Can subsidising job-related training reduce inequality?*

Konstantinos Angelopoulos
University of Glasgow and CESifo

Andrea Benecchi, University of Glasgow
University of Glasgow

James Malley
University of Glasgow and CESifo

January 30, 2018

Abstract

A well-established stylised fact is that employer provided job-related training raises productivity and wages. Using UK data, we further find that firm-provided training is positively related to subsidies aimed at reducing training costs for employers. We also find that there is a positive, albeit quantitatively small, relationship between wage inequality and training inequality in the UK. Motivated by the above, we explore whether policies to subsidise firms' monetary cost of training can improve earnings for the lower skilled and reduce inequality. We achieve this by developing a dynamic general equilibrium model, featuring skilled and unskilled labour, capital-skill complementarity in production and an endogenous training allocation. Our results suggest that training subsidies for the unskilled have a significant impact on the labour income of unskilled workers. These subsidies also increase earnings for skilled workers and raise aggregate income with implied lifetime multipliers exceeding unity. Finally, the positive spill over effects to skilled workers imply that training subsidies are not very effective in reducing inequality, measured as the distance between skilled and unskilled wages and incomes.

Keywords: Job-related training, wage and earning inequality, training subsidies
JEL Classification: E24, J24, J31

*We would like to thank Richard Dennis, Spyridon Lazarakis and Tatiana Kirsanova for helpful comments and suggestions. Any errors are our own.

1 Introduction

Job-related training has been a quantitatively and economically significant activity in the UK in recent decades. According to data from the Quarterly Labour Force Survey (QLFS), about a quarter of workers received some training in each quarter since the mid-1990s. A substantial empirical literature, which includes both academic and policy-related research, has examined the importance and characteristics of employee training. The existing research suggests that job-related training is beneficial to both employers and employees by positively contributing to productivity and wages, although the gains tend to be larger for employers (see, e.g. Blundell *et al.* (1999), Dearden *et al.* (2006) and Konings and Vanormelingen (2015)). Data for the UK from the QLFS, the Continuing Vocational Training Survey (CVTS) and the Employer Skills Survey (ESS), also suggest that the costs of job-related training are to a large extent covered by the employer.[1]

Given the productivity and wage benefits associated with training, the latter could contribute to increasing earnings for workers with lower skills and reducing labour income inequality between groups of workers with different skills. Indeed, a key observation relating to the UK labour market since the 1980s is the existence of pronounced earnings and wage inequality accompanied by a stagnation of wages for the lower income groups since mid-2000s (see e.g. Blundell and Etheridge (2010), Brewer and Wren-Lewis (2016), Belfield *et al.* (2017) and Angelopoulos *et al.* (2017a)). An important dimension of inequality in the UK and in other countries has been the earnings differential between University and non-University educated workers (see e.g. Goldin and Katz (2008) and Heathcote *et al.* (2010) for the US as well as Blundell and Etheridge (2010) and Angelopoulos *et al.* (2017a,b) for the UK). In the UK, wage inequality related to University education increased since 1980, and despite reductions between 1995 and 2005, the skill premium remains high.[2] The UK economy is thus characterised by persistently lower labour market returns for those with lower skills.

A natural response aimed at improving earnings for low skilled labour and thus generating a reduction in inequality by closing the gap from the lower end is to enhance the skills and productivity of those with lower education by improving the quality of basic education (see e.g. Machin and Vignoles (2005), Wößmann and Schütz (2006), and Autor (2014)). Academics and policymakers have considered complementing such efforts by interventions to improve the skills and productivity of workers already in the labour market. These have been applied to individuals with high school degrees through ongoing vocational training and lifelong learning schemes (see e.g. Stevens (1999), Sofer (2004), Bassanini *et al.* (2007) and the Eu-

[1] Further details relating to QLFS, CVTS, and ESS are discussed in the next section and Appendix A.

[2] We will present and discuss data for the UK in more detail in the next Section.

ropean 2020 Strategy).[3] In other words, more intensive training of the lower skilled non-University educated workers could improve their productivity and earnings.

The literature has noted that policy interventions in training can be justified in terms of equality of opportunity (see e.g. Machin and Vignoles (2001), Greenhalgh (2002), Bassanini *et al.* (2007) and Busemeyer (2014)).[4] However, the evidence to date suggests that there is inequality for those who participate in training, i.e. the more skilled and better able workers are trained more. For instance, data from the European Community Household Panel analysed in Bassanini *et al.* (2007)) demonstrate that there is a gap in training participation between workers of different education levels and of different family background. Similarly, data for the UK from QLFS also reveals a gap in training between workers of different education levels.

Given that training is related to labour productivity and returns, it is reasonable to expect training inequality to feed into earnings inequality. Our data analysis below, using QLFS data, finds that education-based training inequality is indeed related to education-based wage inequality. However, the implied elasticity is small, suggesting that changes in training inequality are not likely to have a big impact on wage inequality.

The mechanism by which job-related training increases and determines its subsequent effect on wage growth and wage inequality, is complex. This is mainly because job-related training takes place at the expense of work time and is thus largely dependent on firm's choices, being affected by the structure of production and changes in other inputs. In particular, a firm's decision to train its employees can be expected to depend upon factors such as: (i) the efficiency of training time in creating labour productivity; (ii) the monetary costs for training; and (iii) returns to improved worker productivity for a given increase in worker skills, which in turn depends on the structure of production (e.g. capital-skill complementarity and skill-unskilled substitutability). The government cannot dictate to firms whom and how much to train, but it can try to encourage training by reducing the cost of the investment in training by the firm, and, in particular, the monetary costs associated with employees' training.[5] In our data analysis below, we find that UK firms that receive a higher financial training subsidy, train a higher

[3] Note that here we refer to training of those in employment, as opposed to training unemployed workers or individuals who leave the labour force to study.

[4] In contrast, policy interventions to encourage training for efficiency reasons are more difficult to justify, since under-provision of training as a result of market failures is harder to establish (see e.g. Bassanini *et al.* (2007) and Brunello and de Paola (2009)).

[5] The train-or-pay scheme, where firms face levies if they don't train their workforce, has been abandoned by UK due its unpopularity among entrepreneurs in the 90's (see Bassanini *et al.* (2007)). Also, Dostie (2015) reports that such a scheme does not have a significant impact on training in Canada, one of the few countries that still employ this scheme.

proportion of their employees.

In light of the above, we aim to evaluate the quantitative implications of policies that increase employer's incentives to train workers. We construct a dynamic general equilibrium model that is consistent with the main features of job-related training and wage inequality outlined above and allows for the relevant policy interventions. We focus on the effects of such policies on inequality in training, skill and wages. To model job-related training and skill creation, we build on a large literature of partial and dynamic general equilibrium models with job-related learning and labour productivity in the form of human capital (see e.g. Huggett *et al.* (2006), Kim and Lee (2007), Mejía and St-Pierre (2008), Moreno-Galbis (2012), Manuelli and Seshadri (2014) and Chen and Lai (2015)). The general idea is that, subject to a cost, a portion of the worker's time is invested in learning skills that will improve future productivity, so that job-related training implies both a monetary and a time opportunity cost. In each period, time in training is used with existing job-related skills to improve future job-related skills. In turn, the stock of job-related skills and worker time that is not diverted to training are combined to create the quantity of effective labour input.

What defines this form of skill acquisition as job-related training is that, in our model, the decision to train is made by the employer and training time is explicitly at the cost of work time. In particular, the firm assumes both the monetary and opportunity costs related to training and decides which proportion of employees (or, equivalently, of their time) to train. It simultaneously appropriates the rent from having a more productive stock of labour. Workers benefit in that their wages increase, albeit at a lower rate than their productivity, consistent with the evidence discussed earlier. While this approach adds complexity to the problem of the firm by making it intertemporal, it is nevertheless consistent with the empirical evidence discussed above, showing that it is the firms, rather than workers, that primarily cover the cost of training. It also follows that allowing the firms to decide on training is essential for the evaluation of the effect of policy aimed at redressing training inequality by incentivising job-related training.

We add wage inequality to this setup by allowing for *ex ante* heterogeneity between University and non-University educated workers and a production structure that allows for capital-skill complementarity. In particular, University educated employees work in occupations (or jobs) that are more complementary to capital that those of non-University educated workers.[6] This standard mechanism leads to a University wage premium that has been extensively analysed in the literature, see e.g. Krusell *et al.* (2000), Goldin and Katz (2008) and Acemoglu and

[6]Since we focus on wage (and not wealth or income) inequality, we follow the unemployment literature since Merz (1995) (see e.g. Rogerson and Shimer (2011) for a review) and simplify the model by allowing for perfect consumption insurance between the University and non-University educated members of the single household.

Autor (2011). In the context of job-related training, this further creates different incentives for the firm to train skilled (University educated) and unskilled (non-University educated) employees. Since these employees have different marginal products of effective labour, there is a different (and higher) marginal return to increasing skilled, relative to unskilled, job-related skills and effective labour input. Moreover, the elasticity of skill creation with respect to job-related training is allowed to differ between the two types of workers, and thus is allowed to reflect differences in the efficiency of training.

We calibrate the model to data from the UK and ensure that it generates training and wage inequality that are consistent with the data. We then evaluate policies that target training for the unskilled workers by subsidising the firms and reducing the relevant financial cost. The model predictions regarding the magnitude of the effect of training subsidies to training participation and of the effect of the reduction in training inequality on wage inequality, are consistent with the empirical evidence we collect. In fact, with respect to both relationships, the model predicts effects just below the lower bound of our estimates. However, despite the conservative calibration, there is a significant impact on wages and earnings for workers. In particular, training subsidies significantly increase wages and labour income of the target group, and there are sizeable positive spillover effects from subsidising the training of each group of workers to the other group.

For instance, a policy to subsidise a quarter of the costs to train unskilled workers can increase their wages (earnings) by 0.23% (0.75%, for earnings), 10-years following the implementation of the policy, and by 0.58% (1.06%, for earnings) in the long-run. Moreover, there are sizeable effects on skilled workers, who benefit from the increased productivity of unskilled workers. Continuing with the same example, skilled workers would experience an increase in their wages (earnings) by 0.06% (0.16%, for earnings), 10-years following the implementation of the policy, and by 0.42% (0.51%, for earnings) in the long-run. These positive spillover effects are important in generating wider social gains from a more targeted policy. In addition, they are helpful in reconciling the small effect that training inequality has on wage inequality in the data (and model) with the strong impact that training has on wages in both the empirical literature and the model.

