsiddon, D. (1998). Demographic Transition, Income Distribution, and Economic Growth. *Journal of Economic Growth,* 3(1), pp.29-52.

[27] Davidson, R. and MacKinnon, J. G. (2004). *Econometric Theory and Methods.* Oxford University Press, Chapter 13.

[28] Davidson, J., Meenagh, D., Minford, P. and Wickens, M. (2010). Why Crisis Happen - Nonstationary Macroeconomics. *Cardiff Economics Working Papers,* E2010/13.

[29] Dave, C. and De Jong, D. N. (2007). *Structural Macroeconomics.* Princeton University Press, Chapter 2.

[30] Deininger, K. and Olinto, P. (2000). Asset Distribution, Inequality, and Growth. *Policy Research Working Paper,* 2376.

[31] Deininger, K. and Squire, L. (1996). A New Data Set Measuring Income Inequality. *The World Bank Economic Review,* 10(3), pp.565-591.

[32] Den Haan, W. J. (1996). Heterogeneity, Aggregate Uncertainty, and the Short-Term Interest Rate. *Journal of Business & Economic Statistics,* 14(4), pp.399-411.

[33] Den Haan, W. J. (2010). Assessing the Accuracy of the Aggregate Law of Motion in Models with Heterogeneous Agents. *Journal of Economic Dynamics and Control,* 34(1), pp.79-99.

[34] Diaz-Gimenez, J. and Prescott, E. C. (1992). Liquidity Constraints in Economies with Aggregate Fluctuations: A Quantitative Exploration. *Federal Reserve Bank of Minneapolis,* Staff Report 149.

[35] Ehrhart, C. (2009). The Effects of Inequality on Growth: A Survey of the Theoretical and Empirical Literature. *ECINEQ Working Paper,* 2009-107.

[36] Fishman, A. and Simhon, A. (2002). The Division of Labor, Inequality and Growth.

Journal of Economic Growth, 7(2), pp.117-136.

[37] Foellmi, R. and Zweimüller, J. (2004). Income Distribution and Demand-induced Innovations. *Institute for Empirical Research in Economics, University of Zurich Working Paper,* 212.

[38] Forbes, K.J. (2000). A Reassessment of the Relationship between Inequality and Growth. *The American Economic Review,* 90(4), pp.869-887.

[39] Galor, O. and Moav, O. (2004). From Physical toHuman Capital Accumulation: Inequality and the process of Development. *Review of Economic Studies,* 71, pp.1001-1026.

[40] Galor, O. and Tsiddon, D. (1997). Technological Progress, Mobility, and Economic Growth. *The American Economic Review,* 87(3), pp.363-382.

[41] Gernigon, B., Odero, A. and Guido, H. (2000). ILO Principles Concerning Collective Bargaining. *International Labour Review,* 139(1), pp. 33-55.

[42] Giacomini, R. (2013). The relationship between DSGE and VAR models. In: Fomby, T.B. et al., *Advances in Econometrics,* Emerald Publishing Group Limited, vol. 31.

[43] Gourieroux, C., Monfort, A. and Renault, E. (1993). Indirect Inference. *Journal of Applied Econometrics,* 8, pp.S85–S118.

[44] Greene, William H. (2002). *Econometric Analysis, 5th edition.* Prentice Hall, Appendix B.

[45] Halter, D., Oechslin, M. and Zweimüller, J. (2014). Inequality and Growth: the Neglected Time Dimension. *Journal of Economic Growth,* 19(1), pp.81-104.

[46] Hansen, Gary D. (l985). Indivisible Labor and the Business Cycle. *Journal of Monetary Economics,* 16(3), pp.309-327.

[47] Hayashi, F. (1982). Tobin's Marginal q and Average q: A Neoclassical Interpretation. *Econometrica,* 50(1), pp.213-224.

[48] Huggett, M. (1993). The Risk-free Rate in Heterogeneous-agent Incomplete-insurance Economies. *Journal of Economic Dynamics and Control,* 17, pp.953-969.

[49] Jaimovich, N. and Rebelo, S. (2012). Non-linear Effects of Taxation on Growth. *NBER Working Paper,* 18473.

