# Cardiff Economics Working Papers



Working Paper No. E2018/17


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


*David Meenagh, Patrick Minford and Xiaoliang Ynag*



June 2018

ISSN 1749-6010



Cardiff Business School
Cardiff University
Colum Drive
Cardiff CF10 3EU
United Kingdom
t: +44 (0)29 2087 4000
f: +44 (0)29 2087 4419
business.cardiff.ac.uk

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

David Meenagh*
Cardiff University, UK

Patrick Minford†
Cardiff University, UK and CEPR

Xiaoliang Yang‡
Cardiff University

June 2018

**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 undertaking entrepreneurship. The model assumes wealthy people are more able to afford the costs of entrepreneurship. Wealth concentration therefore stimulates entrepreneurship among the rich and so aggregate growth, whose fruits in turn are largely captured by the rich. This process creates a mechanism by which inequality and growth are correlated. The model is estimated and tested by indirect inference and is not rejected. Policy-makers face a trade-off between redistribution and growth.

Keywords: Heterogeneous-agent Model, Entrepreneurship, Aggregate Growth, Wealth Inequality, Redistribution, Indirect Inferencee

JEL classification: E10; C63; O30; O40

## 1 Introduction

This paper investigates the relationship between capital inequality and aggregate economic growth in the UK during recent decades. This has become a central concern of policy-makers, highlighted by Piketty (2014); this reviewed a large amount of evidence for the fluctuations in inequality and growth over the industrial era in a wide range of countries. Many have concluded from this evidence that there must be some link between inequality and growth; however, a convincing link both in theory and in empirical studies has proved elusive. Most existing studies solely apply either simulation experiments which generate inequality by idiosyncratic randomness in order to fit actual distributional characteristics (e.g. variances of individual output and consumption) like Aiyagari (1994), Krusell and Smith (1998, 2006) and Algan et al. (2008) or reduced form regressions like Alesina and Rodrik (1994), Barro (2000) and Ostry et al. (2014). This paper, however, establishes a heterogeneous-agent growth model to fit the distributional characteristics of the UK economy. We introduce heterogeneity by classifying the population into two groups for simplicity, the rich who own higher capital holdings and the poor. Following Minford et al. (2007)'s endogenous growth mechanism, individuals have entrepreneurship incentives which drive individual productivity growth and further aggregate growth. Our development is to relate individual entrepreneurship incentives to the wealth distribution so that the rich have larger entrepreneurship incentives than the poor. History affords many examples and it is easy to enumerate more successful entrepreneurs born in rich families or the middle class, such as John Pierpont Morgan, Rupert Murdoch, Warren Edward Buffett, William Henry Gates III, Steven Paul Jobs and Mark Zuckerberg etc. This mechanism in our model, the tendency of wealth inequality to create entrepreneurship among the rich, raises average productivity growth, compared with the world of complete equality where all lose entrepreneurship incentives.

However, this same mechanism creates wealth inequality since the productivity growth originates mainly with the rich, who in turn reap larger rewards. The mechanism causes wealth to be gradually concentrated on the rich while also gradually raising the growth rate. Nevertheless this process can be interrupted and even temporarily reversed by aggregate shocks, such as crises and wars, and also by

*MeenaghD@cardiff.ac.uk; Cardiff Business School, Cardiff University, Colum Drive, Cardiff, CF10 3EU, UK
†MinfordP@cardiff.ac.uk; Cardiff Business School, Cardiff University, Colum Drive, Cardiff, CF10 3EU, UK.
‡YangX22@cardiff.ac.uk; Cardiff Business School, Cardiff University, Colum Drive, Cardiff, CF10 3EU, UK.

idiosyncratic shocks to income groups; furthermore it can be, and often is, modified by redistributive policies. We suggest that this is in effect the missing mechanism that can potentially explain much of the data assembled by Piketty and others. In this paper we apply it solely to UK data from the late 1970s in an initial formal test of the model.

To examine whether the model can artificially generate a stable co-movement of wealth inequality and economic growth due to these mechanisms, we conduct an experiment starting with two ad hoc identical groups where posterior heterogeneity solely comes from randomly idiosyncratic shocks. Afterwards we assess whether our model can fit UK stylised facts using the Indirect Inference (II) model test according to the actual UK quarterly data after 1978Q1.

Our empirical study concludes three important findings. First of all, the structural model can generate a stable relationship between inequality and growth. Our benchmark model cannot be rejected by the II test and model simulations can fit the main properties of UK actual data. Second, capital inequality is found to have a stimulating effect on economic growth, especially in the long term. Lastly, although a political redistribution intervention can 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 the tax rate increases, growth reduction follows a gradually increasing trend and thus an appropriately low tax rate instead of a high rate is a priority for policy-makers. The comparison between tax regimes shows that an income transfer regime (from the rich to the poor) is preferred to others.

This paper has the following structure. The current introduction is followed by a literature review in the next section. Our model setting is illustrated in section III. Then section IV describes the Indirect Inference (II) method for model testing and estimation in detail. Section V introduces the data we use. Empirical results are described in section VI. The last section concludes.

## 2 Relevant Literature

Our research firstly relates to the literature on heterogeneous-agent model (HAM), of which Bewley (1980, 1983) are early examples, to consider a variety of individual behaviour 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 computational techniques and numerical algorithms have developed and the popularity of microeconomic data has grown in recent decades, HAM has been more widely used. The development of HAM and of relevant numerical algorithms has proceeded in 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). 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. The well-known Aiyagari (1994) paper uses a 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 the 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 lacking in spite of individual volatility by introducing both idiosyncratic and aggregate shocks. These authors develop a new method to solve models by searching for an equilibrium law of motion for the wealth distribution around which some new numerical algorithms are developed. The distribution is generally described by finite order moments for simplicity and individual decisions are assumed to be made based on the distribution of moments. However, there might be an infinite-dimensionality issue if high order moments are considered for individual optimal decisions (Algan et al, 2014). Hence, the moment order is finally diminished to one which results in the so called "approximate aggregation" suggested by Krusell and Smith that the mean of the wealth distribution and the aggregate shock can well determine 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 may lead to variant sampling errors. The third generation focuses on solving models with a continuum of individuals and moreover attempts to remove the dependence on aggregate laws of motion when solving for individual behaviour. 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, an aggregate law is not essential to the solution for individual behaviour, but instead they search for equilibrium cross-agent distributions in each period described by density functions, so that "approximation aggregation" is not an inevitable finding.

Although the three generation models are helpful in explaining some properties of business cycles, the long-run effects of capital and income distributions on economic growth which now occupies increasingly more attention are not focused on.[1] Conversely, many other scholars establish theoretical models in order to derive an analytical relation between inequality and growth, especially the inequality effect on growth rather than stimulate important individual behaviours and distributional properties as reality, with neither analytical nor 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, have been developed. Their regression methods include cross-country regression for which least squares estimation is frequently used and panel data regression where pooled least squares, 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. 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 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 the estimation method and the 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 somewhat inefficient. Hence, we investigate the relation by structural model based simulations within one country instead of across countries. Our methodology can easily be applied to other economies, but separate model testing in each country is necessary.

# 3 Model Setting

This section illustrates our model with a growth mechanism originating from Minford et al. (2007) by applying the entrepreneurship to existing output methods understood in the most general way. We introduce heterogeneous agents into this growth model and relate individuals' incentives to undertake entrepreneurship to their current wealth so that wealth distribution can endogenously affect economic growth. This idea comes from the fact that rich people who have adequate wealth to support surplus activities generally have larger incentives to undertake entrepreneurship 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 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 An outline of the model

At the heart of this model is the idea that individuals can profit from undertaking entrepreneurship and paying the costs of entrepreneurial entry as opposed to merely pursuing work for others, paid for on a regular basis and providing them with a regular income. In their labour choices they can compute an expected return from this entrepreneurship; at the margin they equate this expected return with the wage they would get from doing a regular job. Individuals' optimal behaviour is affected by their current wealth: someone who is poor will have a high marginal utility of current consumption from regular work; someone who is rich will have a low one. Hence the relative marginal utility provided by expected returns on entrepreneurship will be higher for the rich than the poor. We assume in this that the marginal utility

[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.

of expected future returns will be similar for both since if the entrepreneurship is successful both will expect to be rich and if not both will expect to be poor; rising or falling future returns are accompanied by respectively falling and rising marginal utility as income and consumption move together. The key difference is in the marginal utility of the consumption sacrificed to pay the costs of entrepreneurship against which the expected return to entrepreneurship is measured. For a poor person whose income is low the marginal utility of these costs is higher, whereas for a rich person with high income the marginal utility of these costs is low.

In a world of complete equality, wealth is shared equally and is therefore modest for all households. For all of them the marginal utility of entrepreneurship costs is high relative to expected future returns from entrepreneurship and little entrepreneurship is done. As wealth is redistributed to one set of households, the rich, from the others, the relative expected returns rise for the rich as the marginal utility entrepreneurship costs falls rapidly with the concavity of the utility function: meanwhile the marginal utility of these costs rises rapidly for the others for the same reason. As a result, the rich have a much larger incentive to undertake entrepreneurship, while the others lose their incentive. However, as the rich undertake more entrepreneurship, they augment their future income and wealth, which leads to further increases in the amount of entrepreneurship they undertake. This in turn induces further increases in their wealth so that the process tends upwards without limit, with both wealth and growth rising among the rich. As the poor reduce their entrepreneurship, it gradually falls off asymptotically towards zero and with it the growth in their wealth. So the total amount of entrepreneurship rises as this process continues, because of the asymmetry between the growing rise in entrepreneurship and growth among the rich on the one hand and the fall in both among the poor. Ultimately, as wealth is redistributed more and more the percentage rise in the richer group is applied to a large volume of entrepreneurship while the percentage fall in the poorer group is applied to a very small amount of entrepreneurship and is consequently trivial in absolute terms. This is the mechanism at work in the model, randomly triggered by individual shocks. Thus randomly some will be fortunate and acquire more wealth, others unfortunate and lose wealth. This will create more absolute entrepreneurship, with the fortunate carrying out most of it. Hence it will create rising growth. At the same time it will create rising inequality.

Finally, we have governmental intervention in this process via its tax/subsidy/regulation systems. We can think of this as a political process driven by the fact that the two groups, rich and poor, share the society and must live with each other in a sustainable way. We do not attempt to create a political economy model of this: we simply assume there is some exogenously given redistributive system in place and investigate its consequences for growth and inequality. Plainly there is a trade-off here for society and government: by intervening with redistribution the government will dilute the two mechanisms at work. There will be less inequality and less growth. We could think of the underlying political economy model as one in which the poor, who will be numerically most powerful, recognise this trade-off: they want growth but also want its fruits to be spread to them. At the one extreme they could be totally spread and then there would be no growth; at the other not spread at all and then there would be maximum growth but only enjoyed by the rich. This would be an interesting strand of future work to pursue in understanding the full dynamics of growth and inequality; however, in this paper we leave it to one side.

The set-up we have just set out is embedded in a standard and simplified DSGE model of two agents, who consume and save in the form of capital; there is no borrowing and lending, and the economy is closed. Plainly it is highly unrealistic to close down the financial markets that could provide capital and insurance to entrepreneurs without wealth. In the modern world much innovation, whose risks are readily computable, is handled in an orthodox way by firms using pooled equity and credit of many investors willing to take the risks. However, the basic idea of this paper is that the seriously disruptive ideas that change the business and scientific world tend not to be financially attractive and are pursued by those who have the vision and the means to do it. History affords many examples. Our model emphasises their role to the exclusion of standard R&D processes. However, we are not proposing that our model is accepted on any such a priori grounds. It is a deliberately simplified model but one which we think can explain the data behaviour effectively.

### 3.2 Individual behaviour

Assume the population in an economy is comprised of two groups with constant population weights $\mu_i; i = 1, 2$. Both groups consider the same utility as follows where $N_{i,t}$ and $Z_{i,t}$ are labour input and entrepreneurship time (incentive) respectively.