The increase in lifetime income, both in terms of labour income and in terms of aggregate income, is greater than the present value of the resources required for such a policy, implying associated fiscal multipliers that are greater than unity. What underlies these significant effects is first the strong impact of training on returns to labour and second the spillover effects that work in general equilibrium to enhance the positive effect on any labour input.[7] Subsidies to increase job-

[7] The effect of the increase in training on wages that is implied by the model is consistent with empirical estimates in Blundell *et al.* (1996).

related training of unskilled workers lead to a fall in wage and income inequality, while subsidising training of skilled workers leads to an increase in inequality. On the other hand, the positive impact of training subsidies for skilled workers is bigger in terms of aggregate quantities.

Overall, our results suggest that while subsidising job-related training does not have a big impact in reducing "inequality", it can nevertheless be effective in contributing to improvements of income for the lower skilled. In fact, the conclusion that training subsidies do not significantly reduce "inequality" should not be viewed as a downside of training subsidies, but instead as a welcoming consequence of the positive spillovers that they create for the skilled workers.

The rest of the paper is organized as follows. In Section 2, we review existing empirical findings and present additional empirical evidence on training, training inequality and its relationship with wage inequality, as well as the importance of subsidies for training decisions. In Section 3 we develop the model we use for the quantitative evaluation of the nexus between training, inequality and policy and we discuss its calibration and quantitative relevance. In Section 4, we evaluate the effects of policy aimed at redressing training inequality by incentivising job-related training. Section 5 contains the conclusions.

2 Training costs, returns and inequality

We next review and add to the empirical evidence on the extent of job-related training, its importance for employers/firms and employees/participants as well as its effect on wage inequality. Job-related training refers to training of individuals who are in employment, either as employees in a firm or as self-employed. We will refer to these individuals collectively as "workers". In subplot (1,1) of Figure 1 we plot workers' participation in this type of training in the UK, using quarterly data from the QLFS from 1995.1 to 2015.4.[8] In particular, we calculate the proportion of workers who received training within the 13 weeks prior to the interview date. As can be seen, following a large rise in the 1990s, this proportion has stabilised in recent periods to about 25%, implying that one in four workers receives some type of training every quarter.[9]

[Figure 1 here]

[8] The QLFS provides data using international definitions of employment and unemployment and economic inactivity, together with a wide range of related topics such as occupation, training, hours of work, and personal characteristics of household members aged 16 years and over. Further details regarding the data can be found in Appendix A.

[9] The UK is not an outlier in the European context. In many other European countries training participation is also high (see, e.g. Markowitsch *et al.* (2013) who use the CVTS dataset from Eurostat).

Job-related training on this scale should be motivated by significant returns. Indeed, empirical studies document a strong positive effect from employee training to firm productivity, as well as a positive relationship between wages and training (see e.g. Blundell *et al.* (1999), Haelermans and Borghans (2012), and Méndez and Sepúlveda (2016)). The estimated effects vary between different samples and methods used in the literature, but overall imply benefits to both employers and employees from job-related training (for reviews, see Leuven (2004) and De Grip and Sauermann (2013)). Returns to firms are typically estimated to be higher than returns to workers, and are more robustly significant (see, e.g. Hansson (2008)). A positive effect of training on productivity is also confirmed in studies for the UK (Dearden *et al.* (2006)).

The significant returns that firms realise from training their employees lead them to encourage it.[10] In fact, data from the QLFS, CVTS, and ESS demonstrate that firms in the UK pay for more than 70% of job-related training, and that about half of this training takes place during work time, implying that it is costly in terms of foregone output. Government subsidies cover approximately 4.17% of total training costs (see the evidence from CVTS, waves 3 and 4 (2005 and 2010)).[11] The importance of firms' contribution to training expenses is also confirmed using European-level data. In particular, Bassanini *et al.* (2007) analysed data from CVTS for European countries and finds that employer-provided training represents a major component of training, and that workers do not pay for job-related training through lower initial salaries or flatter wage-tenure profiles. Their results also suggest that training spells paid by employers represent about 70-80% of the total training expenditures (see Bassanini *et al.* (2007)).

Overall, empirical research has linked job-related training to productivity gains. Moreover, existing empirical analyses have also demonstrated that there is inequality in participation and in the returns from training. Bassanini *et al.* (2007)) analyse European data from the European Community Household Panel and demonstrate that there is a gap in training participation between workers of different education levels and of different family background. Moreover, they find that training increases wages more for better educated workers. We further elaborate on training inequality and the relationship between training and wage inequality

[10] In theory, firms are more likely to cover the cost of employee training if the latter is firm-specific. Otherwise, if it is mainly general purpose, then it would be more natural to expect that workers would finance training (see e.g. Becker (1962)), especially if firms engage in poaching of employees from other firms. However, there is empirical evidence, at least for the US, that firms support employee training despite poaching (see e.g. Parent (1999)). The data for the UK suggest significant firm-sponsored training (see e.g. O'Mahony (2012)), consistent with the high returns to firms from training their employees.

[11] Note that this ratio is based on gross receipts over total training costs as reported by surveyed firms.

in the UK in the following sub-sections.

2.1 Inequality in training and wage inequality

In the following analysis we use data from the QLFS on training, wages, employment and hours of work by education groups between 1995.1-2015.4. We split the sample into the group of workers who have at least a bachelor degree or high level qualification (University educated) and those without these qualifications (non-University educated). For each education group, we compute the training participation rate as the ratio of workers who have been trained in the last quarter over the total number of workers. To obtain a measure of training inequality between the two groups, we calculate the ratio of the University educated to non-University educated training participation rates (see, subplot (1,2) in Figure 1). As can be seen, despite significant reductions in the period 1995-2004, training inequality remains high, at about 1.6, without significant reductions since 2005.

Since training contributes to increased productivity and returns, training inequality can contribute to wage inequality. Although this seems like a plausible speculation, we are not aware of existing research demonstrating a direct link between training inequality and wage inequality. The University skill premium has declined in recent decades in the UK, as shown in subplot (2,1) of Figure 1 (see e.g. Blundell and Etheridge (2010) and Brewer, Wren-Lewis (2015), Belfield *et al.* (2017) and Angelopoulos *et al.* (2017a) for an analysis of inequality in the UK).[12] This can be linked to increased University education,[13] which implies that the relative supply of skilled labour over unskilled has grown, as shown in subplot (2,2) in Figure 1.[14] Indeed, as the scatterplot in subplot (3,1) in Figure 1 shows, there is a negative relationship between wage inequality and the relative skill supply in the UK for the period 1995-2015. However, the decline in wage inequality can also be linked to the decline in the training inequality, as the scatterplot in subplot (3,2) in Figure 1 shows. In fact, it is interesting to note that the trend in wage inequality is more consistent with the trend in training inequality. In particular, note that the biggest reduction in wage inequality took place between 1995-2004, the period where training inequality also reduced significantly, whereas after 2005

[12] The skill premium is the ratio of the average skilled to the unskilled wage over the period 1995.1-2015.4. Workers include both employees and self-employed individuals who are between 25 and 65 years old. The wage is computed by dividing weekly labour income by the number of hours worked per week from the main job.

[13] See Goldin and Katz (2008) for evidence on the role education plays in wage inequality.

[14] Using the same definitions for skilled and unskilled as above, the relative skill supply is defined as the ratio of the product of skilled (weekly) working hours and the skilled population share to the product of the same two measures for unskilled workers using QLFS data from 1995.1 to 2015.4.

both series exhibit a smaller slope. In contrast, the growth rate of the relative skill supply increased after 2005, the slope being smaller prior to this date.

To further investigate the relationship between wage inequality and training inequality, we regress the former on the latter and on the relative supply of skilled to unskilled labour. In particular, we consider the following relationship:

$$\frac{w_t^s}{w_t^u} = \alpha_1 + \alpha_2 \frac{p_t^s}{p_t^u} + \alpha_3 \frac{n_t^s}{n_t^u} + \sum\nolimits_{i=1}^{3} \gamma_{it} Q_{it} + \varepsilon_t, \tag{1}$$

where $\frac{w_t^s}{w_t^u}$ is the ratio of wages for skilled or University educated, w_t^s, to unskilled or non-University educated, w_t^u, employees in period t; $\frac{p_t^s}{p_t^u}$ is the ratio of training participation for skilled, p_t^s, to unskilled, p_t^u, employees; and $\frac{n_t^s}{n_t^u}$ is the ratio of skilled, n_t^s, to unskilled, n_t^u, employees. Given that training exhibits quarterly regular variation (see e.g. Felstead *et al.* (2013)), we include a set of quarterly time dummies, $\sum_{i=1}^{3} \gamma_{it} Q_{it}$. Finally, $\varepsilon_t \sim Niid(0, \sigma_\varepsilon^2)$ is the error term.

The results for the coefficients of interest are reported in Table 1.[15] We also report an F-statistic for the joint significance of the three quarterly time dummies. Finally, we present the F-statistic for a test of serial correlation, obtained by regressing the residuals $\widehat{\varepsilon}_t$ on four lagged values and testing for their joint significance. As can be seen, both p_t^s/p_t^u and n_t^s/n_t^u are significant at the 2% and 7.8% levels, and the estimated coefficients have the expected signs. Hence, the results suggest that training inequality is positively related to wage inequality, even after controlling for the change in the education composition of the labour force. Further note that the 95% confidence interval for $\widehat{\alpha}_2$ ranges from 0.022 to 0.249.

Table 1: Wage and Training Inequality

	$\widehat{\alpha}_1$	$\widehat{\alpha}_2$	$\widehat{\alpha}_3$	$\widehat{\gamma}_{1t} = \widehat{\gamma}_{2t} = \widehat{\gamma}_{3t} = 0$		Serial correlation	
estimate	1.464	0.136	-0.093	$F(3,78)$	6.720	$F(4,75)$	0.340
p-value	0.000	0.020	0.078	p-value	0.000	p-value	0.852

2.2 Cost-incentives matter

As discussed above, in the UK, firms assume the larger share of the costs to train workers. On the other hand, the government's contribution to monetary costs is very small. It would thus be useful in the analysis which follows to know whether the decision to train employees is sensitive to subsidies to the direct monetary costs that job-related training entails. We are not aware of existing evidence on

[15] Note that this relationship has also been tested on a panel dataset composed of yearly observations for 1-digit SIC UK sectors. The results of the analysis, available on request, are consistent with the results of the time-series regression performed on the whole UK economy and reported in Table 1.

the link between training subsidies and training participation at the firm level.[16] We thus next explore this link by using sectoral data from the QLFS and the CVTS, editions 2005 and 2010, which report information about training subsidies and training costs for about 4000 companies.

We first compute the per firm nominal average training subsidies and training costs by SIC sector according to the classification reported in each dataset. In 2005, the CVTS employs a classification with 35 sectors, while the 2010 edition classifies firms into 25 different sectors. Due to changes in the classification, we can only match 17 sectors between the two datasets. Thus, to make best use of the available data, we merge them into an unbalanced panel dataset. We use this data to generate the ratio of training subsidies to training costs which is denoted sub_{it} in the regression below.