[50] Johnson. D. S., Aragon, C. R., McGeoch, L. A. and Schevon, C. (1989). Optimization by Simulated Annealing: An Experimental Evaluation, Part I: Graph Partitioning. *Operations Research,* 37(6), pp.865-892.

[51] Keane, M. and Smith, A. A., Jr. (2003). Generalized Indirect Inference for Discrete Choice Model. *Manuscript, Yale University*.

[52] Keefer, P. and Knack, S. (2002). Polarization, Politics and Property Rights: Links between Inequality and Growth. *Public Choice,* 111, pp.127-154.

[53] Knowles, S. (2001). Inequality and Economic Growth: The Empirical Relationship Reconsidered in the Light of Comparable Data. *CREDIT Research Paper,* 01/03.

[54] Kolev, G and Niehues, J. (2016). The Inequality-Growth Relationship: An Empirical

Reassessment. *Cologne Institute for Economic Research IW Report,* 7.

[55] Krusell, P. and Smith, A. A., Jr. (1998). Income and Wealth Heterogeneity in the Macroeconomy. *The Journal of Political Economy,* 106(5), pp.867–896.

[56] Krusell, P. and Smith, A. A., Jr. (2006). Quantitative Macroeconomic Models with Heterogeneous Agents. In Blundell, R., Newey, W.K. and T. Persson, T., *Advances in Economics and Econometrics: Theory and Applications,* Ninth World Congress, Cambridge University Press, vol. 1, pp.298–340.

[57] Kucera, D. (2002). The Effects of Wealth and Gender Inequality on Economic Growth: A Survey for Research Empirical Studies. *International Institute for Labour Studies Geneva Discussion Paper,* 136.

[58] Kuznets, S. (1955). Economic Growth and Income Inequality. *The American Economic Review* 45(1), pp.1-28.

[59] Kydland, Finn E. (1984). Labor Force Heterogeneity and the Business Cycle. *Carnegie-Rochester Conference Series on Public Policy,* 21, pp.173-208.

[60] Le, V. P. M., Meenagh, D., Minford, P. and Wickens, M., (2011). How Much Nominal Rigidity is there in the US Economy? Testing a New Keynesian Model Using Indirect Inference. *Journal of Economic Dynamics and Control,* 35(12), pp.2078-2104.

[61] Le, V. P. M., Meenagh, D. (2013). Testing and Estimating Models Using Indirect Inference. *Cardiff Economics Working Papers,* E2013/8.

[62] Le, V. P. M., Meenagh, D., Minford, P., Wickens, M. and Xu, Y. (2015). Testing Macro Models by Indirect Inference - A Survey for Users. *Cardiff Economics Working Papers,* E2015/9.

[63] Mankiw, N. G. (1986). The Equity Premium and the Concentration of Aggregate Shocks. *Journal of Financial Economics,* 17(1), pp.211-219.

[64] Matthews, K. G. P. and Marwaha, S. (1979). Numerical Properties of the LITP Model, *SSRC Project Working Paper,* 7903 (University of Liverpool).

[65] McCallum, B. T. (1976). Rational Expectations and the Natural Rate Hypothesis: Some Consistent Estimates. *Econometrica,* 44(1), pp.43-52.

[66] Meenagh, D., Minford, P., Nowell, E. and Sofat, P. (2010). Can a Real Business Cycle Model without Price and Wage Stickiness Explain UK Real Exchange Rate Behaviour? *Journal of International Money and Finance,* 29(6), pp.1131-1150.

[67] Meenagh, D., Minford, P. and Wickens, M. (2008). Testing a DSGE Model of the EU Using Indirect Inference. *Cardiff Economics Working Papers,* E2008/11.

[68] Meenagh, D., Minford, P. and Wickens, M. (2012). Testing Macroeconomic Models by Indirect Inference on Unfiltered Data. *Cardiff Economics Working Papers,* E2012/17.

[69] Minford, L. (2016). Tax, Regulation and Growth: A Case Study of the UK. *Cardiff Economics Working Papers,* E2015/16.