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

We assume both the capital market and labour market are perfect. Entrepreneurship has a unit cost $\pi_t$ and the total cost of entrepreneurship for individual $i$ is $\pi_t Z_{i,t}$. This cost is assumed to be the result of government taxes and regulations on entrepreneurial activity; these will yield revenue to the government (some of which is passed on to chosen vested interest groups). We will assume that this revenue is offset by a lump sum transfer payment to all households; in what follows we will assume that both the entrepreneurial cost and the lump sum transfer are indexed to general income, i.e. overall GDP. Agents have the following budget constraint where $b_{i,t}$ is individual bonds. Agents have the following budget constraint where $b_{i,t}$ is individual bonds and $T_t$ is a lump sum transfer to all households of the proceeds of the entrepreneurial tax.

$$(1-\tau)Y_{i,t} + (1+r_{t-1})b_{i,t} - \pi_t Z_{i,t} + T_t = C_{i,t} + b_{i,t+1} + K_{i,t} - (1-\delta)K_{i,t-1} \tag{2}$$

Individuals have a Cobb-Douglas production function (3) where the non-stationary individual productivity $A_{i,t}$ evolves as the process (4) which relies on individual entrepreneurship as well as the aggregate productivity shock $\upsilon_{A,t}$.[2]

$$Y_{i,t} = A_{i,t}(K_{i,t-1})^{\alpha}(N_{i,t})^{1-\alpha} \tag{3}$$

$$\frac{A_{i,t+1}}{A_{i,t}} = \theta_1 + \theta_2 Z_{i,t} + \upsilon_{A,t} \tag{4}$$

As the model is deliberately simple, one can easily obtain the following optimal decisions from first order conditions.

$$(C_{i,t})^{-\Psi_1} = (1+r_t)\beta E_t\left[(C_{i,t+1})^{-\Psi_1}\right] \tag{5}$$

$$(C_{i,t})^{-\Psi_1} = \beta\left\{E_t\left[(C_{i,t+1})^{-\Psi_1}\right]\left[\alpha(1-\tau)\frac{Y_{i,t+1}}{K_{i,t}} + 1 - \delta\right]\right\} \tag{6}$$

$$\frac{1-\Phi}{(1-N_{i,t}-Z_{i,t})^{\Psi_2}} = \Phi\,(C_{i,t})^{-\Psi_1}(1-\tau)(1-\alpha)\frac{Y_{i,t}}{N_{i,t}} \tag{7}$$

$$\frac{1-\Phi}{(1-N_{i,t}-Z_{i,t})^{\Psi_2}} + \Phi\frac{\pi_t}{(C_{i,t})^{\Psi_1}} = \Phi\theta_2\frac{A_{i,t}}{A_{i,t+1}}E_t\left[\sum_{s=1}^{\infty}\beta^s\frac{(1-\tau)Y_{i,t+s}}{(C_{i,t+1})^{\Psi_1}}\right] \tag{8}$$

Equation (8) is an accurately optimal decision rule for $Z_{i,t}$ and can be approximated to equation (9) by approximating $\frac{Y_{i,t}}{C_{i,t}}$ as a random walk before the steady state.[3]

$$(1-\tau)(1-\alpha)\frac{Y_{i,t}}{N_{i,t}} + \pi_t = (1-\tau)\frac{A_{i,t}}{A_{i,t+1}}Y_{i,t}\frac{\beta\theta_2}{1-\beta} \tag{9}$$

where $\Psi_1$ is set to unity for simplicity and this value is also used in our empirical study. Equation (9) gives an approximately optimal decision rule for $Z_{i,t}$. According to the perfect labour market assumption, $(1-\alpha)\frac{Y_{i,t}}{N_{i,t}}$ is the individual implied real wage rate $w_{i,t}$. As individual entrepreneurship has the unit cost $\pi_t$ as well as the unit opportunity cost $w_{i,t}$, we define an "entrepreneurship penalty rate" $\pi'_{i,t} = \frac{\pi_t}{w_{i,t}}$ to reflect the total cost and rewrite (9) as the following

$$\frac{A_{i,t+1}}{A_{i,t}} = \frac{(1-\tau)\beta\theta_2\frac{Y_{i,t}}{w_{i,t}}}{(1-\beta)(1-\tau+\pi'_{i,t})} \tag{10}$$

Notice how the costs of entrepreneurship relative to current productivity, $\pi'_{i,t}$, and so to current income, reduce productivity growth. As current income and consumption falls, the marginal utility of these costs rises, raising the disincentive to entrepreneurship.

[2] We apply an aggregate productivity shock to reflect the effects of aggregate production on individual ones.

[3] This approximation originates from Minford (2007) and the details are shown in Appendix A.

We can easily linearise (10) as (11) by relegating $\frac{Y_{i,t}}{w_{i,t}}$ into the error term.[4]

$$\ln A_{i,t+1} - \ln A_{i,t} = \phi_{1,i} - \phi_{2,i}\pi'_{i,t} + \varepsilon_{A,t} \tag{11}$$

## 3.3 Entrepreneurship penalty rate

Via (10) the individual penalty rate falls with rising income. Income rises directly with capital input. We capture this effect by assuming it to be negatively related with the lagged individual-aggregate capital per capita ratio $\frac{K_{i,t-2}}{K_{t-2}}$ which is a measure of relative wealth and income. This assumes, as noted earlier, that entrepreneurial costs are indexed to the general rise in income (and so wealth). Lastly we assume individuals can observe $\pi'_{i,t}$ but do not know how exactly it is set and thus $\pi'_{i,t}$ is predetermined for optimising individuals.

$\pi'_{i,t}$ evolves as follows.

$$\pi'_{i,t} = \rho_0^\pi + \rho_1^\pi \pi'_{i,t-1} - \rho_2^\pi \cdot Q\left(\frac{K_{i,t-2}}{K_{t-2}}\right) + \varepsilon_{\pi',t} \tag{12}$$

where $\rho_1^\pi \geq 0; \rho_0^\pi, \rho_2^\pi > 0$. Both Banerjee and Duflo (2003) and Kolev and Niehues (2016) find that inequality has empirically significantly nonlinear effects on growth where they consider the effect of inequality square. We hence set $Q(\frac{K_{i,t}}{K_t}) = \frac{\mu_i}{\omega_{Y,i}}(\frac{K_{i,t}}{K_t})^2$ where $\frac{K_{i,t}}{K_t} = 1$ represents a perfect equality while either greater or less than 1 implies inequality.[5] In Appendix B we show that $Q(\frac{K_{i,t}}{K_t})$ has a minimum at perfect equality where $\frac{K_{i,t}}{K_t} = \frac{\mu_i}{\omega_{Y,i}} = 1$, and penalty rates are identical across agents. The expected aggregate growth next period has a reduced form depending on the current aggregate output, current aggregate capital and both the current and the lagged capital inequality. Theoretically, the long-run growth rate has its minimum with complete equality. Short run behaviour, however, is more complicated, as labour input might be temporarily reduced by an increase in entrepreneurship incentives which will have a currently negative effect on output while the rising incentives have a stronger positive in future.

## 3.4 Aggregate economy

Each aggregate variable is the weighted sum of individual ones plus the relevant aggregate uncertainty as follows.

$$Y_t = \mu_1 Y_{1,t} + \mu_2 Y_{2,t} + \mu_3 Y_{3,t} + \epsilon_{Y,t} \tag{13}$$

$$K_t = \mu_1 K_{1,t} + \mu_2 K_{2,t} + \mu_3 K_{3,t} + \epsilon_{K,t} \tag{14}$$

$$C_t = \mu_1 C_{1,t} + \mu_2 C_{2,t} + \mu_3 C_{3,t} + \epsilon_{C,t} \tag{15}$$

Market clearing in goods can be written as

$$Y_t = K_t + C_t - (1-\delta)K_{t-1} + G_t + NX_t \tag{16}$$

where we let $NX_t$ = error and $G_t = \tau Y_t$. The government returns the entrepreneur tax cost as a lump sum transfer to all households. It thus pursues a balanced primary budget and for simplicity we assume has no debt. The bond market clears via Walras' Law. We can therefore rewrite (16) as

$$(1-\tau)Y_t = K_t + C_t - (1-\delta)K_{t-1} + \epsilon_{m,t} \tag{17}$$

The list of linearised model equations is shown in Appendix C.

[4] $\phi_{2,i} = \frac{(1-\tau)\beta\theta_2 \frac{Y_i}{w_i}}{(1-\beta)(1-\tau+\pi'_i)^2}$. The error term in (11) is an aggregate instead of individual one because firstly $\frac{Y_{i,t}}{w_{i,t}}$ is close across agents and secondly individual productivities can be related to the aggregate productivity.

[5] $\frac{\mu_i}{\omega_{Y,i}}$ firstly aims to avoid the penalty policy too beneficial to the rich as the poor generally has a greater population weight relative to their average income share. Second, this setting simplifies aggregation.

### 3.5 Indirect Inference

In this section, we illustrate our methodology of model testing and parameter estimating, Indirect Inference (II), developed by Le et al. (2011).[6] II is based on the idea that if the structural model is true in terms of both specification and parameters, the properties of the actual data should come from the distribution of the properties of the simulated data with some critical minimum probability. The data properties can be captured by a simple auxiliary model such as a VAR, Impulse response functions or moments. Define the parameters of the structural model and the auxiliary model as $\theta$ and $\beta$ respectively. We firstly use the actual data to estimate the auxiliary parameters, say $\widehat{\beta}$. Given the null hypothesis $H_0 : \theta = \theta_0$, we stimulate S samples using the structural model and estimate the auxiliary parameters using each simulated sample to obtain estimators $\widetilde{\beta}_s(\theta_0); s = 1, \cdots, S$. To evaluate whether $\widehat{\beta}$ comes from the distribution of $\left\{\widetilde{\beta}_s(\theta_0)\right\}$, we compute the Wald statistic

$$Wald_a = \left[\widehat{\beta} - \overline{\widetilde{\beta}_s(\theta_0)}\right]' \sum(\theta_0)^{-1} \left[\widehat{\beta} - \overline{\widetilde{\beta}_s(\theta_0)}\right]$$

which asymptotically follows a $\chi^2(k)$ distribution where $k$ is the number of elements in $\beta$ and $\sum(\theta_0)$ is the variance-covariance matrix of $\widehat{\beta} - \widetilde{\beta}_s(\theta_0)$.[7] We can check the allocation of $Wald_a$ in the distribution of simulated $Wald_s; s = 1, \cdots, S$ where $Wald_s$ is computed when using the $s^{th}$ simulated sample to estimate $\widehat{\beta}$. If $Wald_a$ is less than the $c^{th}$ percentile value of $\{Wald_s\}$ sorted from smallest to largest, $H_0$ cannot be rejected in a $c\%$ confidence interval; otherwise the model is false.[8] An alternative way is to compute the transformed Mahalanobis Distance (TMD) and compare it with the critical value of $t$ distribution on the $c\%$ confidence interval.[9]

$$Z = T_c \left[\frac{\sqrt{2Wald_a} - \sqrt{2k-1}}{\sqrt{2Wald_c} - \sqrt{2k-1}}\right]$$

where $T_c$ is the critical value of a one-tail $t$ distribution on the $c\%$ confidence interval.

Generally, a (linearised) DSGE model can be represented as a VARMA or a VAR($\infty$) which can be further represented to a VAR($p$) with a finite order or even a VAR(1) (Dave and De Jong 2007, 9-9; Giacomini 2013; Wickens 2014). However, the long-run solution of our model can only be approximated as a VARX with non-stationary lagged endogenous variables X due to nonstationarity productivities.[10] Both Le et al. (2011) and Le et al. (2015) conduct experiments on examining the power of testing when different orders and VAR coefficients are considered. They find that unity order combined with an appropriate number of variables, e.g. 3 endogenous variables, has a sufficiently great power. Hence, we use a VARX(1) with 3 variables, aggregate output, aggregate capital and capital inequality, combined with the lagged individual productivities as the "X".[11] The auxiliary parameter vector $\beta$ contains 9 VAR coefficients and 3 variances of the VAR residuals.

Given the null hypothesis that the structural model is true, one could back out the structural errors from the model and the actual data and then bootstrap these structural errors to obtain simulated samples. II is also used to estimate the parameters by searching for the parameter values such that the relevant Wald or TMD is smallest.

Le et al. (2012) and Le et al. (2015) conduct Monte Carlo power tests on three testing methods: II, Likelihood ratio test; and "unrestricted Wald" test on different models. II is found to be more powerful than the other classical testing methods. To evaluate the power of II 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 treating each simulated sample from the true model as the observation, we test the false model by II and calculate the rejection rate out of the 500 Monte Carlo experiments. Table 1 shows the result of our power test against the false models with mis-estimation where both structural

[6] II is different from indirect estimating. See details of the latter in Smith (1993), Gourieroux et al. (1993) and Davidson and MacKinnon (2004, 556-44).

[7] Because $\sqrt{S}\left[\widehat{\beta} - \widetilde{\beta}_s(\theta_0)\right]$asymptotically follows a normal distribution, proved by Gourieroux et al. (1993).