Using the QLFS, we next compute the training participation rate for each two digit SIC in 2005 and in 2010 in annual terms. The training participation rate is defined as the ratio between the number of workers who have received training in any quarter and the total number of workers. This variable is denoted s_{it} in the regression below.

We finally combine the sectoral QLFS training participation rate data with the corresponding sectors in the CVTS database. In some cases, we aggregate two or more sub-sectors to match the definition used in the CVTS. In such instances, the number of workers of each sector is used as weight to compute the average participation rate.

To exploit the cross-sectional and time-series dimensions in the sample discussed above, we undertake a panel data random effects analysis. In particular, we estimate the following model:

$$s_{it} = \beta_1 + \beta_2 sub_{it} + \beta_3 size_{it} + \mu_{i,t}, \tag{2}$$

where s_{it} is the share of employees in sector i that received training in period t; and sub_{it}, is the share of training costs that have been received as a training subsidy, on average, by firms of sector i in period t. Given that sectors with bigger firms may train a higher share of their employees to exploit economies of scale in training provision, we also include the average number of employees per firm, $size_{it}$, in the model. Finally, we allow for further unobserved sector heterogeneity captured by the error term and consistent with a random effects specification.

The results from estimating (2) are shown in Table 2. As can be seen, both $\hat{\beta}_2$ and $\hat{\beta}_3$ are positive and significant. Moreover, the results of the Hausman test

[16] In the literature, empirical studies consider several determinants for training provision at sectoral, regional or national level, such as economic density (e.g. Brunello and Gambarotto (2004)), market power (e.g. Bilanakos *et al.* (2017)), and size (e.g. Almeida and Aterido (2015)). Notably, none of these works have been able to control or study the effect of fiscal incentives due to lack of data.

indicate that the random effects cannot be rejected in favor of the fixed effect model. The estimate of coefficient of $\widehat{\beta}_2$, indicates that an increase in the subsidy (as a share of total training cost) of 1%, tends, on average, to increase the share of workers that are trained by about 0.26%, suggesting an inelastic response. Further note that the 95% confidence interval for this coefficient ranges from from 0.051 to 0.473.[17]

Table 2: Training subsidies and participation

	$\widehat{\beta}_1$	$\widehat{\beta}_2$	$\widehat{\beta}_3$	Hausman test	
estimate	19.46	0.262	0.020	$\chi^2(2)$	1.930
p-value	0.015	0.000	0.000	p-value	0.381

3 The model

To evaluate the quantitative implications of policies that raise firms' incentives to train low (high) skilled workers, we construct a dynamic general equilibrium model that coheres with the main stylised facts relating to the UK job-related training and wage inequality data reviewed above. The key features of the model are (i) *ex ante* skill heterogeneity between non-University educated (unskilled) and University (skilled) workers, leading to wage inequality under capital-skill complementarity in production and (ii) job-related training and skill creation undertaken by firms separately for skilled and unskilled workers.

When we analyse the quantitative implications of policies, we focus on their effects on inequality in training, skill accumulation and wages. In particular, we examine subsidies to encourage the productivity of training time and skill accumulation which are financed by a lump-sum tax on the household. To gauge the effects of such policies, we first solve the model at the steady-state, choosing the parameters so that the steady-state is similar to the actual UK economy. We then take this as the initial position of the economy and evaluate the effects of one-off, permanent changes in policy by simulating the economy as it converges to its new long-run equilibrium.

3.1 Representative household

There is an infinitely lived representative household that is comprised of unskilled and skilled members. Superscripts u and s are used in what follows to denote unskilled and skilled respectively. We assume that household members share the household labour and asset income and have equal consumption irrespective of

[17] As robustness check, we include as regressor the average contribution to training funds (as a percent of training costs). This control variable is statistically and economically insignificant and it does not affect the estimation of the β_2 coefficient.

their labour market status (as in e.g. the literature on unemployment since Merz (1995)). This allows us to focus on wage inequality without additional modeling assumptions required to enrich the setup with wealth and consumption inequality. In this setup, the head of the household makes all choices on behalf of its members, differentiated in our case, by labour market skills. In particular, the head of the household maximises discounted lifetime household utility:

$$U = \sum_{t=0}^{\infty} \beta^t \frac{\left\{ c_t^{\psi_1} \left[n^u \left(1 - l_t^u \right) \right]^{\psi_2} \left[n^s \left(1 - l_t^s \right) \right]^{\psi_3} \right\}^{(1-\sigma)}}{1-\sigma}, \tag{3}$$

where, $0 < \beta < 1$ is the time discount factor; c_t is per capita consumption; n^i $(i = u, s)$ is the share of each skill type to total household members $(n^u + n^s = 1)$; l_t^i is labour supply; $\sigma > 1$ is the coefficient of relative risk aversion; and the parameters $0 < \psi_1, \psi_2, \psi_3 < 1$, $\psi_1 + \psi_2 + \psi_3 = 1$, represent the weights that the household attaches to consumption, unskilled leisure, $(1 - l_t^u)$, and skilled leisure, $(1 - l_t^s)$ in utility respectively.

The household's budget constraint is:

$$c_t + k_{t+1} - \left(1 - \delta^k\right) k_t = n^u w_t^u l_t^u + n^s w_t^s l_t^s + r_t k_t + \pi_t - T_t, \tag{4}$$

where k_t is physical capital; $0 < \delta^k < 1$ is the capital depreciation rate; w_t^i is the wage rate; r_t is the net return to capital; π_t is profits; and T_t is a lump-sum transfer/tax. The labour productivity advantages, for University educated workers, work directly via differences in w_t^s versus w_t^u, which in turn capture differences between the marginal productivity of skilled versus unskilled labour input, as it will become apparent when we examine production.

The Lagrangian for the household is given by:

$$\begin{aligned} L = \textstyle\sum_{t=0}^{\infty} \{ & \beta^t \frac{\left\{ c_t^{\psi_1} [n^u (1-l_t^u)]^{\psi_2} [n^s (1-l_t^s)]^{\psi_3} \right\}^{(1-\sigma)}}{1-\sigma} - \beta^t \lambda_t^k [c_t + k_{t+1} - \left(1 - \delta^k\right) k_t - \\ & - n^u w_t^u l_t^u - n^s w_t^s l_t^s - r_t k_t - \pi_t + T_t] \}, \end{aligned} \tag{5}$$

where $\lambda_t^k > 0$, is the Lagrange multiplier. The household chooses $\{c_t, l_t^s, l_t^u, k_{t+1}\}_{t=0}^{\infty}$ taking the initial condition, k_0, the policy variable, $\{T_t\}_{t=0}^{\infty}$, prices, $\{w_t^u, w_t^s, r_t\}_{t=0}^{\infty}$ and profits $\{\pi_t\}_{t=0}^{\infty}$ as given. The static first-order condition (FOC) for consumption:

$$\lambda_t^k = \frac{\psi_1 \left[\left(n^u \left[1 - l_t^u \right] \right)^{\psi_2} \left(n^s \left[1 - l_t^s \right] \right)^{\psi_3} \right]^{1-\sigma}}{c_t^{1-\psi_1(1-\sigma)}}, \tag{6}$$

states that the shadow price of the budget constraint (4) is equal to the marginal utility of consumption, $\frac{\partial U}{\partial c_t}$, at time t.

The *intratemporal* FOCs for unskilled and skilled labour supply:

$$\frac{\psi_2}{\psi_1}\frac{c_t}{n^u\left(1-l_t^u\right)}=w_t^u, \tag{7}$$

$$\frac{\psi_3}{\psi_1}\frac{c_t}{n^s\left(1-l_t^s\right)}=w_t^s, \tag{8}$$

imply that the marginal rates of substitution between leisure (unskilled and skilled) and consumption at time t, i.e. $\frac{\partial U}{\partial\left(1-l_t^i\right)}/\frac{\partial U}{\partial c_t}$, are equal to the unskilled and skilled wage rates, respectively.

Finally, the Euler equation for capital:

$$\frac{1}{\beta}\left[\frac{c_t^{\psi_1}\left(n^u\left[1-l_t^u\right]\right)^{\psi_2}\left(n^s\left[1-l_t^s\right]\right)^{\psi_3}}{c_{t+1}^{\psi_1}\left(n^u\left[1-l_{t+1}^u\right]\right)^{\psi_2}\left(n^s\left[1-l_{t+1}^s\right]\right)^{\psi_3}}\right]^{1-\sigma}\frac{c_{t+1}}{c_t}=1+r_{t+1}-\delta_k \tag{9}$$

says that the marginal rate of substitution between consumption at time t and $t+1$, $\frac{\lambda_t^k}{\lambda_{t+1}^k}\equiv\frac{\partial U}{\partial c_t}/\frac{\partial U}{\partial c_{t+1}}$, is equal to the gross return to capital, $1+r_{t+1}$, net of capital depreciation.

3.2 Representative firm

There is an infinitely lived representative firm, which is owned by the household and employs capital, unskilled and skilled labour to produce a homogeneous final good. Production takes place using the following constant elasticity of substitution (CES) production technology:

$$\widetilde{y}_t^f=A\left\{\mu\left(z_t^u\right)^\alpha+\left(1-\mu\right)\left[\rho\left(k_t^f\right)^\nu+\left(1-\rho\right)\left(z_t^s\right)^\nu\right]^{\frac{\alpha}{\nu}}\right\}^{\frac{1}{\alpha}}, \tag{10}$$

where, $\widetilde{y}_t^f$ is the firm's output; $A>0$ is total factor productivity; $0<\mu,\rho<1$ are the factor share parameters; z_t^i is the effective labour input used in production; k_t^f is the demand for capital; and $\alpha,\nu<1$ are the parameters defining the factor elasticities, i.e. $1/\left(1-\alpha\right)$ is the elasticity of substitution between capital and effective unskilled labour as well as between effective skilled and effective unskilled labour; whereas $1/\left(1-\nu\right)$ is the elasticity of substitution between capital and effective skilled labour. Capital-skill complementarity in production, which is obtained in this setup when $\alpha>\nu$, has been shown to be empirically relevant and a contributor to wage inequality. This is because an increase in capital stock and capital augmenting technology in this setup are skill biased (see e.g. Krusell *et al.* (2000), Hornstein *et al.* (2005), Caselli and Coleman (2006), and Goldin and Katz (2008)).