[70] Minford, P., Matthews, K. and Marwaha, S. (1979). Terminal Conditions as a Means of Ensuring Unique Solutions for Rational Expectations Models with Forward Expectation. *Economic Letter,* 4(2), pp.117-120.

[71] Minford, P., Marwaha, S., Matthews, K. and Sprague, A. (1984). The Liverpool Macroeconomic Model of the United Kingdom. *Economic Modelling,* 1(1), pp.24-62.

[72] Mukoyama, T. and Sahin, A. (2006). Costs of Business Cycles for Unskilled Workers. *Journal of Monetary Economics,* 53(8), pp.2179-2193.

[73] Neves, P. C. and Silva, S.T. (2014). Inequality and Growth: Uncovering the Main Conclusions from the Empirics. *The Journal of Development Studies,* 50(1), pp.1-21.

[74] Ostry, J.D., Berg, A. and Tsangarides, C.G. (2014). Redistribution, Inequality, and Growth. *IMF Staff Discussion Notes,* 14/02.

[75] Oulton, N. and Srinivasan, S. (2003). Capital Stocks, Capital Services, and Depreciation: An Integrated Framework, *Bank of England Working Paper,* 192.

[76] Perotti, R. (1996). Growth, Income Distribution, and Democracy: What the Data Say. *Journal of Economic Growth,* 1(2), pp.149-187.

[77] Persson, T. and Tabellini, G. (1994). Is Inequality Harmful for Growth? *The American Economic Review,* 84(3), pp.600-621.

[78] Piketty T. (1997). The Dynamics of the Wealth Distribution and the Interest Rate with Credit Rationing. *Review of Economic Studies,* 64, pp.173-189.

[79] Piketty, T. (2014). *Capital in the Twenty-First Century,* Harvard University Press.

[80] Piketty, T. and Zucman, G. (2014). Capital is back: Wealth-income Ratios in Rich Countries 1700-2010. *Quarterly Journal of Economics,* 129(3), pp.1255-1310.

[81] Preston, B., and Roca, M. (2007). Incomplete Markets, Heterogeneity and Macroeconomic Dynamics, *NBER Working Papers,* 13260.

[82] Reiter, M. (2009). Solving Heterogeneous-Agent Models by Projection. *Journal of Economic Dynamics and Control,* 33(3), pp.649-665.

[83] Rodríguez, S. B., Díaz-Gimenéz, J., Quadrini, V. and Ríos-Rull, J. (2002). Updated Facts on the U.S. Distribution of Earnings, Income and Wealth. *Federal Reserve of Minneapolis Quarterly Review,* Vol. 26(3), pp.2-35.

[84] Scheinkman, J.A. and Weiss, L. (1986). Borrowing Constraints and Aggregate Economic Activity. *Econometrica,* 54(1), pp.23-45.

[85] Smith, Anthony A., Jr. (1993). Estimating Nonlinear Time–series Models Using Simulated Vector Autoregressions. *Journal of Applied Econometrics,* 8(S1), pp.S63–S84.

[86] Stiglitz, J.E. (2012). *The Price of Inequality: How Today's Divided Society Endangers Our Future, 1st Edition.* W. W. Norton & Company.

[87] Telmer, C. (1993). Asset Pricing Puzzles and Incomplete Markets. *The Journal of Finance,* 48(5), pp.1803-1832.

[88] Voitchovsky, S. (2005). Does the Profile of Income Inequality Matter for Economic Growth?: Distinguishing Between the Effects of Inequality in Different Parts of the Income Distribution. *Journal of Economic Growth,* 10, pp.273-296.

[89] Voitchovsky, S. (2011). Inequality and Economic Growth. In: Nolan, B., Salverda, W. and Smeeding, T. M., *The Oxford Handbook of Economic Inequality,* Oxford

University Press.

[90] Wickens, M. (2012). *Macroeconomic Theory: A Dynamic General Equilibrium Approach, 2nd Edition*. Princeton University Press, chapter 16.7.

[91] Wickens, M. (2014). How useful are DSGE Macroeconomic Models for Forecasting? *Open Economies Review,* 25(1), pp.171-193.