[8] We do not directly compare Wald_a with a certain critical value of $\chi^2(k)$because of the possible simulated sample-variant issue.

[9] $\sqrt{2Wald_a}$asymptotically follows a normal distribution $N(\sqrt{2k-1}, 1)$.

[10] See more details of the VARX approximation in Appendix D.

[11] Appendix B has shown that the aggregate growth is determined by aggregate output, aggregate capital and capital inequality while the aggregate growth is the difference in logarithm of aggregate output and thus in the VARX(1), series of aggregate output could implicitly describe growth.

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.

| Parameter Falseness | True | 1% | 3% | 5% | 7% | 9% |
|---|---|---|---|---|---|---|
| Rejection Rate with 0.95 Confidence | 0.05 | 0.134 | 0.27 | 0.672 | 0.82 | 1.00 |

Table 1: Power test against numerical falsity of parameters

We then test the power of II against a mis-specified model in which the basic mechanism of wealth inequality on entrepreneurship is turned off. Namely, the equations of penalty rate (C15)-(C16) in Appendix C are replaced by a simple AR(1) process. Wealth inequality in this false model is generated by randomness. We keep parameters the same as the full-estimated values in the benchmark model. As the rich are still more sensitive to a reduction in the penalty rate ($\phi_{2,1} > \phi_{2,2}$), 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 0.95 confidence is as high as 0.994 and hence we can conclude that II provides huge power against the mis-specified models which attempt to mimic the results from the true model.

# 4 Model Data

We collect seasonally adjusted quarterly data in the UK from 1978Q1 to 2015Q4 with no stationarisation.[12] Aggregate output and consumption are GDP and household final consumption expenditure respectively, measured by chain volume with base year 2012 from UK national statistics. The real interest rate is one quarter of the difference between annual nominal interest and annual next-period inflation rate where nominal interest rate is the 90-day Yields Rate reported by the Bank of England. Aggregate labour is the number employed in the UK aged 16 and over divided by the sum of the total claimant counts and the number of UK workforce jobs and then divided by two, sourced by ONS. Aggregate capital stock is estimated by the perpetual inventory method. We firstly collect the data on UK investment from ONS to calculate the average output-investment ($Y$-$I$) ratio. Given the annual capital-income ratio in the UK, 4.715, we can get the average quarterly capital-investment ($K$-$I$) ratio. Then the capital depreciation rate $\delta$ can be backed out from the equilibrium investment equation using the estimated $K$-$I$ ratio. We lastly use the investment equation $K_{t+1} = I_t + (1 - \delta)K_t$ to iterate capital.

To estimate the data on the entrepreneurship penalty rate $\pi'_{i,t}$, we follow Minford (2016)'s idea that the costs of incentives are determined by two factors, the "labour market regulation" (LMR) and the tax rate. LMR describes the degree of supervision in the labour market, protection to labour welfare and employment costs and we use the average of two indicators "centralized collective bargaining" (CCB) and "mandated cost of worker dismissal" (MCD) reported by Fraser Institute as the proxy of LMR.[13] For the other determinant, different from Minford (2016), we choose corporation tax rate (CTR). Finally, the quarterly $\pi'_{i,t}$ can be achieved by averaging the transformed LMR (TLMR) and the CTR as shown in Figure 1.[14]

With regard to the individual data, due to the lack of the data tracking the same agents for decades, we assume no cross-group mobility for individuals. 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 which are classified by the income deciles, top 10% and bottom 90%. Data on income distributions and wealth distributions are collected from "World Wealth & Income Database" (WID). The distribution of "taxes on final goods and services" reported by UK ONS is used as the proxy of consumption distribution. The data on individuals' labour are estimated by the aggregate labour and the labour ratio of each group. According to the implicit wage equal to the marginal productivity of labour, the group labour ratio can be approximated by the group income ratio over the wage ratio.[15] Individual $\pi'_{i,t}$ can be estimated from the aggregate $\pi'_{i,t}$ and

[12] As VARX is used as the auxiliary model, the non-stationary variables could be directly considered. Besides, stationarisation could affect the shocks. For example, HP filter will affect volatilities of shocks and introduce extra cyclical properties.

[13] CCB describes the procedure for both employers and employees to make a collective agreement on contracts (Gernigon et al, 2000). It is measured by a value from 0 (hardest to agreements) to 10 (easiest), so is MCD.

[14] See more details of the data on LMR and CRT in Appendix E.

[15] We find the 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.

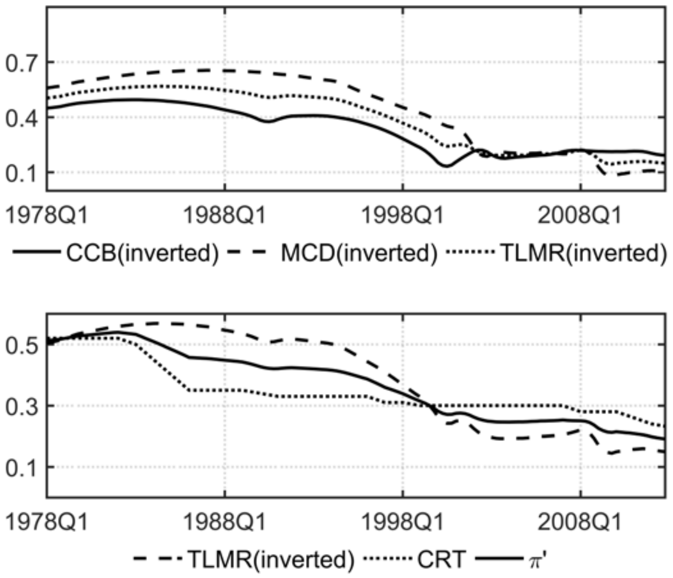


Figure 1: Generated data on penalty rate

the individual wage distribution.[16] Lastly, individual productivities are the individual Solow residuals. All the data are plotted in Appendix E. We measure inequality by the share of the top 10% decile as in Figure 2 instead of the Gini-coefficient because firstly the number of groups is not sufficiently large to compute Gini-coefficient and also Gini-coefficient cannot tell where the inequality comes from.

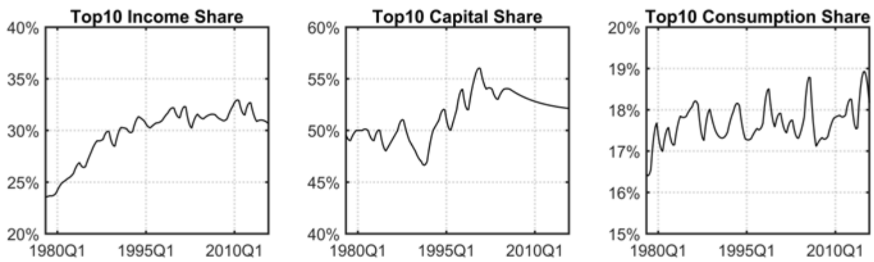


Figure 2: Inequality indicators

# 5 Empirical Results

## 5.1 Detection on tendency

This section aims to examine whether the two interactive mechanisms of our model work properly. We start from two identical groups with the same population weights. Inequality is generated solely by innovations from normal distributions.[17] As we expect the tendency is parameter-independent if it exists, parameter values are calibrated and freely set.[18] To ensure the time period is sufficiently long to observe tendency, data are extended to 250 periods.

Figure 3 shows one typical simulation where the capital inequality is calculated by capital per capita in group one (the rich) over the aggregate capital per capita minus one ("0" represents a perfect capital equality while more deviation from 0 implies higher inequality). Aggregate growth rate deviation is

[16] We estimate $\pi'_{i,t}$by $\pi'_{i,t} = \frac{w_t}{w_{i,t}}\pi'_t = \frac{\mu_i}{\omega_{w,i}}\pi'_t$where $\omega_{w,i} = \frac{\mu_i \cdot w_{i,t}}{w_t}$is the wage share of group $i$.

[17] We do not bootstrap innovations backed out from the model and the actual data because the backed-out innovations are same across identical agents. For convenient computation, we set the distribution mean of innovations slightly different across groups to make one group, say group one, more likely to be richer.

[18] The key parameters freely set are $\theta_2 = 0.5; \rho_2^{\pi} = 0.001; \phi_{2,i} = 0.5$.

measured by how the simulated data deviate from the actual data which exhibits lower volatilities than quarterly simulated growth rate, considering the non-negligible effects of actual data on simulated growth. Obviously there exhibits a stable relationship between capital inequality and aggregate growth and both have almost the same tendency. The synchronously gradual increase in both inequality and growth reflects the two interactive mechanisms of our model that inequality can stimulate growth while growth can further lead to a higher inequality degree.

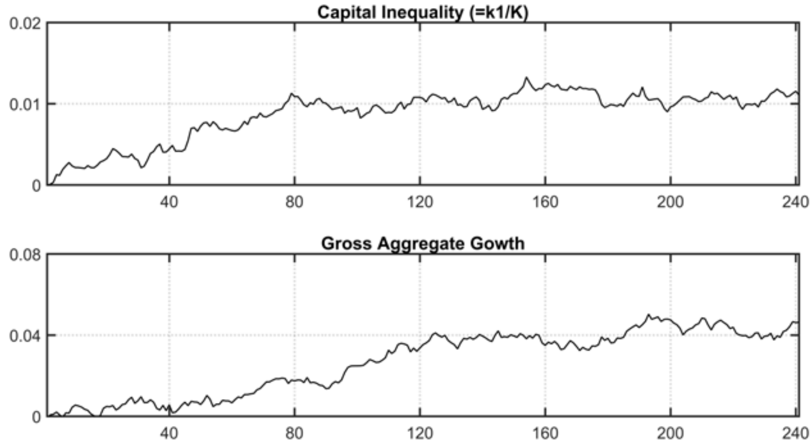


Figure 3: Tendency of aggregate growth and capital inequality

## 5.2 Empirical results on actual data

In this section, we employ the UK quarterly data on the 10%t-90% income segmentation to examine if the benchmark model can mimic important characteristics of UK economic growth and inequality in recent decades. Redistribution policies will also be considered.

Table 2 shows our parameter calibration according to the sampled UK quarterly data. The elasticity of consumption in utility $\Psi_1$ is set to 1 following the proof in Appendix A. Elasticity of labour $\Psi_2$ is also set to 1 for simplicity. The constant term $\rho_0^\pi$ in the individual equation of $\pi'_{i,t}$ is set to the sample average of the aggregate penalty rate.

| | | |
|---|---|---|
| Share of capital in Cobb-Douglas production | $\alpha$ | 0.300 |
| Utility discount rate | $\beta$ | 0.970 |
| Capital depreciation rate | $\delta$ | 0.015 |
| Share of conumption preference in CRRA utility[19] | $\Phi$ | 0.500 |
| Elasticity of consumption in CRRA utility | $\Psi_1$ | 1.000 |
| Elasticity of labourr in CRRA utility | $\Psi_2$ | 1.000 |
| Steady-state net growth rate of productivity | $\theta_1$ | 0.004 |
| drift in linearised productivity equation for group $i$ | $\phi_{1,i}$ | 0.004 |
| drift in individual entrepreneurship penalty equations | $\rho_0^\pi$ | 0.369 |
| Steady-state output share by top 10% income decile | $\omega_{Y,1}$ | 0.300 |
| Steady-state capital share by top 10% income decile | $\omega_{K,1}$ | 0.500 |
| Steady-state consumption share by top 10% income decile | $\omega_{C,1}$ | 0.300 |
| Steady-state aggregate output/consumption ratio | $Y/C$ | 1.693 |
| Steady-state aggregate capital/consumption ratio | $K/C$ | 25.66 |

Table 2: Calibrated parameter values

We emphasise estimation on $\rho_2^\pi$ and $\phi_{2,i}$ where the first tells us how capital distribution affects policy attitude on entrepreneurship penalty rate and the second tells us how this intermediate factor drives productivity growth. The estimated $\phi_{2,1}$ is greater than $\phi_{2,2}$ in Table 3 implies that individual productivity is more sensitive to changes in penalty rate for the rich.