The firm hires $l_t^{f,i}$ hours from the labour market, but not all of it is used for production, as some of the workers' time is used for training purposes. By denoting the share of workers' time in job-related training by t_t^i, this implies that the net time actually used for production is given by $l_t^{f,i}(1-t_t^i)$, whereas $l_t^{f,i}t_t^i$ is the actual time devoted to job-related training. Training increases next period's labour productivity. In particular, building on the human capital tradition since Ben Porath (1967), and following e.g. Huggett *et al.* (2006), we assume that labour productivity, or else the stock of skills accumulated via job-related training evolves according to the following laws of motion:

$$h_{t+1}^u = (1-\delta^u)\, h_t^u + H^u \left(l_t^{f,u} t_t^u h_t^u\right)^{\gamma^u}, \tag{11}$$

$$h_{t+1}^s = (1-\delta^s)\, h_t^s + H^s \left(l_t^{f,s} t_t^s h_t^s\right)^{\gamma^s}, \tag{12}$$

where $0 < \delta^u, \delta^s < 1$ are the depreciation rates for skills accumulated by unskilled and skilled workers respectively; $H^u \left(l_t^{f,u} t_t^u h_t^u\right)^{\gamma^u}$ is new skills created at time t; $H^i > 0$ is the productivity in new skill creation; and $\gamma^i < 1$ captures the elasticity of new skills with respect to existing skills and training time. Note that both H^i and γ^i are related to workers' learning ability (see Huggett *et al.* (2006)), i.e. the ability of the workers to use existing skills with their time for training to create new labour skills. This ability is fixed at the point of their entry in the labour market. Since both sets of parameters relate to the same economic concept, we will normalise in what follows H^i to be unity and let γ^i capture differences in learning ability associated with University education.

The restriction that $\gamma^i < 1$ guarantees that there is well-defined (bounded) steady-state value for h^i, thus precluding growth in the stock of skills in the long-run. At the same time, $\gamma^i < 1$ leaves open the possibility of increasing or decreasing returns to scale in creating labour productivity. Importantly, following a basic assumption largely employed in the literature since the seminal work of Mincer (1992), we allow learning ability to differ between skilled and unskilled workers, reflecting their different education status prior to entering the labour market.

The firm thus incurs an opportunity cost in terms of foregone workers' time when it decides to train its employees. Moreover, we assume that it incurs a monetary cost. The benefit for the firm is that the labour productivity generated by job-related training increases effective labour input. In particular, the effective labour input z_t^i is a function of workers' time and of labour productivity:

$$z_t^s = \left[l_t^{f,s}(1-t_t^s)\right]^{\omega} [h_t^s]^{1-\omega}, \tag{13}$$

$$z_t^u = \left[l_t^{f,u}(1-t_t^u)\right]^{\omega} [h_t^u]^{1-\omega}, \tag{14}$$

where $0 < \omega < 1$ measures the elasticity of effective labour with respect to production time. Note that the constant returns to scale restriction in (13)-(14) implies that the production function (10) is also constant returns to scale in its five inputs $\left[l_t^{f,i}\left(1-t_t^i\right)\right]$, h_t^i , for $(i = s, u)$, and k_t^f

This setup implies that it is the firm, and not the worker, which assumes the costs of training and also owns job-related skills associated with h_t^i, thus capturing firm-specific skills that are augmented by job-related training.[18] As explained in Section 2, this is consistent with empirical evidence which suggests that (i) firms pay for the majority of job-related training of their employees and (ii) that the returns to productivity and firm profitability/returns from job-related training are estimated to be larger than the effect of job-related training on workers' wages, implying significant rents for the firms associated with job-related training. Indeed, in this specification, and given that the production function in (10) is constant returns to scale, the compensation to labour productivity in the form of h_t^i is captured by the firm as a rent associated with training its employees, and takes the form of profits. Therefore, the higher the contribution of the firm-owned factor h_t^i in production, which is captured by a lower ω, the higher the firm's profitability associated with investment in employee training.

The firm's problem is formalised as follows. The representative firm aims to maximize the present discounted value of lifetime profits (e.g. Chen and Lai (2015)[19]):

$$\Pi^f = \sum_{t=0}^{\infty} Q_t \pi_t, \tag{15}$$

where

$$Q_t = \prod_{j=0}^{t-1} \left(1 + r_{j+1} - \delta^k\right)^{-1}, \tag{16}$$

defines the discount factor[20] and

$$\pi_t^f = \widetilde{y}_t^f - r_t k_t^f - w_t^u l_t^{f,u} - w_t^s l_t^{f,s} - \phi^u t_t^u l_t^{f,u}\left(1-\tau^u\right) - \phi^s t_t^s l_t^{f,s}\left(1-\tau^s\right), \tag{17}$$

[18] This is therefore different from partial or general equilibrium studies where on-the-job training is modelled as a household's decision variable, as in e.g. Huggett *et al.* (2006) or Kim and Lee (2007).

[19] Note that in the setup in Chen and Lai (2015), all new hires are unskilled and firms train automatically all new recruits who then become skilled in the second period. Hence, in their setup, training does not increase the productivity of skilled and unskilled workers in their tasks, but rather serves as a means to move workers through tasks.

[20] We follow the convention: $Q_0 = \prod_{j=0}^{-1}\left(1 + r_{j+1} - \delta^k\right) = 1$.

denotes profits, which are defined as the revenue from selling the final good, minus the costs of capital, the costs of unskilled and skilled labour, as well the monetary training costs for unskilled and skilled labour. The parameter $0 < \phi^i < 1$ refers to the fixed cost per training hour; and τ^i is a subsidy or tax on training activities.

The intertemporal trade-off associated with training time is evident in equations (10)-(14) and (17). In particular, *ceteris paribus*, an increase in training time raises new skills at time t and the stock of skills in $t+1$. Hence, effective labour and output in $t+1$ increase. However, training incurs a resource outlay. In addition, by lowering the time available for work at time t, effective labour and output at time t fall.

This setup further creates different incentives for the firm to train its skilled and unskilled employees which we observe in the UK data. In particular, since the employees have different marginal products of effective labour, there is a different (and higher) marginal return to increasing skilled, relative to unskilled, job-related skills and effective labour input. Moreover, if the learning ability for skilled workers is higher, i.e. $\gamma^s > \gamma^u$, then the increase in labour productivity is higher, for a given amount of training time, for skilled versus unskilled workers (see e.g. Almeida and Faria (2014)). On the other hand, if training skilled workers implies a relatively higher monetary cost (i.e. if $\phi^u < \phi^s$), then the firm has a disincentive to train skilled, versus unskilled workers. In this case, relative training between skilled and unskilled depends on the quantitative evaluation of this trade-off.

Taking the initial conditions, $\left\{k_0^f, h_0^s, h_0^u\right\}$, prices, $\{w_t^s, w_t^u, r_t\}_{t=0}^{\infty}$, and the discount factor $\{Q_t\}_{t=0}^{\infty}$ as given, the firm chooses $\{k_t^f, l_t^{f,u}, l_t^{f,s}, t_t^u,\ t_t^s, h_{t+1}^u,\ h_{t+1}^s\}_{t=0}^{\infty}$ to maximise (15), subject to (11) and (12).[21] The Lagrangian for the firm is given by:

$$\begin{aligned}\Lambda = \textstyle\sum_{t=0}^{\infty}\{ & Q_t\{y_t^f - r_t k_t^f - w_t^u l_t^{f,u} - w_t^s l_t^{f,s} - \phi^u t_t^u l_t^{f,u}\left(1-\tau^u\right) - \\ & -\phi^s t_t^s l_t^{f,s}\left(1-\tau^s\right)\} - \\ & -Q_t\lambda_t^u[h_{t+1}^u - \left(1-\delta^u\right)h_t^u - H^u\left(l_t^{f,u}t_t^u h_t^u\right)^{\gamma^u}] - \\ & -Q_t\lambda_t^s[h_{t+1}^s - \left(1-\delta^s\right)h_t^s - H^s\left(l_t^{f,s}t_t^s h_t^s\right)^{\gamma^s}]\},\end{aligned} \tag{18}$$

where λ_t^i are the shadow prices associated the skill accumulation constraints (11)

[21] This is equivalent to a setup where: (i) a branch of the firm faces a static problem and decides on capital and labour demand, taking training time and labour productivity as given; and (ii) another branch faces the intertemporal problem of choosing training time and labour skill acquisition, as long as both branches have the same objective function in (17).

and (12); and y_t^f is given by:

$$y_t^f = A_t \left\{ \mu \left(\left[l_t^{f,u} \left(1 - t_t^u\right) \right]^{\omega} \left[h_t^u\right]^{1-\omega} \right)^{\alpha} + \right.$$
$$\left. + (1-\mu) \left[\rho \left(k_t^f \right)^{\nu} + (1-\rho) \left(\left[l_t^{f,s} \left(1 - t_t^s\right) \right]^{\omega} \left[h_t^s\right]^{1-\omega} \right)^{\nu} \right]^{\frac{\alpha}{\nu}} \right\}^{\frac{1}{\alpha}}. \tag{19}$$

The static FOCs with respect to capital, unskilled and skilled labour:[22]

$$r_t = \frac{\partial y_t^f}{\partial k_t^f}, \tag{20}$$

$$w_t^u + \phi^u t_t^u \left(1 - \tau^u\right) = \frac{\partial y_t^f}{\partial l_t^{f,u}} + \lambda_t^u \frac{\partial h_{t+1}^u}{\partial l_t^{f,u}}, \tag{21}$$

$$w_t^s + \phi^s t_t^s \left(1 - \tau^s\right) = \frac{\partial y_t^f}{\partial l_t^{f,s}} + \lambda_t^s \frac{\partial h_{t+1}^s}{\partial l_t^{f,s}}, \tag{22}$$

equate their respective marginal costs to their marginal products. In the presence of job-related training and skill accumulation, marginal costs are comprised of the wage costs, w_t^i, and the marginal increase in monetary costs of training, $\phi^i t_t^i$, net of the tax or subsidy, τ^i. The corresponding marginal products are comprised of the marginal product of labour in output, $\frac{\partial y_t^f}{\partial l_t^{f,i}}$, plus the marginal product of labour in skill accumulation, $\frac{\partial h_{t+1}^i}{\partial l_t^{f,i}}$, valued by its corresponding shadow price, λ_t^i. Hence, the second term in the right hand side of these two FOCs captures the benefit to the firm from increasing work time since this allows for more time to train and thus for increased future labour labour productivity.