[92] Young, Eric R. (2010). Solving the Incomplete Markets Model with Aggregate Uncertainty Using the Krusell–Smith Algorithm and Non-stochastic Simulations. *Journal of Economic Dynamics and Control,* 34(1), pp.36-41.

[93] Zweimüller, J. (2000). Inequality, Redistribution, and Economic Growth. *Institute for Empirical Research in Economics, University of Zurich Working Paper,* 31.

Appendix

Appendix 1: Prove could also be approximated to a random walk

Define and rewrite individual budget constraint as

(AP1. 1)

The condition of no Ponzi that present value of all the future increment of bonds should be equal to implies

Euler equation (5) can be approximated by

For simplification, we set which will also be employed in the empirical study. Then (AP1.3) can be simplified to . In fact, we could also use this simplified form of (Ap1.3) as long as is close to unity which is true in many empirical papers.

Now (AP1.2) can be written as

where (AP1.4) means that current consumption should equal the sum of current bond gross return and present value of permanent income in all the future, denoted by , discounted by the rate .

The steady-state condition of bonds is and generally follows an AR process before steady state, denoted by which is non-stationary due to the coefficient .[63] This can be transformed to which implies that before steady state approximately follow a unit root because the random growth rate is generally close to . Hence, (AP1.4)

[63] For example, if consider an open economy, generally equal the net export which is a random in terms of output.

implies that could also be approximated to a random walk.

Appendix 2: Derive the relation between the aggregate growth and inequality

Given the linearised aggregate output equation, aggregate growth is

Individual output growth is yielded using (24) and (25)

Substituting out , and using (26)-(27), (30)-(31) and (32)-(33) yields

Taking on the equation above yields

Substituting out , and using (14) and (28) yields

To get rid of the past , set to obtain [64]

Aggregating the equations above across individuals with assumption that is same dnoted by across groups for simplification yields

We set where also measures capital inequality and define the aggregate term . Then

Note if while if and means group 2 has higher average capital, still inequality. It concludes that has minimum at perfect equality ( ). Equation (AP2.1) now can be rewritten as

The aggregate growth above is still complicated due to both current and lagged inequality terms. However, if we consider a mid-term or a long-term growth by summing up temporary growth rates within a long time period, we yield the following [65]

---

[64] This assumption will not change the direction of inequality effect on growth but make this effect magnified.

[65] For the long-run growth rate, we use an approximation that

Since the other growth determinant real interest rate endogenously depends on lagged output and capital, both short-run and long-run growth rate in (AP2.2) and (AP2.3) could be approximated as a reduced form of lagged output, capital and inequality which is the usual form of existing empirical regression studies. Note that and for long term. Since , this long-run growth rate is minimised when capital distribution is always perfectly equality. Importantly, the stimulating effect in very short term is not as clear as that in long term because when inequality stimulates entrepreneurship incentives, labour input in production will have a decline which implies a negative but temporary effect on growth.

Appendix 3: Data source of ITB, IPW and AMW

ITB does not include data on wealth distribution. IPW only covers annual data from 2001 to 2015. And AMW collects annual data from 1976 to 2005. If the 50%-50% segmentation is considered, one has to combine IPW and AMW to estimate because WID only provides top 1% and top 10% shares. Furthermore, the wealth share of bottom 50% agents is quite small, inconvenient for model use.

Source of ITB:
https://www.ons.gov.uk/peoplepopulationandcommunity/personalandhouseholdfinances/incomeandwealth/datasets/averageincomestaxesandbenefitsbydecilegroupsofallhouseholds

Source of IPW:
https://www.gov.uk/government/statistics/table-138-identified-personal-wealth-analysis-by-decile

Source of AMW:
http://webarchive.nationalarchives.gov.uk/20120406010108/http://www.hmrc.gov.uk/stats/personal_wealth/archive.htm