Given these parameters, AR coefficients of the structural errors are summarised below.[20]

[20] Aggregate productivity is tested by both ADF test and Phillips-Perron test and is found to be I(1).

| | | |
|---|---|---|
| Marginal effect of entrepreneurship time on individual productivity growth | $\theta_2$ | 0.4970 |
| (Negative) Marginal effect of capital on individual entrepreneurship penalty rate | $\rho_2^{\pi}$ | 0.0023 |
| (Negative) Marginal effect of penalty rate on productivity growth for the rich | $\phi_{2,1}$ | 0.5670 |
| (Negative) Marginal effect of penalty rate on productivity growth for the poor | $\phi_{2,2}$ | 0.4930 |

Table 3: Indirect Inference estimators

| $\frac{d\varepsilon_{Y,t}}{d\varepsilon_{Y,t-1}}$ | $\frac{d\varepsilon_{K,t}}{d\varepsilon_{K,t-1}}$ | $\frac{d\varepsilon_{C,t}}{d\varepsilon_{C,t-1}}$ | $\frac{d\Delta\varepsilon_{A,t}}{d\Delta\varepsilon_{A,t-1}}$ | $\frac{d\varepsilon_{C1,t}}{d\varepsilon_{C1,t-1}}$ | $\frac{d\varepsilon_{C2,t}}{d\varepsilon_{C2,t-1}}$ | $\frac{d\varepsilon_{N1,t}}{d\varepsilon_{N1,t-1}}$ | $\frac{d\varepsilon_{N2,t}}{d\varepsilon_{N2,t-1}}$ | $\frac{d\varepsilon_{\pi'_t}}{d\varepsilon_{\pi'_{t-1}}}$ |
|---|---|---|---|---|---|---|---|---|
| 0.9651 | 0.9453 | 0.8953 | 0.0026 | 0.8892 | 0.8773 | 0.9397 | 0.9424 | 0.9705 |

Table 4: AR coefficients of structural errors

Our II testing result below shows that the null hypothesis that structural parameters are equal to our estimators cannot be rejected on the 5% confidence interval.

| Wald Statistic | Transformed MD | Wald Percentile |
|---|---|---|
| 25.205 | 1.6358 | 0.9481 |

Table 5: II wald test results

As illustrated in the II section, our auxiliary model VARX takes the form of

$$\begin{bmatrix} Y_t \\ K_t \\ IQ_t \end{bmatrix} = \beta \begin{bmatrix} Y_{t-1} \\ K_{t-1} \\ IQ_{t-1} \end{bmatrix} + \alpha X_t + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \varepsilon_{3,t} \end{bmatrix}$$

where $X_t$ is a vector of the exogenous non-stationary variables. Table 6 shows the coefficients of the auxiliary model with the actual data compared to those with the simulated data.[21]

| Auxiliary Coefficients | Actual | | Simulated | |
|---|---|---|---|---|
| | | Mean | Lower 2.5% | Upper 2.5% |
| $\beta_1$ | 0.7745 | 0.8414 | 0.6747 | 1.0005 |
| $\beta_2$ | 0.0475 | 0.0205 | −0.1065 | 0.1668 |
| $\beta_3$ | 0.0875 | 0.0519 | −0.1753 | 0.2706 |
| $\beta_4$ | 0.0251 | 0.0567 | 0.0091 | 0.1169 |
| $\beta_5$ | 0.9656 | 0.9585 | 0.9165 | 1.0069 |
| $\beta_6$ | −0.0060 | 0.0068 | −0.0575 | 0.0796 |
| $\beta_7$ | −0.0495 | 0.0063 | −0.0666 | 0.0943 |
| $\beta_8$ | 0.0548 | 0.0048 | −0.0897 | 0.0757 |
| $\beta_9$ | 0.9773 | 0.9505 | 0.8590 | 1.0164 |
| $Var(\widehat{\varepsilon}_1)$ | 0.0000300 | 0.0001441 | 0.0000995 | 0.0001979 |
| $Var(\widehat{\varepsilon}_2)$ | 0.0000003 | 0.0000021 | 0.0000009 | 0.0000048 |
| $Var(\widehat{\varepsilon}_3)$ | 0.0000132 | 0.0000358 | 0.0000213 | 0.0000567 |

Table 6: Coefficients of the auxiliary model

Now we analyse how different shocks affect both the aggregate economy and individual behaviours. As the aggregate shocks have the typical effects as the basic RBC models, here we focus on the impulse response functions (IRF) of the key variables to the individual shocks and the shock of the penalty rate (See the rest of the IRF figures in Appendix F).[22] A sudden rise in consumption preference of the rich in Figure 4 firstly reduces individual capital accumulation of the rich, followed by a decrease in both individual output of the rich and the capital inequality. This negative effect is also spread to the poor due to the rising interest rate so that the capital inequality gradually rises back. After all, the aggregate growth rate has a humped shape.[23] A temporary rise in the labour supply of the rich in

[21] The auxiliary coefficient vector to compute the Wald statistic contains the elements in $\beta \begin{bmatrix} \beta_1 & \beta_2 & \beta_3 \\ \beta_4 & \beta_5 & \beta_6 \\ \beta_7 & \beta_8 & \beta_9 \end{bmatrix}$ and the variances of the VARX regression residuals $Var(\widehat{\varepsilon}_i); i = 1, 2, 3$.

[22] Some IRFs are not hump-shaped within the forecasting periods mainly because individual capital is too sensitive to the real interest rate. One can try the "capital adjustment cost" originating from Hayashi (1982) in individual's budget constraint to improve them.

[23] The aggregate growth in the initial forecasting period is negative; however, the simulation starting from that period can only give the growth rate starting from the next period.

Figure 5 has contrary effects on both growth and inequality. As we see, any individual shock can have a cross-group effect via the capital market (affecting the interest rate) and the goods market in our model. The effects of the penalty rate are of most interest to us because $\pi'_{i,t}$, like a tax rate in some sense, can be easily controlled by governments. If $\pi'_{i,t}$ falls due to a negative shock like in Figure 6, both groups will undertake more entrepreneurship. However, the distributions of capital will gradually become more unequal because individual productivity is more sensitive to changes in the penalty rate for the rich ($\phi_{2,1} > \phi_{2,2}$) so that the rich have a higher individual productivity growth. Meanwhile, the aggregate growth will also rise.

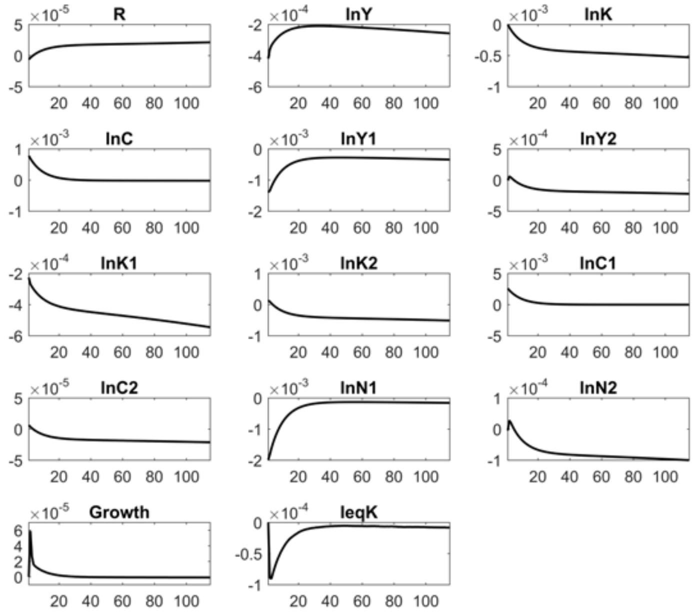


Figure 4: IRF to a one s.d. individual consumption temporary shock on the rich

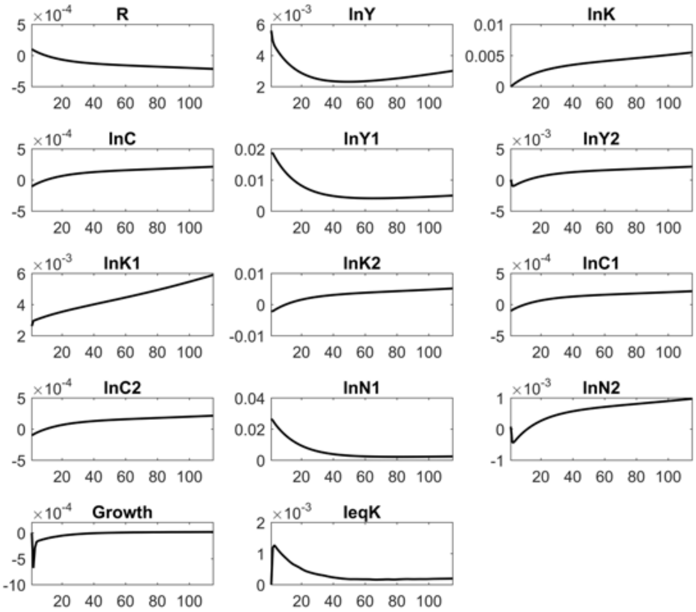


Figure 5: IRF to a one s.d. individual labour supply temporary shock on the rich

We conduct a variance decomposition analysis to investigate how various innovations affect volatilities of growth and capital inequality across time.[24] Table 7 indicates that the volatilities of growth and

[24]See the details of the decomposition analysis in Appendix G.

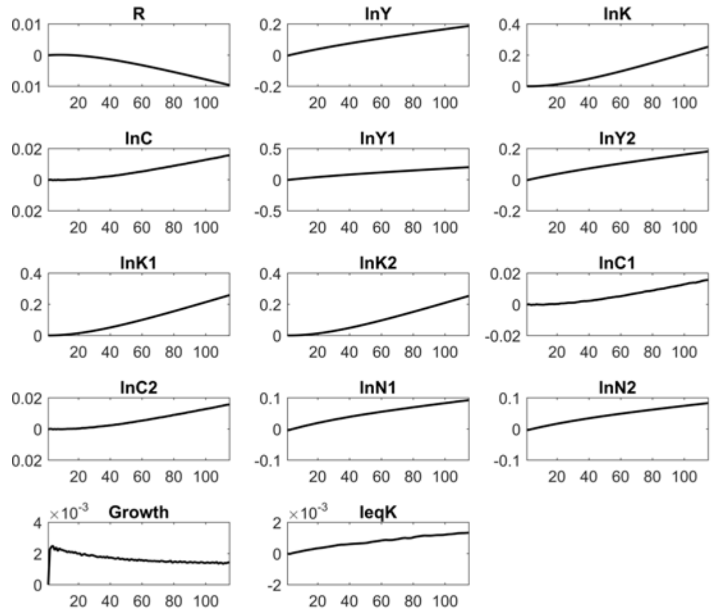


Figure 6: IRF to a negative one s.d. temporary penalty shock

inequality across time are both dominated by innovations of the penalty rate and not all the innovations have significant on the inequality fluctuations. Compared with others, innovations of aggregate capital have moderate effects on both the growth fluctuations and the inequality fluctuations which suggests that government can adopt capital intervention to stabilise the capital inequality and that this is more feasible than adjustment of the penalty rate.

## 5.3 Redistribution and policy measurement

As the rising wealth concentration is unlikely to spontaneously die out due to the interactive relation between inequality and growth, a redistribution policy is needed if policy-makers are concerned with social equality. Now we consider how such policy intervention works.

### 5.3.1 Tax rate adjustment

We first investigate the policy effect by resetting the income tax rate.[25] Figure 7 and Figure 8 summarise the IRF of the growth rate and capital inequality to all the 9 shocks, given different income tax rates. Both growth and inequality have similar tendency of IRF to all the shocks despite the gaps in magnitude when different tax rates are set. Overall, a lower tax rate which essentially encourages inequality leads to higher deviations from the actual level for both inequality (for those shocks which can have considerable effects on inequality) and growth. For instance, a negative shock on the penalty rate results in a higher inequality and growth, associated with a lower tax rate.

Now we consider growth and inequality in the full simulation with interactive shocks, given different income tax rates. A reduction in the tax rate $\tau$ from 0.4 to 0.2 in Figure 9 can stimulate growth by sacrificing social equality, especially in capital equality. The inverse policy by raising $\tau$ to reduce inequality therefore harms growth.

### 5.3.2 Income transfer from the rich to the poor

In this section we consider a more realistic and probably more efficient redistribution policy, income transfer from the rich to the poor.[26] For simplicity, we transfer a certain proportion $\tau_{Y1}$ of the income of the rich to the poor and assume no government spending is funded by income tax revenue to avoid multi-types of tax enforced. Given the same time seeds of bootstrapping, we compare the model with income transfer rate $\tau_{Y1} = 0.4$ with both the benchmark model and the model with no tax at all in Figure 10. As we see, compared with the model with no tax, income transfer also has a lower growth

[25] As the initial tax rate is as high as 0.4, we consider a reduction in $\tau$.