The intratemporal FOCs with respect to unskilled and skilled training time:

$$\phi^u l_t^{f,u} \left(1 - \tau^u\right) - \frac{\partial y_t^f}{\partial t_t^u} = \lambda_t^u \frac{\partial h_{t+1}^u}{\partial t_t^u}, \tag{23}$$

$$\phi^s l_t^{f,s} \left(1 - \tau^s\right) - \frac{\partial y_t^f}{\partial t_t^s} = \lambda_t^s \frac{\partial h_{t+1}^s}{\partial t_t^s}, \tag{24}$$

equate their respective marginal costs to their marginal products. Marginal costs are equal to the opportunity cost of foregone output, $\frac{\partial y_t^f}{\partial t_t^i}$, due to time being diverted from work, plus as above, the marginal increase in monetary costs of training time, net of the tax or subsidy. The corresponding marginal products are the marginal product of training time in skill accumulation, $\frac{\partial h_{t+1}^i}{\partial t_t^i}$, valued by its corresponding shadow price, λ_t^i.

[22] All of the derivatives listed in the following FOCs are defined in Appendix B.

Finally the Euler equations for unskilled and skilled skills acquisition:

$$\lambda_t^u = \frac{Q_{t+1}}{Q_t}\left(\frac{\partial y_{t+1}^f}{\partial h_{t+1}^u} + \lambda_{t+1}^u \frac{\partial h_{t+2}^u}{\partial h_{t+1}^u}\right), \tag{25}$$

$$\lambda_t^s = \frac{Q_{t+1}}{Q_t}\left\{\frac{\partial y_{t+1}^f}{\partial h_{t+1}^s} + \lambda_{t+1}^s \frac{\partial h_{t+2}^s}{\partial h_{t+1}^s}\right\}, \tag{26}$$

state that the shadow price of skill acquisition at time t, λ_t^i is equal to the discounted value of the net benefits to skill accumulation, $\frac{\partial y_{t+1}^f}{\partial h_{t+1}^i} + \lambda_{t+1}^u \frac{\partial h_{t+2}^i}{\partial h_{t+1}^i}$, where $\frac{\partial y_{t+1}^f}{\partial h_{t+1}^i}$ is the increase in output due to increased labour skills at $t+1$ and $\frac{\partial h_{t+2}^i}{\partial h_{t+1}^i}$ is the increased labour skills in $t+2$ that result from increased skills in $t+1$, valued by its corresponding shadow price in $t+1$, λ_{t+1}^i.

3.3 Government budget

To focus on policies to reduce training inequality, we assume the following balanced-budget constraint for the government:

$$T_t = \tau^u\left(\phi^u t_t^u l_t^{f,u}\right) + \tau^s\left(\phi^s t_t^s l_t^{f,s}\right), \tag{27}$$

which equates the lump-sum transfer/tax, T_t, with revenue/expenditure for the monetary costs of training time, $\phi^i t_t^i l_t^{f,i}$. To ensure that the government budget is balanced, T_t, is the residual policy instrument in the analysis below.

3.4 Market clearing conditions

The market clearing conditions for physical capital, unskilled and skilled labour, dividends and goods markets are respectively:

$$k_t^f = k_t, \tag{28}$$

$$l_t^{f,u} = n^u l_t^u, \tag{29}$$

$$l_t^{f,s} = n^s l_t^s, \tag{30}$$

$$\pi_t^f = \pi_t, \tag{31}$$

$$y_t^f = c_t + k_{t+1} - (1-\delta_k)\,k_t + \phi_u t_t^u l_t^{f,u} + \phi_s t_t^s l_t^{f,s}. \tag{32}$$

3.5 Decentralized Equilibrium

Given initial conditions, the decentralized equilibrium is defined to be an allocation $\left\{c_t, l_t^u, l_t^s, \pi_t, l_t^{f,u}, l_t^{f,s}, k_t^f, \pi_t^f, t_t^u, t_t^s, h_{t+1}^u, h_{t+1}^s\right\}_{t=0}^{\infty}$, prices $\{r_t, w_t^u, w_t^s\}_{t=0}^{\infty}$, shadow prices $\left\{\lambda_t^k, \lambda_t^u, \lambda_t^s\right\}_{t=0}^{\infty}$, and the policy instrument, $\{T_t\}_{t=0}^{\infty}$, such that (i) households and firms undertake their respective optimisation problems taking aggregate outcomes as given; (ii) all constraints are satisfied; and (iii) all markets clear.

Using Walras's law we discard the household's budget constraint, thus the DE consists of the following 19 equations: (i) the household's 4-FOCs, equations (6)-(9); (ii) the firm's 2-skill accumulation equations (11)-(12); (iii) the firm's 7-FOCs, equations (20)-(26); (iv) the government's budget constraint, equation (27); and (v) the 5-market clearing conditions, equations (28)-(32).

3.6 Model calibration and steady-state

We set the parameters appearing in the DE equations with the overall aim that the model generates a steady-state solution which implies model generated quantities similar to the actual data for the UK. The calibrated parameters are summarised in Table 3. More details on data sources used can be found in Appendix A.

The productivity parameters which work as scaling factors $\{A, H^u, H^s\}$ are all normalised to unity. Also, following many dynamic general equilibrium studies, we set the coefficient of relative risk aversion $\sigma = 2$. Similarly, we set the depreciation rate of capital, $\delta^k = 2.5\%$, which is commonly used in dynamic general equilibrium studies for the UK economy, see e.g. Harrison and Oomen (2010). Given that the depreciation of job-related skills is hard to measure, we assume $\delta^s = \delta^u = \delta^k$. The literature on work-related human capital, e.g. Blundell *et al.* (1999), suggests that this depreciates within a decade or so, which implies a yearly depreciation rate of about 10%. Indeed, Mincer and Ofek (1982) estimated the annual rates of individual-level depreciation to be between 3.3% and 7.6%, while Heckman (1976) reports a confidence interval between 3.7% and 8.9%. To these figures, one needs to add the value of human capital stock lost because of retirees, which, according to Stokey and Rebelo (1995), amounts to 2.5% up to 4% of the total stock. Based on this evidence, the quarterly depreciation rate should lie between 1.45% and 3.26%. Thus, our assumption of 2.5% is in-between these estimates. We set the quarterly discount factor of $\beta = 0.995$ to ensure that the annualized risk-free interest rate net of depreciation is equal to 2 percentage points in the steady-state. The latter is the value obtained from the real rate of discount on 3 month Treasury bills, net of inflation, averaged over the periods 1992q1-2015q4. Finally the population shares n^u and n^s are obtained from the QLFS dataset, and correspond to the average shares over the period 2000q1-2015q4.

Data are available for training subsidies to firms from CVTS 3 & 4. We divide these subsidies by training costs (average per firm, in a given year) and find that the subsidies amount, on average, to about 4.17% of firms' training costs. The CVTS dataset does not distinguish training subsidies for skilled workers separately from those for unskilled workers, and current fiscal policies do not discriminate between training recipients with respect to job-related training paid by companies. We thus set $\tau^u = \tau^s = 4.17\%$.

Table 3: Calibration

Symbol	Value	Definitions
		Household
β	0.995	quarterly time discount factor
ψ_1	0.320	consumption weight in utility
ψ_2	0.370	unskilled leisure weight in utility
ψ_3	0.310	skilled leisure weight in utility
δ^k	0.025	quarterly capital depreciation rate
σ	2.000	coefficient of relative risk aversion
n^s	0.340	share of unskilled to total household members
n^u	0.660	share of skilled to total household members
		Firm
ν	-0.495	effective skilled labour to capital substitution parameter
α	0.401	effective unskilled labour substitution parameter
ω	0.9416	elasticity of effective labour with respect to time
A	1.000	total factor productivity
δ^u	0.025	depreciation rate for accumulated skills (unskilled)
δ^s	0.025	depreciation rate for accumulated skills (skilled)
H^u	1.000	productivity of new skill creation (unskilled)
H^s	1.000	productivity of new skill creation (skilled)
μ	0.589	share of composite input to output
ρ	0.881	share of capital to the composite input
ϕ^u	3.234	fixed cost per training hour (unskilled)
ϕ^s	4.445	fixed cost per training hour (skilled)
γ^u	0.589	returns to scale for creating new skills (unskilled)
γ^s	0.622	returns to scale for creating new skills (skilled)
		Policy
τ^u	0.042	public subsidy for training activities (unskilled)
τ^s	0.042	public subsidy for training activities (skilled)

The parameters $\{\psi_2, \psi_3\}$ (recall that ψ_1 follows from $\psi_1 + \psi_2 + \psi_3 = 1$) are calibrated to match labour supply for skilled and unskilled workers. In particular, the QLFS database reports the average weekly hours of work of skilled and

of unskilled workers over the periods 1994.1-2015.4. We normalize these by the number of daytime hours (i.e. 16×7) in a week to calculate the labour supply of skilled and unskilled workers as 0.31 and 0.29, respectively. Conditional on the remaining parameters, $\{\psi_2, \psi_3\}$, are obtained from the labour supply conditions to ensure $l^s = 0.31$ and $l^u = 0.29$.

We next move to the group of parameters relating to training and production $\{\nu, \alpha, \mu, \rho, \phi^u, \phi^s, \gamma^u, \gamma^s, \omega\}$. We start with the parameter ω, which, as discussed in the previous section, is directly linked to firms' profitability or returns associated with job-related training and the resulting rents to firms. To the best of our knowledge, data on firms' returns, in terms of profitability, associated with firms' expenses on job-related training do not exist in the UK. Blundell *et al.* (1999) estimate the private return to participating to job-related training in the UK to be up to 10% and Dearden *et al.* (2006) estimate the partial effect of training time to firms' profits, alongside other factor inputs in a regression analysis. However, it is difficult to express such partial effects in model-relevant quantities. We thus choose ω by relating firm profitability to a monetary valuation of the investment in job-related training, as measured by the ratio of firm's profits over total monetary costs of training, including both direct and indirect costs, i.e. $\frac{\pi_t}{\phi^u t_t^u l_t^{f,u}(1-\tau^u)+\phi^s t_t^s l_t^{f,s}(1-\tau^s)+w_t^u t_t^u l_t^{f,u}+w_t^s t_t^s l_t^{f,s}}$. The advantage of using this ratio is that it is free of units of measurement, and thus useful for model calibration purposes. Almeida and Carneiro (2009) estimate this return to be between 8.6 and 13.8 percentage points for training firms in Portugal. Given this available information, we choose ω so that, in conjunction with the remaining parameters, firms' returns on investment in training, defined as above, are about 10%.