Appendix 4: Impulse Responses in the benchmark model not used in the body

Figure 25: IRFs to a +5% one-period aggregate output shock

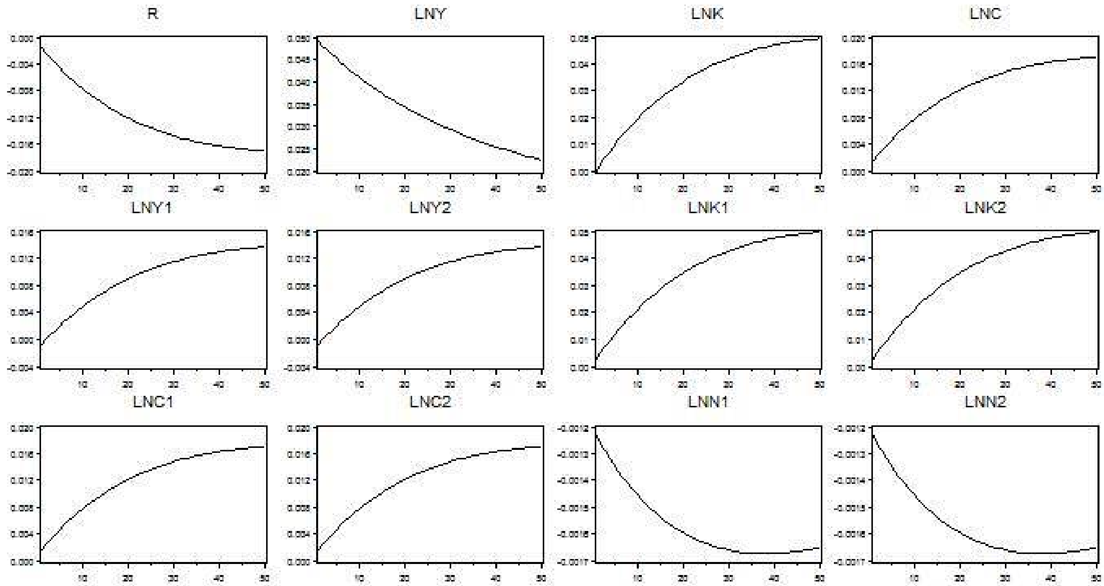


Figure 26: IRs to a +5% one-period aggregate capital shock

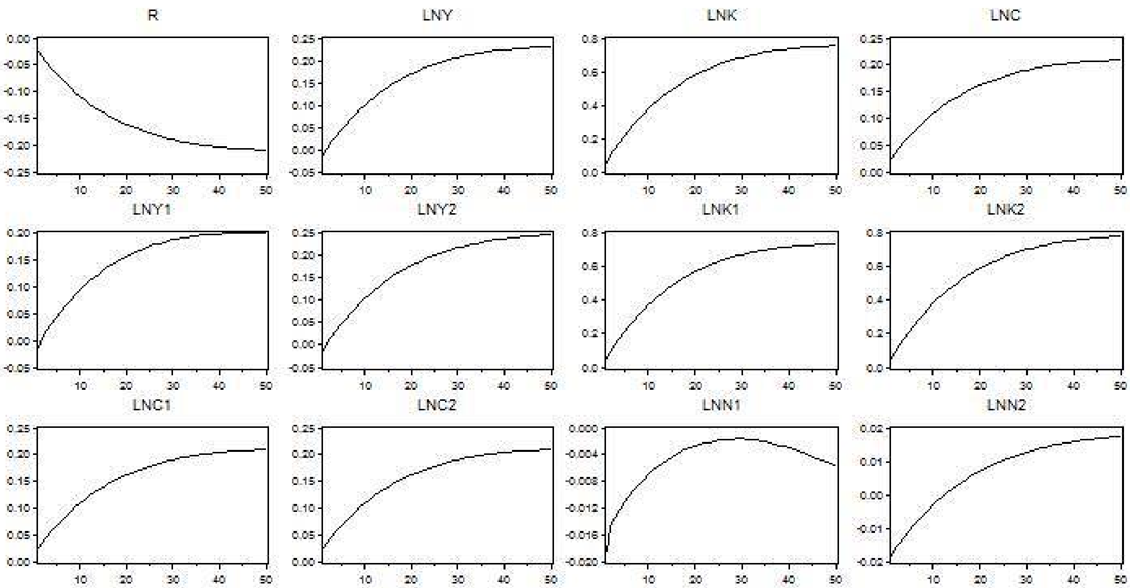
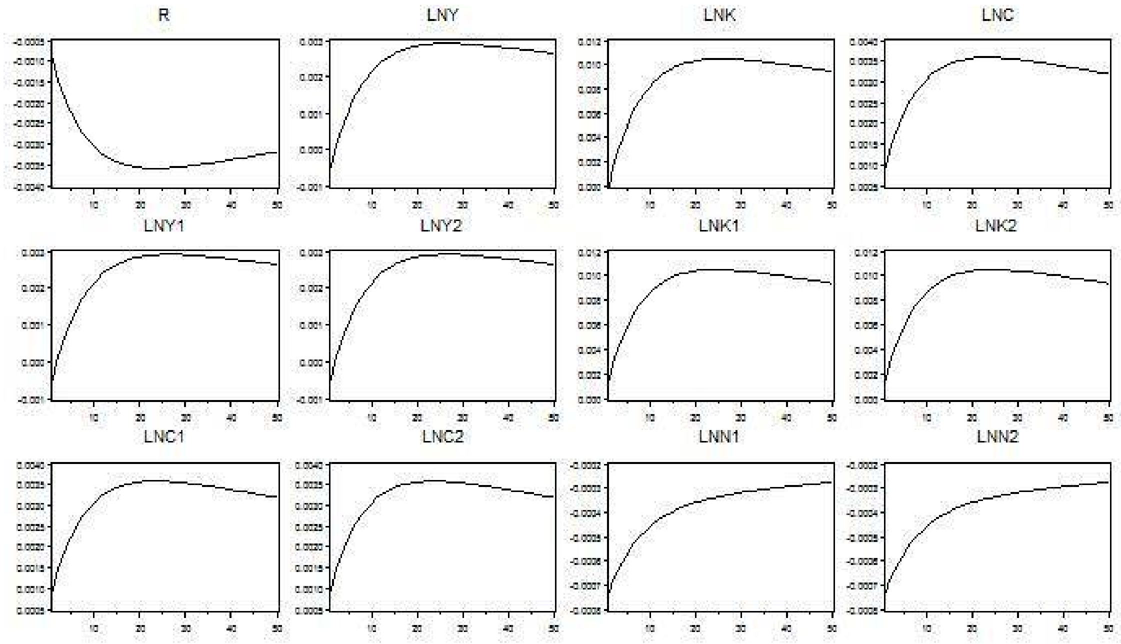


Figure 27: IRs to a +5% one-period aggregate consumption (market clearing) shock



Appendix 5: Variance decomposition using the IRs way

Table 12: Usual way (IRs way) to decompose variances of interest rate

Periods									
1	1.38%	31.34%	0.48%	46.34%	1.17%	0.24%	6.64%	4.30%	8.13%
2	2.97%	63.11%	0.86%	23.98%	0.45%	0.45%	4.19%	1.20%	2.78%
3	3.51%	72.30%	0.87%	6.88%	0.07%	0.48%	1.71%	0.04%	14.14%
4	3.73%	74.50%	0.81%	1.10%	0.00%	0.47%	0.68%	0.15%	18.56%
5	3.97%	76.65%	0.76%	0.01%	0.02%	0.46%	0.26%	0.66%	17.22%