[26] The realisation of the income transfer needs an approximation in our model. See more details in Appendix H.

| Variables | Aggregate output shock | Aggregate capital shock | Aggregate consumption shock | Aggregate productivity shock | Consumption shock of group1 | Consumption shock of group2 | Labour shock of group1 | Labour shock of group2 | Aggregate penalty shock |
|---|---|---|---|---|---|---|---|---|---|
| R | 0.09% | 1.64% | 0.94% | 17.59% | 0.16% | 0.10% | 1.47% | 1.68% | 76.33% |
| Y | 0.06% | 1.10% | 0.69% | 11.51% | 0.06% | 0.07% | 0.72% | 2.41% | 83.38% |
| K | 0.09% | 1.57% | 1.01% | 16.84% | 0.13% | 0.11% | 1.04% | 1.25% | 77.97% |
| C | 0.04% | 0.76% | 0.43% | 8.15% | 2.21% | 3.70% | 0.68% | 0.78% | 83.24% |
| Y1 | 0.06% | 0.92% | 0.64% | 10.66% | 0.39% | 0.07% | 3.92% | 0.52% | 82.83% |
| Y2 | 0.06% | 1.16% | 0.70% | 11.64% | 0.11% | 0.07% | 0.95% | 3.66% | 81.65% |
| K1 | 0.09% | 1.54% | 0.90% | 16.68% | 0.08% | 0.09% | 0.65% | 1.49% | 78.48% |
| K2 | 0.09% | 1.59% | 0.91% | 16.88% | 0.16% | 0.10% | 1.41% | 0.83% | 78.03% |
| C1 | 0.04% | 0.75% | 0.43% | 8.05% | 7.11% | 0.04% | 0.67% | 0.77% | 82.14% |
| C2 | 0.05% | 0.80% | 0.46% | 8.65% | 0.08% | 0.05% | 0.72% | 0.82% | 88.36% |
| N1 | 0.05% | 0.80% | 0.58% | 9.50% | 1.25% | 0.06% | 12.84% | 0.41% | 74.50% |
| N2 | 0.06% | 1.07% | 0.65% | 10.68% | 0.10% | 0.07% | 0.87% | 11.91% | 74.59% |
| Growth | 0.22% | 6.05% | 6.76% | 0.92% | 1.32% | 0.85% | 9.98% | 11.98% | 61.92% |
| Inequality | 0.00% | 4.39% | 0.00% | 0.00% | 2.71% | 0.00% | 25.48% | 20.28% | 47.14% |

Table 7: Variance decomposition of key variables (static proportions)

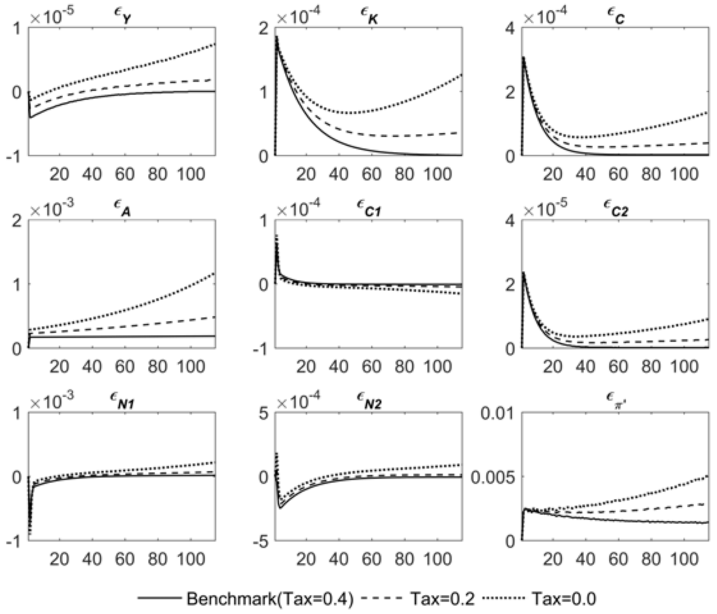


Figure 7: IRFs of the Growth Rate with different Income Tax Rates

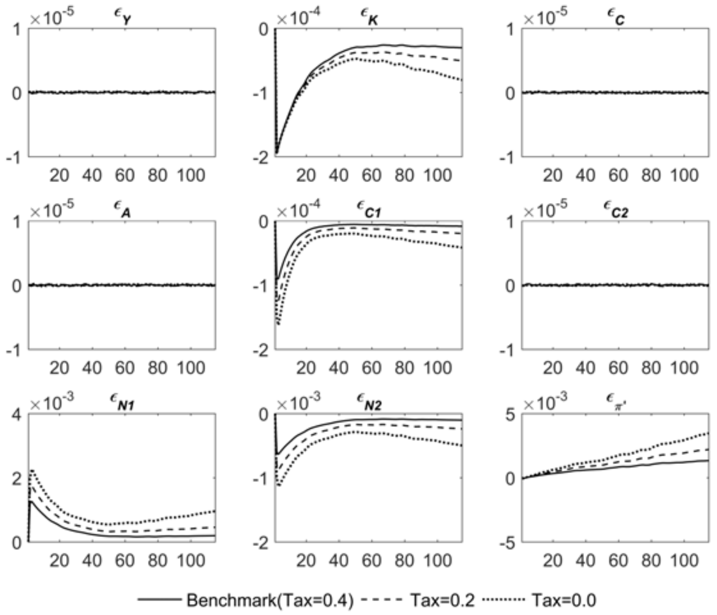


Figure 8: IRFs of the Capital Inequality with different Income Tax Rates

rate in spite of a more equal wealth distribution. However, compared with the proportional income tax (the benchmark model), income transfer generates both a higher growth rate and lower inequality. The reason for this is that tax takes money from the rich more than from the poor and then spends it on a government good that has no effect on household utility or therefore on any personal incentives; nor does it enter the production function as an input. Resources are therefore wasted: by taking the same from the rich (with the same effect on growth) direct transfer to the poor can generate a bigger reduction in inequality, improving the growth-inequality trade-off. The government can therefore obtain both more growth and less inequality by shifting from a pure income tax system to a direct transfer system.

Now we still have one question to consider. Is a rapid redistribution policy better or a moderate policy better? To answer this, we compare the redistributive consequences and costs with different $\tau_{Y1}$ in Figure 11. It can be seen the more redistributive policy is (moving from black through to green) the

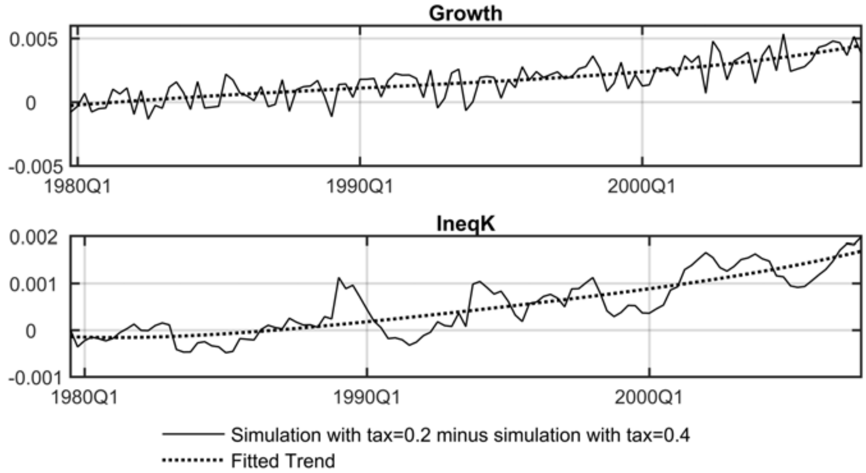


Figure 9: Redistribution by Changing the Income Tax Rate

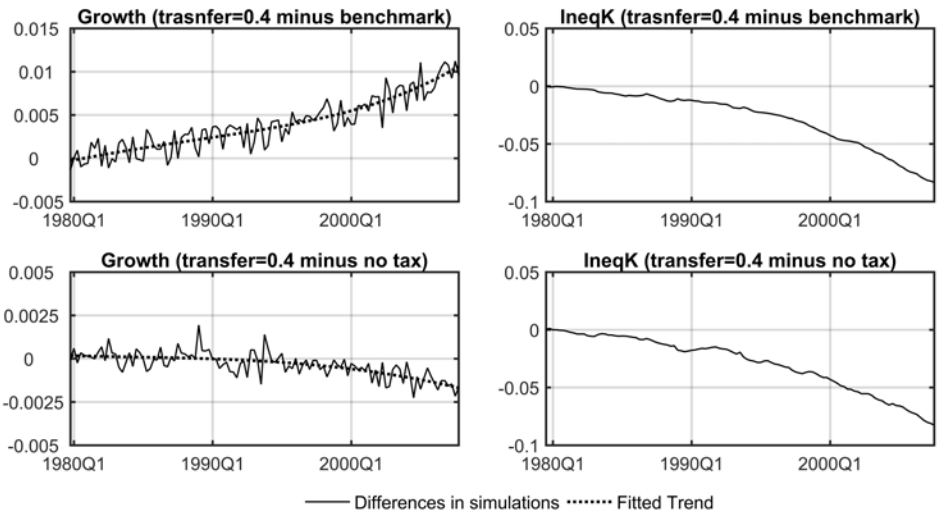


Figure 10: Redistribution by Income Transfer

higher the ratio of the effect on growth to the effect on inequality. Thus, an initial movement towards redistribution (from none to a rate of 0.1) lowers inequality a lot (nearly 0.10) but reduces growth only a little (about 0.002); whereas once the redistributive rate is as high as 0.3, the same movement reduces growth by nearly 0.01 and inequality by only 0.01. Hence the trade-off between the effect on inequality and that on growth becomes steadily worse as redistribution increases. This exhibits a 'diminishing return' of inequality to growth reduction. That is, the attempt to rapidly reduce the inequality by an extremely high transfer rate is improper.

# 6 Conclusions

This paper constructs a theoretical framework to investigate the relationship between capital inequality and economic growth in the UK during the recent decades. In our model, wealth inequality enhances entrepreneurship incentives of the rich to stimulate growth and the growth in turn aggravates inequality. When considering the 10%-90% income segmentation, our benchmark model cannot be rejected by the Indirect Inference test and can fit the main characteristics of the UK data. Policy-makers have to face a trade-off between wealth equalisation and economic growth when redistribution policy is conducted. For this an income transfer regime provides the best trade-off between growth and inequality. The trade-off worsens as the transfer rate rises which suggests that governments will limit the transfer rate. While our model is deliberately simplified and could be added to in various ways, it is striking that it can match UK data behaviour closely, while offering a theoretically attractive mechanism for explaining the broad

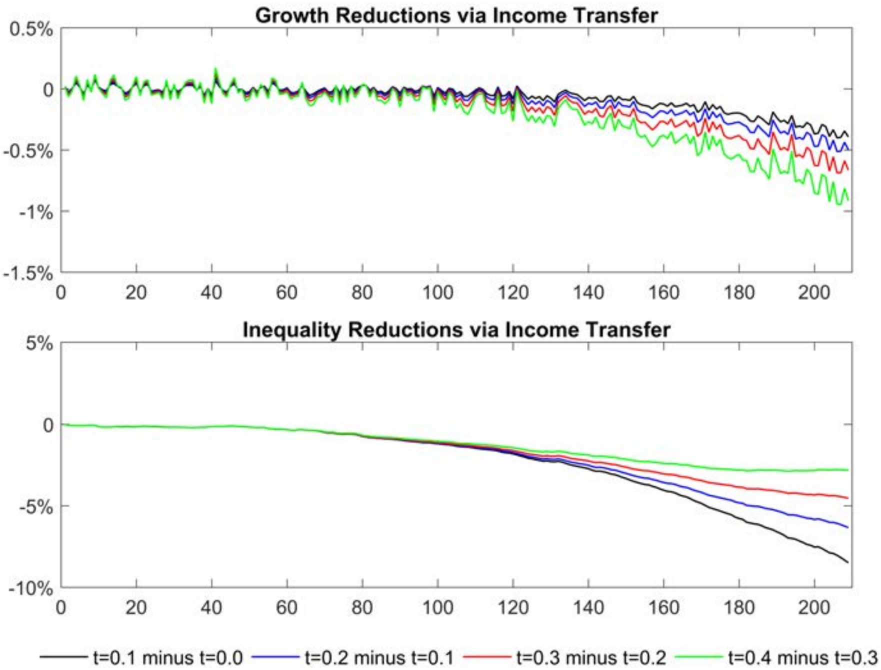


Figure 11: Redistribution by Different Income Transfer Rates

international correlations between inequality and growth.

# References


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

[2] 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): 875-908.