We also have data on the: (i) labour income share, $\frac{n^s l^s w^s + n^u l^u w^u}{y}$; (ii) capital-to-output ratio, $\frac{k}{y}$; (iii) skill premium, $\frac{w^s}{w^u}$; (iv) training costs as a percent of GDP, $\frac{\phi_s t^s l^s n^s + \phi_u t^u l^u n^u}{y}$; (v) unskilled training share, t^u; and (vi) skilled training share, t^s. These are obtained, respectively, from: (i) data from the OECD (2015) report; (ii) GDP and capital stock series published by the ONS; (iii) our own calculations from the UK QLFS data, averaging the ratio of the hourly wage of University educated workers and that of non-University educated workers over the period 1995q1-215q4; (iv) ONS data on gross value added (GVA) and the estimates of the total training costs reported in the 2011 ESS; (v) our own calculations, based on ESS estimates of total training time per employee, on QLFS population shares, and on the average ratio of training participation rate of University educated and that of non-University educated workers derived from the QLFS over the period 1995q1-2015q4; (vi) same as (v). Thus, these data provide six targets.

Following common practice in the literature using general equilibrium calibrated models with the CES production function (see e.g. Lindquist (2004) and Pourpourides (2011)), we set the elasticities of substitution $\nu = -0.495$ and

$\alpha = 0.401$, based on the estimates by Krusell *et al.* (2000). We then choose the remaining parameters in the production function so that the model's steady-state solution is consistent with factor income shares and inequality indices. In particular, we choose $\{\mu, \rho, \phi^u, \phi^s, \gamma^u, \gamma^s\}$ so that the model's steady-state predictions regarding $\left\{\frac{n^s l^s w^s + n^u l^u w^u}{y}, \frac{k}{y}, \frac{w^s}{w^u}, \frac{\phi_s t^s l^s n^s + \phi_u t^u l^u n^u}{y}, t^u, t^s\right\}$ are similar to the data.

The steady-state solution implied by the parameters in Table 3 is summarised below in Table 4. As can be seen, the model's predictions for the long-run quantities are very close to the data. Moreover, we can use this steady-state to evaluate the predictions of the model regarding the elasticity of training (average across the two types of workers) with respect to changes in subsidies. Recall that the empirical evidence in Section 2 demonstrates a significant, but small effect of an increase in subsidies on training shares, i.e. 0.26% with a 95% confidence interval implying a range from 0.05 to 0.47. Given that we cannot differentiate between skilled and unskilled workers in the data, this estimate refers to an average effect, across worker types. Thus, we examine the response of the model solution to increasing both τ^u and τ^s by 1%, starting from the solution in Table 4, and find that on average, across skilled and unskilled workers, training increases by 0.03%. Thus, the elasticity of training time with respect to training subsidies that is implied by the model is fairly consistent with the empirical evidence, and it lies just below the lower bound of our estimation in Section 2.

Table 4: Steady-State

Variable	Definition	Model	Data
$\frac{w^s}{w^u}$	skill premium	1.675	1.671
t^s	skilled training to total time share	0.023	0.023
t^u	unskilled training to total time share	0.013	0.013
t^s/t^u	training differential	1.743	1.746
$\frac{t^s l^{f,s} + t^u l^{f,u}}{(1-t^s) l^{f,s} + (1-t^u) l^{f,u}}$	training to work time share	0.019	0.017
l^s	skilled labour to total hours	0.316	0.310
l^u	unskilled labour to total hours	0.292	0.290
k/y	capital-to-output	10.25	10.30
$\frac{\phi_s t^s l^s n^s + \phi_u t^u l^u n^u}{y}$	monetary training costs-to-output	0.025	0.025
T/y	public spending on training costs-to-output	0.0006	0.0006
rk/y	capital income-to-output	0.306	0.285
$\frac{n^s l^s w^s + n^u l^u w^u}{y}$	labour income-to-output	0.665	0.685

4 Policy analysis

We next examine the dynamic effects of training subsidies on training, wages and earnings. To solve for the transition paths, we work as follows. We assume that

the economy is at its steady-state, as summarised by Table 4, when a one-off permanent change takes place in either τ^u or τ^s. We then solve for the dynamic paths of the endogenous variables of the system as this moves towards the new steady-state by obtaining the dynamic solution of the non-linear DGE system of equations for T periods, which is solved non-linearly using standard numeric methods in *Dynare* (see Adjemian *et al.* (2011)). We set $S = 1000$ to ensure that convergence is achieved.

4.1 Income and inequality effects

The effects of a permanent increase in τ^u from 0.042 to 0.5, which implies that the government subsidises half of the training cost for unskilled workers, are shown in Figure 2. As can be seen, an increase in τ^u increases training time for unskilled workers (see subplot (1,2) for t_t^u). As expected, given that the cost to train workers is lower for the firm, this creates additional incentives for the firm to increase t_t^u.

[Figure 2 here]

To see the effect of τ^u on training time more formally, recall the first-order condition (23), which is re-written as:

$$\phi^u l_t^{f,u} (1 - \tau^u) = \lambda_t^u \frac{\partial h_{t+1}^u}{\partial t_t^u} - \frac{\partial y_t^f}{\partial t_t^u}. \tag{33}$$

This implies that a reduction in the training costs requires that the right-hand side of (33) must also fall. Given the concavity of the skill creation function (11), $\frac{\partial h_{t+1}^u}{\partial t_t^u}$ is a decreasing function of t_t^u, hence a rise in τ^u tends to create a rise in t_t^u via its effect on future labour productivity. However, $\frac{\partial y_t^f}{\partial t_t^u}$ is also a decreasing function of t_t^u,[23] hence the rise in τ^u also tends to generate a fall in t_t^u via its direct effect on production, capturing the opportunity cost of time taken away from production. Quantitatively, the effects associated with $\frac{\partial h_{t+1}^u}{\partial t_t^u}$ in this case dominate, but they are mediated by the opportunity cost effects via $\frac{\partial y_t^f}{\partial t_t^u}$.

The increase in t_t^u leads to higher worker skills (see subplot (2,1) for h_t^u). In turn, the increased labour productivity works to increase labour demand, since the

[23] It can be shown from the production function (10) that $\frac{\partial \widetilde{y}_t^f}{\partial z_t^u} > 0$, and from equation (14) that $\frac{\partial z_t^u}{\partial t_t^u} < 0$, $\frac{\partial^2 z_t^u}{\partial^2 t_t^u} < 0$. Note then that $\frac{\partial y_t^f}{\partial t_t^u} = \frac{\partial \widetilde{y}_t^f}{\partial z_t^u} \frac{\partial z_t^u}{\partial t_t^u} < 0$ and that $\frac{\partial^2 y_t^f}{\partial^2 t_t^u} = \frac{\partial \left[\frac{\partial \widetilde{y}_t^f}{\partial z_t^u} \frac{\partial z_t^u}{\partial t_t^u} \right]}{\partial t_t^u} = \frac{\partial \widetilde{y}_t^f}{\partial z_t^u \partial t_t^u} \frac{\partial z_t^u}{\partial t_t^u} + \frac{\partial \widetilde{y}_t^f}{\partial z_t^u} \frac{\partial^2 z_t^u}{\partial^2 t_t^u} = \frac{\partial \widetilde{y}_t^f}{\partial z_t^u} \frac{\partial^2 z_t^u}{\partial^2 t_t^u} < 0$.

marginal product of labour, $\frac{\partial y_t^f}{\partial l_t^{f,u}}$, is an increasing function of h_t^u.[24] At the same time, the increased training also tends to reduce $l_t^u(1-t_t^u)$, which puts pressure on firms to increase the unskilled labour input, $l_t^{f,u}$, to accommodate the higher training without reducing drastically hours used in production (see subplot (2,3) for $l_t^u(1-t_t^u)$). On the other hand, the increase in the unskilled labour input tends to decrease the marginal product of unskilled labour (see subplot (5,1)), since the marginal product of labour, $\frac{\partial y_t^f}{\partial l_t^{f,u}}$, is a decreasing function of $l_t^{f,u}$.[25] The trade-off is resolved by an increase in $l_t^{f,u}$ (see subplot (4,3)) and in unskilled wages (see subplot (3,3)), despite the initial reduction in the marginal product of unskilled labour. Recall from the first-order condition:

$$w_t^u + \phi^u t_t^u (1-\tau^u) = \frac{\partial y_t^f}{\partial l_t^{f,u}} + \lambda_t^u \frac{\partial h_{t+1}^u}{\partial l_t^{f,u}}, \tag{34}$$

that wages increase with the marginal product of labour, $\frac{\partial y_t^f}{\partial l_t^{f,u}}$, and with $\frac{\partial h_{t+1}^u}{\partial l_t^{f,u}}$, which is a decreasing function of $l_t^{f,u}$, given the concavity of the skill creation function, but also increase when training costs decrease. Here, although the rise in $l_t^{f,u}$ tends to decrease the right-hand side of the first-order condition, the reduction in training costs dominates quantitatively, so that wages increase. Consequently, earnings for unskilled workers, $w_t^u l_t^u$ increase (see subplot (5,3)), since both labour input and wages increase.

The positive developments in the labour market for unskilled labour, and, in particular, the increase in the effective labour input of the unskilled (see the increase in z_t^u in subplot (3,1)), have positive spillover effects on the productivity and returns to skilled labour. In particular, after an initial decline, the marginal product of skilled labour and thus skilled wages increase (see subplots (5,2) and (3,4) respectively). Following these dynamics, capital stock is also increasing (see subplot (4,2)).[26] Hence, the increased labour productivity and employment for

[24] It can be shown from the production function (10) that $\frac{\partial \widetilde{y}_t^f}{\partial z_t^u} > 0$, and from equation (14) that $\frac{\partial z_t^u}{\partial l_t^{f,u}} > 0$, $\frac{\partial z_t^u}{\partial l_t^{f,u} \partial h_t^u} > 0$. Note then that $\frac{\partial y_t^f}{\partial l_t^{f,u}} = \frac{\partial \widetilde{y}_t^f}{\partial z_t^u} \frac{\partial z_t^u}{\partial l_t^{f,u}} > 0$ and that $\frac{\partial y_t^f}{\partial l_t^{f,u} \partial h_t^u} = \frac{\partial \left[\frac{\partial \widetilde{y}_t^f}{\partial z_t^u} \frac{\partial z_t^u}{\partial l_t^{f,u}} \right]}{\partial h_t^u} = \frac{\partial \widetilde{y}_t^f}{\partial z_t^u \partial h_t^u} \frac{\partial z_t^u}{\partial l_t^{f,u}} + \frac{\partial \widetilde{y}_t^f}{\partial z_t^u} \frac{\partial z_t^u}{\partial l_t^{f,u} \partial h_t^u} = \frac{\partial \widetilde{y}_t^f}{\partial z_t^u} \frac{\partial z_t^u}{\partial l_t^{f,u} \partial h_t^u} > 0$.