6	4.24%	79.31%	0.73%	0.84%	0.05%	0.45%	0.08%	1.27%	13.03%
7	4.52%	81.74%	0.69%	2.57%	0.10%	0.45%	0.01%	1.89%	8.03%
8	4.75%	83.19%	0.66%	4.73%	0.13%	0.44%	0.00%	2.46%	3.63%
9	4.90%	83.01%	0.62%	7.01%	0.17%	0.42%	0.02%	2.93%	0.92%
10	4.94%	81.00%	0.57%	9.14%	0.19%	0.40%	0.05%	3.27%	0.44%
11	4.87%	77.33%	0.52%	10.94%	0.20%	0.37%	0.09%	3.47%	2.21%
12	4.72%	72.46%	0.47%	12.34%	0.21%	0.34%	0.13%	3.54%	5.81%
13	4.51%	66.96%	0.41%	13.33%	0.20%	0.31%	0.16%	3.52%	10.62%
14	4.26%	61.30%	0.36%	13.97%	0.20%	0.27%	0.18%	3.42%	16.04%
15	4.01%	55.86%	0.32%	14.33%	0.19%	0.25%	0.20%	3.28%	21.57%
16	3.77%	50.82%	0.28%	14.50%	0.18%	0.22%	0.21%	3.13%	26.88%
17	3.55%	46.31%	0.25%	14.55%	0.17%	0.20%	0.23%	2.97%	31.78%
18	3.35%	42.34%	0.23%	14.54%	0.16%	0.18%	0.23%	2.81%	36.16%
19	3.17%	38.91%	0.20%	14.49%	0.15%	0.16%	0.24%	2.67%	40.01%
20	3.03%	35.94%	0.19%	14.44%	0.15%	0.15%	0.24%	2.54%	43.33%