[3] Algan, Y., Allais, O. and Den Haan, W. J. and Rendahl, P. (2014). 'Solving and Simulating Models with Heterogeneous Agents and Aggregate Uncertainty.' In Handbook of Computational Economics, edited by Schmedders, K. and Judd, K. L., 277-48. Oxford: North Holland.

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

[5] Aiyagari, S. R. (1994). 'Uninsured Idiosyncratic Risks and Aggregate Saving.' Quarterly Journal of Economics, 109(3): 659-684.

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

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

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

[9] Bewley, T. F. (1980). 'Stationary Equilibrium.' Journal of Economic Theory, 24(2): 265-295.

[10] Bewley, T. F. (1983). 'A Difficulty with the Optimum Quantity of Money.' Econometrica, 51(5): 1485-1504.

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

[12] Davidson, R. and MacKinnon, J. G. (2004). Econometric Theory and Methods. Oxford New York: Oxford University Press.

[13] Dave, C. and De Jong, D. N. (2007). Structural Macroeconomics. New Jersey: Princeton University Press.

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

[15] 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.

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

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

[18] Giacomini, R. (2013). 'The relationship between DSGE and VAR models.' In Advances in Econometrics, edited by Fomby, T. B., Kilian, L. and Murphy, A., 1-25. Emerald Publishing Group Limited.

[19] Gourieroux, C., Monfort, A. and Renault, E. (1993). 'Indirect Inference.' Journal of Applied Econometrics, 8: 85-118.

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

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

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

[23] Kolev, G and Niehues, J. (2016). 'The Inequality-Growth Relationship: An Empirical Reassessment.' Cologne Institute for Economic Research IW Report 7.

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

[25] Krusell, P. and Smith, A. A., Jr. (2006). 'Quantitative Macroeconomic Models with Heterogeneous Agents.' In Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress, edited by Blundell, R., Newey, W.K. and T. Persson, T., 298-43. London: Cambridge University Press.

[26] 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): 2078-2104.

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

[28] 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.

[29] Levine, R. and Rubinstein, Y. (2013). 'Smart and Illicit: Who Becomes an Entrepreneur and Do They Earn More?' NBER Working Paper 19276.

[30] Meenagh, D., Minford, P. and Wang, J. (2007) 'Growth and relative living standards — testing barriers to riches on post-war panel data.' Cardiff Economics Working Papers E2007/12.

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

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

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

[34] 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): 117-120.

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

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

[37] Piketty, T. (2014). Capital in the Twenty-First Century. Translated by Arthur Goldhammer. Cambridge, Massachusetts London, England: Harvard University Press.

[38] Piketty, T. and Zucman, G. (2014). 'Capital is back: Wealth-income Ratios in Rich Countries 1700-2010.' Quarterly Journal of Economics, 129(3): 1255-1310.

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

[40] Smith, Anthony A., Jr. (1993). 'Estimating Nonlinear Time–series Models Using Simulated Vector Autoregressions.' Journal of Applied Econometrics, 8(S1): 63–84.

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

[42] 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): 36-41.


**Appendix A: Approximating C-Y ratio by a random walk**

Define $\widetilde{Y}_{i,t} = Y_{i,t} + (1-\delta)K_{i,t-1} - K_{i,t} - \pi_t Z_{i,t}$. Rewrite individual budget constraint as

$$(1+r_{t-1})b_{i,t} = C_{i,t} + b_{i,t+1} - \widetilde{Y}_{i,t} \tag{A1}$$

The condition of no Ponzi that present value of all the future increment of bonds should be equal to $(1+r_{t-1})b_{i,t}$ implies

$$(1+r_{t-1})b_{i,t} = C_{i,t} - \widetilde{Y}_{i,t} + E_t \sum_{j=1}^{\infty} \frac{C_{i,t+j} - \widetilde{Y}_{i,t+j}}{\Pi_{s=1}^{j}(1+r_{t+s-1})} \tag{A2}$$

Euler equation (5) can be approximated by

$$(C_{i,t})^{\Psi_1} = \frac{1}{\beta} E_t \left[ \frac{(C_{i,t+1})^{\Psi_1}}{(1+r_t)} \right] \approx \frac{1}{\beta^j} E_t \left[ \frac{(C_{i,t+j})^{\Psi_1}}{\Pi_{s=1}^{j}(1+r_{t+s-1})} \right] \tag{A3}$$

For simplicity, we set $\Psi_1 = 1$ which is also employed in the empirical study. Then (A3) can be simplified to $\frac{1}{\beta^j} E_t \left[ \frac{(C_{i,t+j})^{\Psi_1}}{\Pi_{s=1}^{j}(1+r_{t+s-1})} \right] = C_{i,t}$.[27] Rewrite (A2) as

$$\frac{C_{i,t}}{Y_{i,t}} = (1-\beta) \left[ (1+r_{t-1})\frac{b_{i,t}}{Y_{i,t}} + \frac{\widetilde{Y}_{i,t}}{Y_{i,t}} + \frac{1}{Y_{i,t}} E_t \sum_{j=1}^{\infty} \frac{\widetilde{Y}_{i,t+j}}{\Pi_{s=1}^{j}(1+r_{t+s-1})} \right] \tag{A4}$$

[27] In fact, we can also use this simplified form of (A3) as long as $\Psi_1$is close to unity which is true in many empirical papers.

where (A4) 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 $\widehat{Y_{i,t}} = \widetilde{Y}_{i,t} + E_t \sum_{j=1}^{\infty} \frac{\widetilde{Y}_{i,t+j}}{\Pi_{s=1}^{j}(1+r_{t+s-1})}$, discounted by the rate $(1-\beta)$.

The steady state of bonds is $(1+r_{T-1})b_{i,T} = b_{i,T+1}$. Bonds generally follow an AR process, $b_{i,t+1} = (1+r_{t-1})b_{i,t} + x_{i,t}$, before the steady state, which is nonstationary. This can be transformed to $\frac{b_{i,t+1}}{Y_{i,t+1}} \frac{Y_{i,t+1}}{Y_{i,t}} = (1+r_{t-1})\frac{b_{i,t}}{Y_{i,t}} + \frac{x_{i,t}}{Y_{i,t}}$ which implies that $\frac{b_{i,t}}{Y_{i,t}}$ before the steady state approximately has a unit root because the random growth rate $\frac{Y_{i,t+1}}{Y_{i,t}}$ is generally close to $r_{t-1}$. Hence, (A4) implies that $\frac{C_{i,t}}{Y_{i,t}}$ can also be approximated to a random walk.

**Appendix B: Derive the relation between the aggregate growth and inequality**

Given the linearised aggregate output equation, aggregate growth is

$$g_{t+1} \equiv \Delta \ln Y_{t+1} = \omega_{Y1}\Delta \ln Y_{1,t+1} + \omega_{Y2}\Delta \ln Y_{2,t+1} + \varepsilon_{Y,t+1}$$

Individual output growth is yielded using the linearised equations (C5) and (C6) in Appendix C.

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

Using (C7)-(C8) and (C11)-(C14) yields

$$\begin{aligned} \Delta \ln Y_{i,t+1} &= \alpha\left(\ln Y_{i,t} - \varphi r_t - \ln K_{i,t-1}\right) - \phi_{2,i}\pi'_{i,t} \\ &\quad + \frac{(1-\alpha)}{(1+\Psi_2)}\left[\Delta \ln Y_{i,t+1} - \Psi_1 \Delta \ln C_{i,t+1} + \frac{2\Psi_2\phi_{2,i}}{\theta_2}\Delta\pi'_{i,t+1}\right] + error \end{aligned} \tag{B1}$$

Taking $E_t$ on (B1) and Substituting out $\Delta E_t \ln C_{i,t+1}, \pi'_{i,t+1}$ and $\pi'_{i,t}$ using (12) and (C9) yields

$$\begin{aligned} \left(\frac{\alpha+\Psi_2}{1+\Psi_2}\right) E_t \Delta \ln Y_{i,t+1} &= \alpha(\ln Y_{i,t} - \ln K_{i,t-1}) - \left(\alpha\varphi + \frac{1-\alpha}{1+\Psi_2}\right) r_t \\ &\quad - \phi_{2,i}\left[\frac{2(1-\alpha)\Psi_2}{(1+\Psi_2)\,\theta_2} + 1\right]\left[\rho_1^{\pi}\pi'_{i,t-1} - \rho_2^{\pi} Q(\frac{K_{i,t-2}}{K_{t-2}})\right] \\ &\quad + \frac{2(1-\alpha)\Psi_2\phi_{2,i}}{(1+\Psi_2)\,\theta_2}\left[\rho_1^{\pi}\pi'_{i,t} - \rho_2^{\pi} Q(\frac{K_{i,t-1}}{K_{t-1}})\right] + error \end{aligned}$$

To get rid of the past $\pi'_{i,t-s}; s>1$, set $\rho_1^{\pi} = 0$ to obtain[28]

$$\begin{aligned} \left(\frac{\alpha+\Psi_2}{1+\Psi_2}\right) E_t \Delta \ln Y_{i,t+1} &= \alpha(\ln Y_{i,t} - \ln K_{i,t-1}) - \left(\alpha\varphi + \frac{1-\alpha}{1+\Psi_2}\right) r_t \\ &\quad + \rho_2^{\pi}\phi_{2,i}\left\{\left[\frac{2(1-\alpha)\Psi_2}{(1+\Psi_2)\,\theta_2} + 1\right] Q(\frac{K_{i,t-2}}{K_{t-2}}) - \frac{2(1-\alpha)\Psi_2}{(1+\Psi_2)\,\theta_2} Q(\frac{K_{i,t-1}}{K_{t-1}})\right\} + error \end{aligned}$$

Aggregating the equations above across individuals with assumption that $\phi_{2,i}$ is same dnoted by $\phi_{2,2}$ across groups for simplicity yields

$$\begin{aligned} \left(\frac{\alpha+\Psi_2}{1+\Psi_2}\right) E_t g_{t+1} &= \alpha(\ln Y_t - \ln K_{t-1}) - \left(\alpha\varphi + \frac{1-\alpha}{1+\Psi_2}\right) r_t \\ &\quad + \rho_2^{\pi}\phi_{2,2}\left\{\left[\frac{2(1-\alpha)\Psi_2}{(1+\Psi_2)\,\theta_2} + 1\right] \Sigma\omega_{Yi} Q(\frac{K_{i,t-2}}{K_{t-2}}) - \frac{2(1-\alpha)\Psi_2}{(1+\Psi_2)\,\theta_2} \Sigma\omega_{Yi} Q(\frac{K_{i,t-1}}{K_{t-1}})\right\} + error \end{aligned} \tag{B2}$$

We set $Q(\frac{K_{i,t-1}}{K_{t-1}}) = \frac{\mu_i}{\omega_{Yi}}(\frac{K_{i,t-1}}{K_{t-1}})^2$ where $q_t = \frac{K_{i,t-1}}{K_{t-1}}$ also measures capital inequality and define the aggregate term $Q_t \equiv \Sigma\omega_{Yi} Q(\frac{K_{i,t-1}}{K_{t-1}})$. Then

$$Q_t = \sum \omega_{Yi}\frac{\mu_i}{\omega_{Yi}}(\frac{K_{i,t-1}}{K_{t-1}})^2 = \mu_1 (q_t)^2 + \mu_2\left[\frac{1}{\mu_2}(1-\mu_1 q_t)\right]^2$$

[28] This assumption will only magnify the effect of inequality on growth without changing the sign.

$$\frac{dQ_t}{dq_t} = 2\mu_1 \left[ q_t \left( 1 + \frac{\mu_1}{\mu_2} \right) - \frac{1}{\mu_2} \right] = \frac{2\mu_1}{\mu_2} (q_t - 1)$$

Note $\frac{dQ_t}{dq_t} > 0$ if $q_t > 1$ while $\frac{dQ_t}{dq_t} < 0$ if $q_t < 1$. Hence, $Q_t$ has has minimum at perfect equality ($q_t = 1$). Equation (B2) now can be rewritten as

$$\left( \frac{\alpha + \Psi_2}{1 + \Psi_2} \right) E_t g_{t+1} = \rho_2^\pi \phi_{2,2} \left\{ \left[ \frac{2(1-\alpha)\Psi_2}{(1+\Psi_2)\,\theta_2} + 1 \right] Q_{t-1} - \frac{2(1-\alpha)\Psi_2}{(1+\Psi_2)\,\theta_2} Q_t \right\} + \cdots \tag{B3}$$