[25] It can be shown from the production function (10) that $\frac{\partial \widetilde{y}_t^f}{\partial z_t^u} > 0$, and from equation (14) that $\frac{\partial z_t^u}{\partial l_t^{f,u}} > 0$, $\frac{\partial^2 z_t^u}{\partial^2 l_t^{f,u}} < 0$. Note then that $\frac{\partial y_t^f}{\partial l_t^{f,u}} = \frac{\partial \widetilde{y}_t^f}{\partial z_t^u} \frac{\partial z_t^u}{\partial l_t^{f,u}} > 0$ and that $\frac{\partial^2 y_t^f}{\partial^2 l_t^{f,u}} = \frac{\partial \left[\frac{\partial \widetilde{y}_t^f}{\partial z_t^u} \frac{\partial z_t^u}{\partial l_t^{f,u}} \right]}{\partial l_t^{f,u}} = \frac{\partial \widetilde{y}_t^f}{\partial z_t^u \partial l_t^{f,u}} \frac{\partial z_t^u}{\partial l_t^{f,u}} + \frac{\partial \widetilde{y}_t^f}{\partial z_t^u} \frac{\partial^2 z_t^u}{\partial^2 l_t^{f,u}} = \frac{\partial \widetilde{y}_t^f}{\partial z_t^u} \frac{\partial^2 z_t^u}{\partial^2 l_t^{f,u}} < 0$.

[26] To see that $\frac{\partial y_t^f}{\partial l_t^{f,f}}$ is increasing in both z_t^u and k_t^f, note the following. It can be shown

unskilled workers initially crowds out capital, skilled training and skilled labour productivity (see subplots (1,3) and (2,2) for t_t^s and h_t^s, respectively). However, as more resources are diverted towards unskilled labour during the initial phase of the adjustment towards the new steady-state, the increased effective unskilled labour input eventually crowds in capital and skilled training as well as skilled hours (see subplot (4,4) for $l_t^{f,s}$). The changes in wages and hours imply that earnings are also increased (see subplot (5,4) for $w_t^s l_t^s$).

In summary, increased training subsidies for unskilled workers create benefits to both skilled and unskilled workers, in terms of wages and earnings. The effect is stronger for unskilled workers, so that wage inequality is reduced. Hence, this is a policy intervention which, in terms of labour income, is beneficial for all the agents and it reduces inequality.

The same dynamics can be observed in the case that the government increases subsidies to skilled training. In this case, the spillovers come from the positive developments in the labour market for skilled labour, and produce an improvement of working conditions of unskilled workers. Figure 3 shows the effects of a permanent increase in τ^s from 0.042 to 0.5, which implies that the government subsidises half of the training costs for unskilled workers. As can be seen, although increasing τ^s is Pareto improving in terms of labour income, it increases inequality.

[Figure 3 here]

In Table 5, we summarise the effects of different increases in τ^u and τ^s on training, wages and earnings for both types of workers, as well as on the relevant measures of inequality. For each tax instrument, we consider three different magnitudes of training subsidies, in particular $\tau^i = 0.25$, $\tau^i = 0.5$ (which was analysed in Figures 2 and 3), and $\tau^i = 1$ for $i = u, s$. The increase in τ^u to 0.25, increases training for unskilled workers by nearly 18%, implying an increase from 3.4 days of average training to 4. Similarly, the increase in τ^u to 0.5 implies an increase from 3.4 days of average training to 5 days.[27] In terms of implied elasticities, these effects suggest that an increase by 1% in τ^u increases training for the unskilled

from the production function (10) that $\frac{\partial \widetilde{y}_t^f}{\partial z_t^s} > 0$, $\frac{\partial \widetilde{y}_t^f}{\partial z_t^s \partial z_t^u} > 0$, $\frac{\partial \widetilde{y}_t^f}{\partial z_t^s \partial k_t^f} > 0$, and from equation (14) that $\frac{\partial z_t^u}{\partial l_t^{f,s}} > 0$. Note then that $\frac{\partial y_t^f}{\partial l_t^{f,s}} = \frac{\partial \widetilde{y}_t^f}{\partial z_t^s} \frac{\partial z_t^s}{\partial l_t^{f,s}} > 0$ and that $\frac{\partial y_t^f}{\partial l_t^{f,s} \partial z_t^u} = \frac{\partial \left[\frac{\partial \widetilde{y}_t^f}{\partial z_t^z} \frac{\partial z_t^s}{\partial l_t^{f,s}} \right]}{\partial z_t^u} = \frac{\partial \widetilde{y}_t^f}{\partial z_t^z \partial z_t^u} \frac{\partial z_t^s}{\partial l_t^{f,s}} + \frac{\partial \widetilde{y}_t^f}{\partial z_t^z} \frac{\partial z_t^s}{\partial l_t^{f,s} \partial z_t^u} = \frac{\partial \widetilde{y}_t^f}{\partial z_t^s \partial z_t^u} \frac{\partial z_t^s}{\partial l_t^{f,s}} > 0$,

$\frac{\partial y_t^f}{\partial l_t^{f,s} \partial k_t^f} = \frac{\partial \left[\frac{\partial \widetilde{y}_t^f}{\partial z_t^z} \frac{\partial z_t^s}{\partial l_t^{f,s}} \right]}{\partial k_t^f} = \frac{\partial \widetilde{y}_t^f}{\partial z_t^z \partial k_t^f} \frac{\partial z_t^s}{\partial l_t^{f,s}} + \frac{\partial \widetilde{y}_t^f}{\partial z_t^z} \frac{\partial z_t^s}{\partial l_t^{f,s} \partial k_t^f} = \frac{\partial \widetilde{y}_t^f}{\partial z_t^z \partial k_t^f} \frac{\partial z_t^s}{\partial l_t^{f,s}} > 0.$

[27] Using ESS and QLFS data, we estimate average unskilled training time to be 3.4 days per year, by combining the information about the average days of training per worker, the population share of skilled and unskilled workers, and the ratio of unskilled training participation to skilled training participation rate.

workers by 0.02%, which is at the lower bound of the estimates in Table 3. Hence, although consistent with the lower values of the empirical estimate for the effect of financial incentives on training, training subsidies produce sizeable increases in training.

In turn, these lead to smaller, but quantitatively significant increases in wages for the unskilled. The effect of the increase in training on wages is also consistent with previous econometric evidence (see e.g. Table 2 in Blundell *et al.* (1999)). In particular, we find in Table 5 (using the case for $\tau^u = 0.5$ as an illustration) that an increase in training by about 1.6 days increase wages by 1.45%. Since the average days of training in a year are 3.4, this implies that the incidence of training for a worker increases her wage by about 3.1%, which is consistent with the estimates in Blundell *et al.* (1999) regarding the effect of employer-provided training courses on the wages for a worker who undertook training in the year.

The effects of job-related training subsidies for unskilled workers on wage inequality reduction are smaller, because of the concurrent increase in wages for the skilled. Earnings inequality is reduced by more, because of the positive effects of increased training on unskilled hours. The relationship between wage inequality and training inequality in Table 5 is also consistent with the empirical estimates in Table 2. In particular, the results in Table 5 imply that a decrease in training inequality by 1% leads to a fall in wage inequality by about 0.011%, which is at the lower bound of the confidence interval for $\widehat{\beta}_2$ from Table 2.

Table 5: Steady-state effects of increasing the training subsidies

	Permanent increase in τ^u				Permanent increase in τ^s		
$\tau^u=$	0.25	0.5	1	$\tau^s=$	0.25	0.5	1
$\%\Delta t^u$	17.85	49.58	214.9		0.41	0.99	2.60
$\%\Delta t^s$	0.27	0.65	1.81		16.54	44.88	172.37
$\%\Delta \frac{t^s}{t^u}$	-14.92	-32.71	-67.67		16.06	43.46	165.47
$\%\Delta w^u$	0.58	1.45	4.39		0.61	1.46	3.85
$\%\Delta w^s$	0.42	1.04	2.89		0.89	2.20	6.40
$\%\Delta \frac{w^s}{w^u}$	-0.16	-0.41	-1.44		0.28	0.73	2.46
$\%\Delta w^u l^u$	1.06	2.76	9.83		0.64	1.61	5.11
$\%\Delta w^s l^s$	0.51	1.30	4.49		1.54	3.95	13.05
$\%\Delta \frac{w^s l^s}{w^u l^u}$	-0.55	-1.42	-4.86		0.9	2.30	7.56

The effect of a subsidy to skilled training on training time is slightly lower than that of subsidies to unskilled training. However, the effect on wages is larger. This can mainly be attributed to the skill-capital complementarity that allows a greater increase in overall labour productivity. Even though the policy produces higher inequality, we observe important spillovers especially with respect to the unskilled wage.

The message from Table 5 is that while the results are on the conservative side of the estimates regarding the effects of training subsidies on training and wage inequality, they nevertheless imply significant gains in terms of wages and income for unskilled workers. Therefore, although subsiding job-related training may not be the most effective policy tool in reducing inequality, it has strong potential to support the income of the lower skilled. In the next sub-section, we explore further the resource effectiveness of these income gains.

4.2 Multiplier analysis

We next measure the effectiveness of job-related training subsidies with respect to increases in income and other monetary values quantities, compared to the resources required for their funding. To do so, we compute multipliers based on the simulation exercise described above. In particular, we define the impact multiplier for the variable x as the difference between x_t and its value in the initial steady-state x, divided by the difference in government spending at time t and its initial steady-state, $T_t - T$, which is the period in which the new fiscal policy is introduced. Similarly, and following the large strand of literature on fiscal policy evaluation (see e.g. Leeper *et al.* (2010)), we compute the lifetime multiplier of e.g. the variable x according to the formula:

$$\text{lifetime multiplier} = \frac{\sum_{t=0}^{S} Q_t \left(x_t - x\right)}{\sum_{t=0}^{S} Q_t \left(T_t - T\right)} \tag{35}$$

where Q_t is the household discount factor introduced in (16). We simulate $S = 2000$ periods after the shock to compute (35). The multipliers for the case of subsidies to unskilled training are reported in Table 6.

Table 6: Multiplier effects of increasing the unskilled training subsidy

	Permanent increase in τ^u							
		impact multiplier				lifetime multiplier		
income measures	$\tau^u=$	0.25	0.5	1	$\tau^u=$	0.25	0.5	1
$w^u l^u$		0.75	0.72	0.66		1.74	1.60	1.25
$w^s l^s$		0.09	0.10	0.10		1.23	1.12	0.82
$w^u l^u + w^s l^s$		0.32	0.31	0.29		1.40	1.28	0.97
$(1 + r - \delta_k)k$		-0.19	-0.14	-0.07		7.99	7.46	5.93
y		0.24	0.24	0.24		1.62	1.46	1.07

As can be seen, all multipliers (except on impact for capital income) are positive and the lifetime multipliers are generally greater than one. Therefore, over the

lifetime, the increase in benefits is greater than the increase in resources required to finance the policy. As expected, given the dynamic analysis in Figure 2, and since the benefits increase over time, the lifetime multipliers are greater than the impact multipliers, but it is noteworthy that the benefits materialise even in the short-run. It is also interesting to note that the multipliers are decreasing with interventions that have larger fiscal implications, which implies decreasing returns on income from the increase in training that is induced by training subsidies.