Table 13: Usual way (IRs way) to decompose variances of aggregate output

Periods									
1	15.21%	1.56%	0.02%	60.00%	0.73%	0.01%	3.19%	11.96%	7.33%
2	14.90%	0.10%	0.00%	63.63%	0.44%	0.00%	1.93%	14.05%	4.95%
3	10.77%	0.54%	0.02%	49.40%	0.25%	0.01%	1.06%	10.87%	27.09%
4	7.32%	1.38%	0.04%	35.82%	0.16%	0.02%	0.66%	7.42%	47.18%
5	5.15%	2.03%	0.04%	26.70%	0.12%	0.02%	0.46%	5.12%	60.36%
6	3.81%	2.47%	0.05%	20.86%	0.09%	0.03%	0.34%	3.69%	68.66%
7	2.96%	2.74%	0.05%	17.01%	0.08%	0.03%	0.28%	2.79%	74.08%
8	2.39%	2.93%	0.04%	14.39%	0.07%	0.03%	0.23%	2.19%	77.74%
9	2.00%	3.05%	0.04%	12.53%	0.06%	0.03%	0.20%	1.78%	80.32%
10	1.72%	3.13%	0.04%	11.18%	0.05%	0.03%	0.18%	1.49%	82.20%
11	1.51%	3.18%	0.04%	10.16%	0.05%	0.03%	0.16%	1.27%	83.60%
12	1.35%	3.22%	0.04%	9.40%	0.04%	0.02%	0.15%	1.11%	84.67%
13	1.23%	3.25%	0.03%	8.81%	0.04%	0.02%	0.14%	0.99%	85.49%
14	1.14%	3.28%	0.03%	8.36%	0.04%	0.02%	0.13%	0.89%	86.11%
15	1.07%	3.30%	0.03%	8.01%	0.04%	0.02%	0.13%	0.82%	86.59%
16	1.01%	3.33%	0.03%	7.75%	0.04%	0.02%	0.12%	0.76%	86.95%
17	0.97%	3.36%	0.03%	7.56%	0.04%	0.02%	0.12%	0.71%	87.21%
18	0.93%	3.39%	0.03%	7.43%	0.04%	0.02%	0.12%	0.67%	87.38%
19	0.91%	3.43%	0.03%	7.35%	0.03%	0.02%	0.12%	0.64%	87.48%
20	0.89%	3.47%	0.03%	7.31%	0.03%	0.02%	0.12%	0.61%	87.51%