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, $g_L$, by summing up temporary growth rates within a long period, we yield the following[29]

$$\left( \frac{\alpha + \Psi_2}{1 + \Psi_2} \right) g_L \approx \rho_2^\pi \phi_{2,2} \left\{ \left[ \frac{2(1-\alpha)\Psi_2}{(1+\Psi_2)\,\theta_2} + 1 \right] \Sigma_{t=0}^{T} Q_{t-1} - \frac{2(1-\alpha)\Psi_2}{(1+\Psi_2)\,\theta_2} \Sigma_{t=0}^{T} Q_t \right\} + \cdots \tag{B4}$$

Since the other growth determinant real interest rate endogenously depends on lagged output and capital, both short-run and long-run growth rate in (B3) and (B4) can be approximated as a reduced form of lagged output, capital and inequality which is the usual form of existing empirical regression studies. Note that $Q_t > 0$ and $\Sigma_{t=0}^{T} Q_{t-1} \approx \Sigma_{t=0}^{T} Q_t$ for a long term. Since $\frac{2(1-\alpha)\Psi_2}{(1+\Psi_2)\theta_2} + 1 > \frac{2(1-\alpha)\Psi_2}{(1+\Psi_2)\theta_2}$, the long-run growth rate is minimised when capital distribution is perfectly equal. 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 C: Linearisation**

(C1) and (C9) are obtained from individual Euler equation (5) and the consumption aggregation equation. $\omega_{Y1}$ and $\omega_{K1}$ in (C2) and (C3) are the steady-state income share and capital share for the rich. To linearise individual capital equation (6), we use (5) to rewrite it as $K_{i,t} = \alpha(1-\tau)\frac{E_t Y_{i,t+1}}{\delta + r_t}$. We use the approximation $E_t \left( \frac{Y_{i,t+1}}{C_{i,t+1}} \right) \approx \frac{Y_{i,t}}{C_{i,t}}$ to linearise it as $\ln K_{i,t} \approx \frac{1}{\delta + r} \left\{ \alpha(1-\tau) \frac{Y_i}{K_i} \left[ \ln Y_{i,t} + E_t \ln C_{i,t+1} - \ln C_{i,t} \right] - r_t \right\}$, regardless of some constant terms. Finally use (5) to obtain (C7) and (C8). To linearise labour equation (7), we firstly take the $1^{st}$ order Taylor expansion with the approximation $1 - N_i - Z_i \approx 0.5$ and $N_i \approx 0.5$ and then substitute $Z_{i,t}$ out using (4) and (11). Lastly, individual bonds are removed from equation list because they take small share over individual capital resource which we are not interested in.

The list of linearised equations is[30]

$$r_t = \Psi_1 (E_t \ln C_{2,t+1} - \ln C_{2,t}) - \ln \beta \tag{C1}$$

$$\ln Y_t = \omega_{Y1} \ln Y_{1,t} + \omega_{Y2} \ln Y_{2,t} + \varepsilon_{Y,t} \tag{C2}$$

$$\ln K_t = \omega_{K1} \ln K_{1,t} + \omega_{K2} \ln K_{2,t} + \varepsilon_{K,t} \tag{C3}$$

$$\ln C_t = (1-\tau)\frac{Y}{C} \ln Y_t - \frac{K}{C} \left[ \ln K_t - (1-\delta) \ln K_{t-1} \right] + \varepsilon_{C,t} \tag{C4}$$

$$\ln Y_{1,t} = \alpha \ln K_{1,t-1} + (1-\alpha) \ln N_{1,t} + \ln A_{1,t} \tag{C5}$$

$$\ln Y_{2,t} = \alpha \ln K_{2,t-1} + (1-\alpha) \ln N_{2,t} + \ln A_{2,t} \tag{C6}$$

$$\ln K_{1,t} = \varphi_1 \left( \ln Y_{1,t} + \frac{1}{\Psi_1} \ln \beta \right) - \varphi_2 r_t \text{ [31]} \tag{C7}$$

[29] For the long-run growth rate, we use an approximation that $\Pi_{t=1}^{T}(1 + g_t) \approx 1 + \Sigma_{t=1}^{T} g_t$

[30] Some linearisation is conducted around the steady state in despite of nonstationarity because the steady state in our model is the one conditional on the nonstationary predetermined variables or exogenous shocks ("conditional steady state").

[31] $\varphi_1 = \frac{(1-\tau)\alpha Y_1}{(\delta+r)K_1} = \frac{(1-\tau)\alpha Y}{(\delta+r)K}$; $\varphi_2 = \frac{1}{(\delta+r)} - \frac{\varphi_1}{\Psi_1}$

$$\ln K_{2,t} = \varphi_1 \left( \ln Y_{2,t} + \frac{1}{\Psi_1} \ln \beta \right) - \varphi_2 r_t \tag{C8}$$

$$\ln C_{1,t} = E_t \ln C_{1,t+1} - \frac{1}{\Psi_1}(r_t + \ln \beta) + \varepsilon_{C1,t} \tag{C9}$$

$$\ln C_{2,t} = \frac{1}{\omega_{C2}}(\ln C_t - \omega_{C1} \ln C_{1,t}) + \varepsilon_{C2,t} \tag{C10}$$

$$\ln N_{1,t} = \frac{1}{(1+\Psi_2)} \left( \ln Y_{1,t} - \Psi_1 \ln C_{1,t} + 2\frac{\Psi_2 \phi_{2,1}}{\theta_2} \pi'_{1,t} \right) + \varepsilon_{N1,t} \tag{C11}$$

$$\ln N_{2,t} = \frac{1}{(1+\Psi_2)} \left( \ln Y_{2,t} - \Psi_1 \ln C_{2,t} + 2\frac{\Psi_2 \phi_{2,1}}{\theta_2} \pi'_{2,t} \right) + \varepsilon_{N2,t} \tag{C12}$$

$$\ln A_{1,t+1} = \ln A_{1,t} + \phi_{1,1} - \phi_{2,1} \pi'_{1,t} + \varepsilon_{A,t} \tag{C13}$$

$$\ln A_{2,t+1} = \ln A_{2,t} + \phi_{1,1} - \phi_{2,1} \pi'_{2,t} + \varepsilon_{A,t} \tag{C14}$$

$$\pi'_{1,t} = \rho_0^{\pi} + \rho_1^{\pi} \pi'_{1,t-1} - \rho_2^{\pi} \cdot Q(\frac{K_{1,t-2}}{K_{t-2}}) + \varepsilon_{\pi',t} \tag{C15}$$

$$\pi'_{2,t} = \rho_0^{\pi} + \rho_1^{\pi} \pi'_{2,t-1} - \rho_2^{\pi} \cdot Q(\frac{K_{2,t-2}}{K_{t-2}}) + \varepsilon_{\pi',t} \tag{C16}$$

**Appendix D: Auxialiary model**

Meenagh et. al. (2012) show that the solution of a DSGE model with nonstationary variables can be approximated as a VARX. Suppose $A(L)y_t = BE_t y_{t+1} + C(L)x_t + D(L)e_t$ is a structural model where $y_t$, $x_t$ and $e_t$ are vectors of endogenous variables, nonstationary variables and *i.i.d.* shocks respectively. Given nonstationary $y_t$ and $x_t$ (as $y_t$ depends on $x_t$), there should be cointegrations which requires both $y_t$ and $x_t$ are I(1). Then a short-run equilibrium (stationary shocks are zeros) can be written as $y_t = \Pi x_t + g$, a conditional steady state (also the form of our terminal conditions). As the long-run solution for $x_t$ comprises a deterministic trend and a stochastic trend, model solution finally can be approximated as $y_t = C + Py_{t-1} + Qx_{t-1} + \phi t + \nu_t$ which is a VARX conditional on nonstationary $x_{t-1}$, deterministic trend $\phi t$ and stochastic term $\nu_t$.

**Appendix E: Model details on data**

As the available data on both CCB and MCD are annual ones and are incomplete in our time horizon, we firstly supplement the omitted annual observations using 3-points quadratic estimating interpolation. Then to generate quarterly series, we follow Minford (2016) to use a Denton method with a highly frequent instrument, "trade union membership" rate (TUM) which is the fraction of the number of employees who are trade union members out of the total number of employees. Since both CCB and MCD describe how lax the regulation is, we define the final instrument by the inverted "1-TUM". Afterwards, we can yield quarterly data of CCB and MCD using Denton method. Lastly, these quarterly data are inverted (in order to describe costs instead of benefits) and scaled less than unity (to keep consistent with the magnitude of the tax rate).

**Appendix F: Impulse response functions in the benchmark model**

**Appendix G: Variance decomposition in the benchmark model**

We firstly draw random innovations from normal distributions with model-implied standard deviations on one structural error individually to obtain 1,000 simulated samples. Then we calculate the variance of each endogenous variable along the time horizon and average variances over samples for each variable. Repeat this on each structural error and calculate the proportions corresponding to different errors for each variable. This decomposition tells us whether the change in a variable caused by a certain innovation is stable across time.

**Appendix H: Income transfer regime**

The realisation of the income transfer needs approximation in our model; otherwise the transferred income to the poor will not affect their individual capital accumulation and consumption because the individual budget constraints themselves are not used as model equations. We use an approximation to avoid this trap. Suppose a constant income tax rate $\tau_{Y1}$ is enforced on the rich. The tax revenue

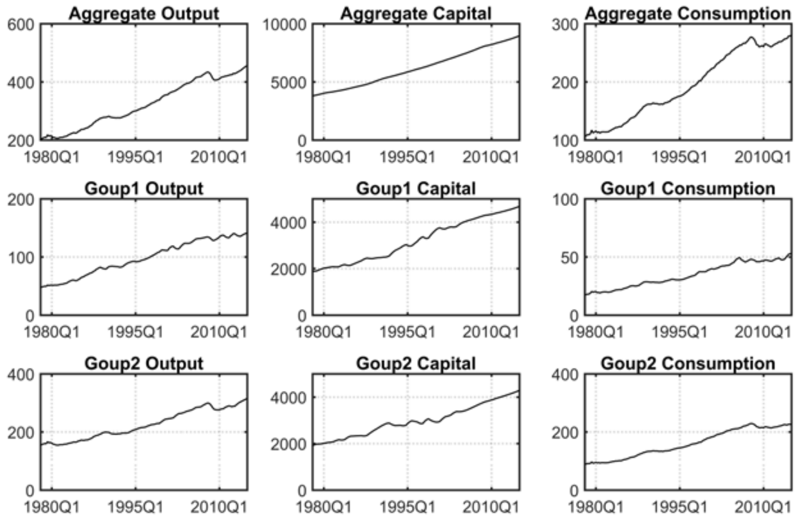


Figure 12: Data on Output, Capital and Consumption (Billion GBP)

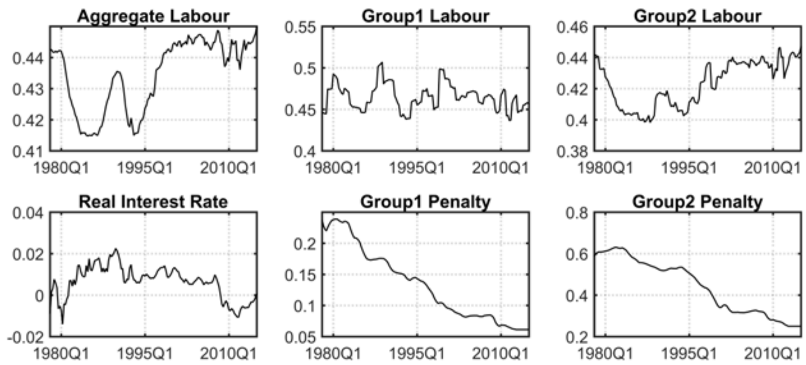


Figure 13: Data on Real Interest Rate, Individual Labour and Individual Penalty

per capita across the whole population now is $\tau_{Y1}\mu_1 Y_{1,t}$ which is transferred to the poor with the population weight $\mu_2$, i.e. $\tau_{Y1}\mu_1 Y_{1,t} = \tau_{Y2}\mu_2 Y_{2,t}$. Then we can approximate the subsidy rate for the poor $\tau_{Y2} = \frac{\mu_1 Y_{1,t}}{\mu_2 Y_{2,t}}\tau_{Y1}$ by $\tau_{Y2} \approx \frac{\omega_1}{\omega_2}\tau_{Y1}$. Hence individual income of the rich and the poor after transfer are respectively $(1-\tau_{Y1})Y_{1,t}$ and $(1+\frac{\omega_1}{\omega_2}\tau_{Y1})Y_{2,t}$. Here we use the approximation $\frac{\omega_1}{\omega_2} \approx \frac{\omega_{1,t}}{\omega_{2,t}}$. Some model equations should be modified to meet the transfer setting. The parameter $\varphi_1$ in (C7) and (C8) now are $\frac{(1-\tau_{Y1})\alpha Y_{1,t}}{(\delta+r)K}$ and $\frac{(1+\frac{\omega_1}{\omega_2}\tau_{Y1})\alpha Y_{1,t}}{(\delta+r)K}$ respectively.[32] Parameters $\phi_{2,1}$ and $\phi_{2,2}$ now become $\phi_{2,1} = \frac{(1-\tau_{Y1})\beta\theta_2\frac{Y_1}{w_1}}{(1-\beta)(1-\tau_{Y1}+\pi_1')^2}$ and $\phi_{2,1} = \frac{(1+\frac{\omega_1}{\omega_2}\tau_{Y1})\beta\theta_2\frac{Y_2}{w_2}}{(1-\beta)(1+\frac{\omega_1}{\omega_2}\tau_{Y1}+\pi_2')^2}$ respectively. To reflect the effect of transfer on $\phi_{2,i}$, we back out $\pi_i'$ from the benchmark estimated $\phi_{2,i}$ and recalculate it with all the other components unchanged. Table 8 shows the updated values of $\phi_{2,i}$ given different transfer rates.

| Marginal penalty effects | $\tau_{Y1} = 0.1$ | $\tau_{Y1} = 0.2$ | $\tau_{Y1} = 0.3$ | $\tau_{Y1} = 0.4$ |
|---|---|---|---|---|
| $\phi_{2,1}$ | 0.5032 | 0.5248 | 0.5464 | 0.5666 |
| $\phi_{2,2}$ | 0.4433 | 0.4374 | 0.4316 | 0.4257 |

Table 8: Marginal penalty effects on productivities given different transfer rate

[32] The linearised individual labour equations (C11) and (C12) now have extra constant terms involving transfer tax which could be computed by "Type II Fix".