Table 14: Usual way (IRs way) to decompose variances of aggregate capital

Periods									
1	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
2	1.01%	89.62%	0.35%	6.53%	0.24%	0.17%	0.39%	1.12%	0.57%
3	1.79%	81.87%	0.54%	11.75%	0.36%	0.28%	0.57%	2.12%	0.72%
4	2.21%	75.13%	0.58%	14.73%	0.38%	0.31%	0.61%	2.72%	3.33%
5	2.39%	68.04%	0.56%	16.19%	0.36%	0.31%	0.59%	2.98%	8.58%
6	2.42%	60.74%	0.50%	16.62%	0.33%	0.29%	0.56%	3.02%	15.53%
7	2.36%	53.70%	0.44%	16.41%	0.29%	0.26%	0.51%	2.91%	23.12%
8	2.25%	47.27%	0.37%	15.85%	0.26%	0.23%	0.47%	2.74%	30.56%
9	2.12%	41.62%	0.32%	15.13%	0.23%	0.21%	0.43%	2.53%	37.41%
10	2.00%	36.78%	0.27%	14.38%	0.20%	0.18%	0.39%	2.33%	43.47%
11	1.88%	32.69%	0.23%	13.66%	0.18%	0.16%	0.36%	2.15%	48.71%
12	1.77%	29.24%	0.20%	13.00%	0.16%	0.14%	0.33%	1.98%	53.18%
13	1.67%	26.35%	0.18%	12.41%	0.15%	0.13%	0.31%	1.83%	56.98%
14	1.59%	23.93%	0.15%	11.90%	0.13%	0.11%	0.29%	1.70%	60.19%
15	1.52%	21.89%	0.14%	11.47%	0.12%	0.10%	0.28%	1.59%	62.90%
16	1.46%	20.17%	0.12%	11.12%	0.11%	0.09%	0.26%	1.49%	65.17%
17	1.41%	18.72%	0.11%	10.82%	0.10%	0.09%	0.25%	1.41%	67.08%
18	1.37%	17.49%	0.10%	10.60%	0.10%	0.08%	0.25%	1.34%	68.68%
19	1.34%	16.46%	0.10%	10.42%	0.09%	0.07%	0.24%	1.28%	70.00%
20	1.32%	15.58%	0.09%	10.31%	0.09%	0.07%	0.24%	1.23%	71.09%

Table 15: Usual way (IRs way) to decompose variances of aggregate consumption

Periods									
1	0.55%	19.14%	0.19%	26.00%	9.00%	26.08%	2.63%	1.70%	14.72%
2	0.89%	27.82%	0.26%	9.75%	4.39%	13.02%	1.25%	0.36%	42.27%
3	1.10%	32.48%	0.27%	2.69%	2.02%	6.28%	0.53%	0.01%	54.62%
4	1.36%	38.52%	0.30%	0.39%	1.05%	3.42%	0.25%	0.06%	54.66%
5	1.74%	47.01%	0.33%	0.08%	0.58%	2.02%	0.11%	0.29%	47.83%
6	2.23%	58.00%	0.38%	1.06%	0.32%	1.23%	0.04%	0.67%	36.07%
7	2.76%	69.32%	0.43%	3.14%	0.16%	0.72%	0.01%	1.16%	22.31%
8	3.17%	76.59%	0.44%	5.90%	0.06%	0.37%	0.00%	1.64%	11.83%
9	3.21%	74.95%	0.41%	8.29%	0.01%	0.15%	0.01%	1.92%	11.04%
10	2.90%	65.29%	0.34%	9.47%	0.00%	0.05%	0.03%	1.92%	20.01%
11	2.43%	52.78%	0.26%	9.49%	0.00%	0.01%	0.04%	1.73%	33.26%
12	1.97%	41.41%	0.19%	8.89%	0.01%	0.00%	0.05%	1.48%	46.00%
13	1.60%	32.44%	0.15%	8.09%	0.02%	0.00%	0.06%	1.25%	56.41%
14	1.31%	25.72%	0.11%	7.31%	0.02%	0.01%	0.06%	1.05%	64.41%
15	1.09%	20.75%	0.09%	6.62%	0.02%	0.01%	0.06%	0.90%	70.46%
16	0.93%	17.05%	0.07%	6.04%	0.02%	0.01%	0.05%	0.77%	75.05%
17	0.80%	14.26%	0.06%	5.55%	0.02%	0.01%	0.05%	0.67%	78.56%
18	0.71%	12.13%	0.05%	5.15%	0.02%	0.02%	0.05%	0.59%	81.29%
19	0.63%	10.47%	0.04%	4.82%	0.02%	0.02%	0.05%	0.53%	83.43%
20	0.57%	9.16%	0.04%	4.54%	0.02%	0.02%	0.05%	0.48%	85.14%

List of Figures

List of Tables