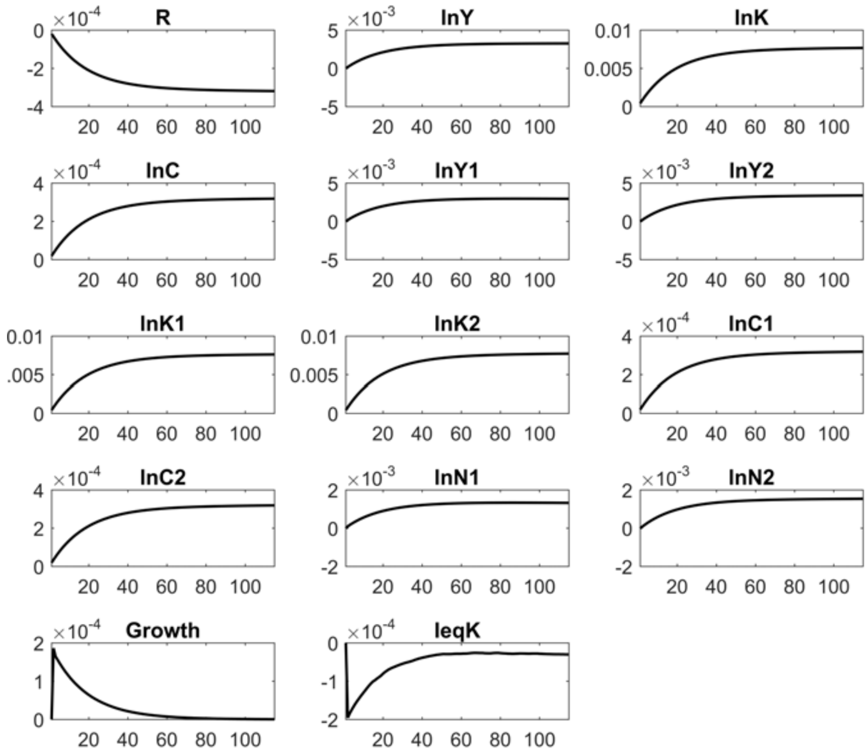


Figure 14: IRF to a one s.d. aggregate capital temporary shock

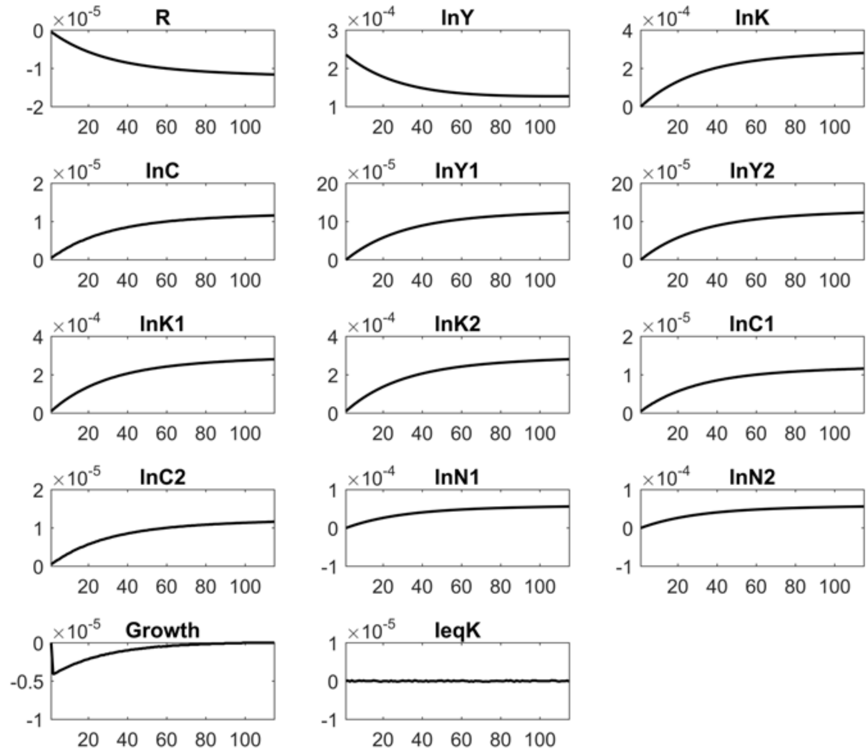


Figure 15: IRF to a one s.d. aggregate output temporary shock

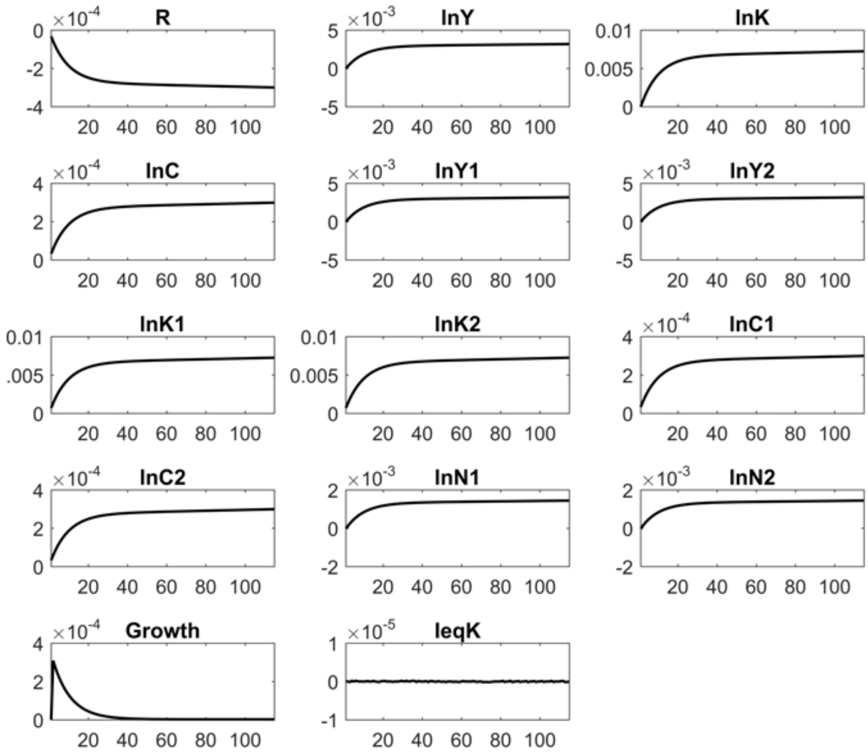


Figure 16: IRF to a one s.d. aggregate consumption (market clearing) shock

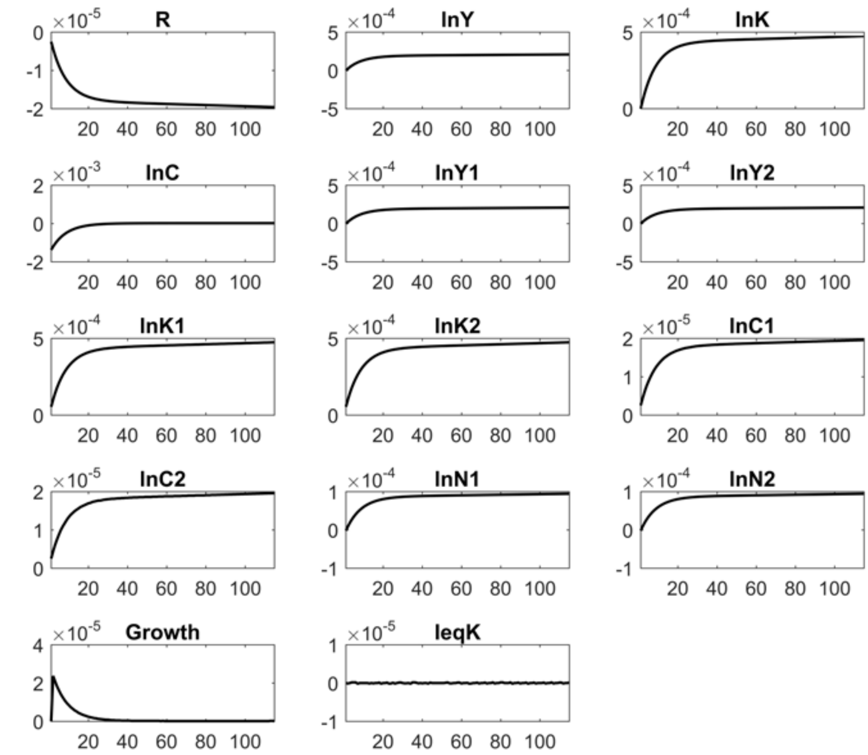


Figure 17: IRF to a one s.d. individual consumption temporary shock on the poor

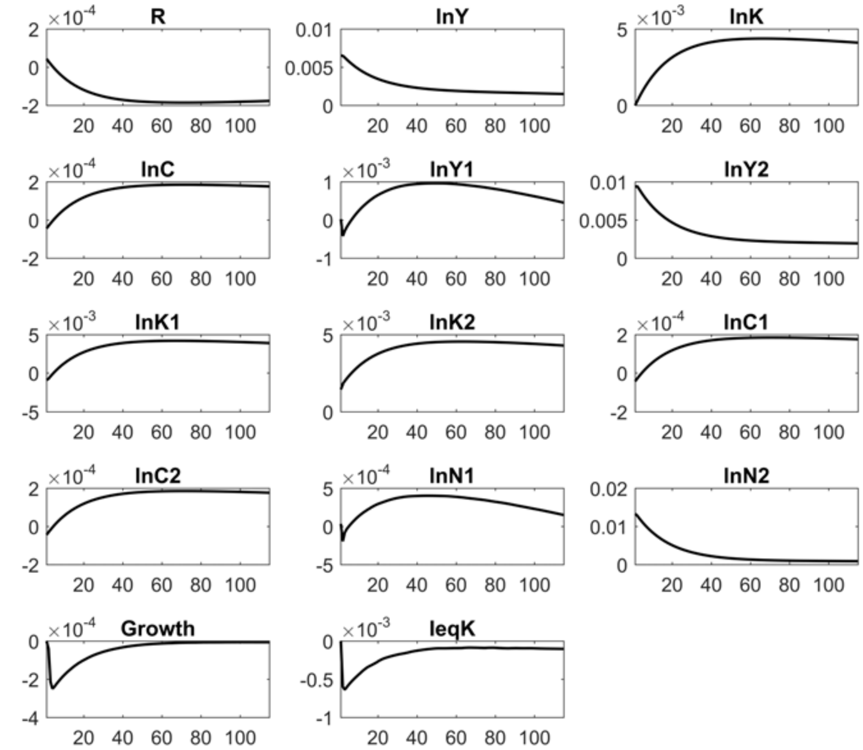


Figure 18: IRF to a one s.d. individual labour supply temporary shock on the poor