In Table 7, we report the multiplier effects for permanent increases of training subsidies to skilled training, τ^s. The results are broadly similar to those in Table 6, although in general the positive effects are stronger at the aggregate level. This is explained by the central role of skilled labour in production, since its complementarity with capital acts as an amplification mechanism for the policy intervention at the aggregate level.

Table 7: Multiplier effects of increasing the skilled training subsidy

	Permanent increase in τ^s							
	impact multiplier				lifetime multiplier			
income measures	$\tau^s=$	0.25	0.5	1	$\tau^s=$	0.25	0.5	1
$w^u l^u$		0.32	0.29	0.22		0.76	0.69	0.52
$w^s l^s$		1.82	1.68	1.41		3.29	3.02	2.38
$w^u l^u + w^s l^s$		1.31	1.21	1.00		2.43	2.23	1.74
$(1 + r - \delta_k)k$		0.03	0.00	-0.08		25.6	22.3	14.6
y		0.40	0.36	0.27		2.14	1.90	1.32

5 Conclusions

To understand whether subsides to job-related training could improve earnings for the lower skilled workers and reduce wage inequality, as measured by the distance between wages and earnings of the skilled and unskilled workers, we developed a dynamic general equilibrium model for the UK. This model, incorporating skilled and unskilled labour, capital-skill complementarity in production and an endogenous training allocation, performed well with respect to replicating key long-term relationships in the UK data.

Our quantitative policy analysis suggested that training subsidies for the unskilled have a significant impact on their labour income. These subsidies also increase earnings for skilled workers and raise aggregate income with implied lifetime multipliers exceeding unity. The latter result implies that the increase in benefits accruing from the policy is greater than the increase in resources required to finance the policy. It should be noted, however, that a given increase in resources to finance training subsidies can have additional cost implications for the society depending on the type of revenue-generating policy implemented.

Training subsidies to skilled workers, while again increasing skilled and unskilled earnings, raise the former by more and worsen wage inequality. Therefore, there is a trade-off associated with subsidies to skilled training. In contrast, training subsidies to unskilled workers improve earnings for both skilled and unskilled workers without a negative impact on inequality.

The positive spillover effects to skilled workers imply that the effects of training subsidies on inequality are small. As a result, training subsidies that are targeted to unskilled workers are not a very effective tool for reducing inequality. However, this finding is a consequence of the effectiveness of the policy to propagate benefits throughout the labour force and thus works to increase the social value of such interventions.

References

[1] Acemoglu, D., and D. Autor (2011). 'Skills, tasks and technologies: implications for employment and earnings'. In: Ashenfelter, O., and D. Card (eds.). Handbook of Labor Economics, vol. 4B, chapter 12, pp. 1043–1171. Elsevier Science B.V., Amsterdam.

[2] Angelopoulos, K., Lazarakis, S., and J. Malley (2017a). 'Asymmetries in earnings, employment and wage risk in Great Britain'. CESifo Working Paper No. 6400.

[3] Angelopoulos, K., Lazarakis, S., and J. Malley (2017b). 'Wealth inequality and externalities from ex-ante skill heterogeneity'. CESifo Working Paper No. 6572.

[4] Adjemian, S., Bastani, H., Juillard, M., Mihoubi, F., Perendia, G., Ratto, M., and S. Villemot (2011). 'Dynare: reference manual'. Version 4.

[5] Almeida, R., and R. Aterido (2015). 'Investing in formal on-the-job training: are SMEs lagging much behind?'. *IZA Journal of Labor and Development*, 4, 1–23.

[6] Almeida, R., and P. Carneiro (2009). 'The return to firm investments in human capital'. *Labour Economics*, 16, 97–106.

[7] Almeida, R., and M. Faria (2014). 'The wage returns to on-the-job training: evidence from matched employer-employee data'. *IZA Journal of Labor and Development,* 3, 1–33.

[8] Autor, D. (2014). 'Skills, education, and the rise of earnings inequality among the other 99 percent'. *Science*, 344, 843–851.

[9] Bassanini, A., Booth, A., Brunello, G., Paola, M., and E. Leuven (2007). 'Workplace training in Europe'. In: Brunello, G., Garibaldi, P., and E. Wasmer (eds). Education and Training in Europe. Oxford University Press, Oxford, 187–231.

[10] Becker, G. (1962). 'Investment in human capital: a theoretical analysis'. *Journal of Political Economy* 70, 9–49.

[11] Belfield, C., Crawford, C., and L. Sibieta (2017). 'Long run comparisons of spending per pupil across different stages of education'. IFS report No. 126, Institute for Fiscal Studies, London, 1–45.

[12] Ben-Porath, Y. (1967). 'The production of human capital and the life cycle of earnings'. *Journal of Political Economy,* 75, 352–365.

[13] Bilanakos, C., Green, C. P., Heywood, J. S., and N. Theodoropoulos (2017). 'Do dominant firms provide more training?'. *Journal of Economics and Management Strategy*, 26, 67–95

[14] Blundell, R., Dearden, L., and C. Meghir (1996). 'The determinants and effects of work-related training in Britain'. IFS report R50, Institute for Fiscal Studies, London, 1–101.

[15] Blundell, R., Dearden, L., Meghir, C., and B. Sianesi (1999). 'Human capital investment: the returns from education and training to the individual, the firm and the economy'. *Financial Studies*, 20, 1–23.

[16] Blundell, R., and B. Etheridge (2010). 'Consumption, income and earnings inequality in Britain'. *Review of Economic Dynamics*, 13, 76–102.

[17] Brewer, M., and L. Wren-Lewis (2016). 'Accounting for changes in income inequality: decomposition analyses for the UK, 1978–2008'. *Oxford Bulletin of Economics and Statistics*, 78, 289–322.

[18] Brunello, G., and F. Gambarotto (2004). 'Agglomeration effects on employer-provided training: evidence from the UK'. CESifo Working Paper No. 1055.

[19] Brunello, G., and M. de Paola (2009). 'Is there under-provision of training?'. *Empirical Research in Vocational Education and Training*, 1, 1–18.

[20] Busemeyer, M. R. (2014). 'Skills and inequality'. Cambridge University Press, Cambridge.

[21] Caselli, F., and W. Coleman (2006). 'The world technology frontier'. *American Economic Review,* 96, 499–522.

[22] Chen, K., and C. Lai (2015). 'On-the-job learning and news-driven business cycles'. *Journal of Money, Credit and Banking,* 47, 261–294.

[23] De Grip, A., and J. Sauermann (2013). 'The effect of training on productivity: the transfer of on-the-job training from the perspective of economics'. *Educational Research Review*, 8, 28–36.

[24] Dearden, L., Reed, H., and J. Van Reenen (2006). 'The impact of training on productivity and wages: evidence from British panel data'. *Oxford Bulletin of Economics and Statistics*, 68, 397–421.

[25] Dostie, B. (2015). 'Do train-or-pay schemes really increase training levels?'. *Industrial Relations: A Journal of Economy and Society*, 54, 240–255.

[26] Felstead, A., Green, F., Jewson, N., and P. Casey (2013). 'Training in recession: the impact of the 2008-2009 recession on training at work'. UKCES Evidence Report No. 72, UK Commission for Employment and Skills.

[27] Goldin, C., and L. Katz (2008). 'Transitions: career and family life cycles of the educational elite'. *American Economic Review*, 98, 363–369.

[28] Greenhalgh, C. (2002). 'Does an employer training levy work? The incidence of and returns to adult vocational training in France and Britain'. *Fiscal Studies*, 23, 223–263.

[29] Haelermans, C., and L. Borghans (2012). 'Wage effects of on-the-job training: a meta-analysis'. *British Journal of Industrial Relations*, 50, 502–528.

[30] Hansson, B. (2008). 'Job-related training and benefits for individuals: a review of evidence and explanations'. *Education*, 19, 1–49.

[31] Harrison, R., and Ö. Oomen (2010). 'Evaluating and estimating a DSGE model for the United Kingdom'. Working Paper No. 380, Bank of England, 1–64.

[32] Heathcote, J., Storesletten, K., and G. Violante (2010). 'The macroeconomic implications of rising wage inequality in the United States'. *Journal of Political Economy*, 118, 681–722.

[33] Heckman, J. (1976). 'A life-cycle model of earnings, learning, and consumption'. *Journal of Political Economy*, 84, 9–44.

[34] Hornstein, A., Krusell, P., and G. Violante (2005). 'The effects of technical change on labor market inequalities'. In: Durlauf, S., and P. Aghion (eds.). Handbook of Economic Growth, vol. 1, chapter 20, pp. 1275–1370, Elsevier, Amsterdam.

[35] Huggett, M., Ventura, G., and A. Yaron (2006). 'Human capital and earnings distribution dynamics'. *Journal of Monetary Economics,* 53, 265–290.

[36] Kim, D., and C. Lee (2007). 'On-the-job human capital accumulation in a real business cycle model: implications for intertemporal substitution elasticity and labor hoarding'. *Review of Economic Dynamics,* 10, 494–518.

[37] Konings, J. and S. Vanormelingen (2015). 'The impact of training on productivity and wages: firm-level evidence'. *Review of Economics and Statistics*, 97, 485–497.

[38] Krusell, P., Lee, E., Ríos-Rull, J., and G. Violante (2000). 'Capital-skill complementarity and inequality: a macroeconomic analysis'. *Econometrica*, 68, 1029–1053.

[39] Leeper, E., Walker, T., and S. Yang (2010). 'Government investment and fiscal stimulus'. *Journal of Monetary Economics*, 57, 1000–1012.

[40] Leuven, E. (2004). 'A review of the wage returns to private sector training'. Proceedings of EC-OECD seminar on human capital and labour market performance, Brussels, December 8.

[41] Lindquist, M. (2004). 'Capital-skill complementarity and inequality over the business cycle'. *Review of Economic Dynamics*, 7, 519–540.

[42] Machin, S., and A. Vignoles (2001). 'The economic benefits of training to the individual, the firm and the economy: the key issues'. Cabinet Office, London.

[43] Machin, S., and A. Vignoles (2005). 'What's the good of education? The economics of education in the UK'. Princeton University Press, Princeton.

[44] Markowitsch, J., Käpplinger, B. and G. Hefler (2013). 'Firm-provided training in Europe and the limits of national skills strategies'. *European Journal of Education*, 48, 281–291.

[45] Mejía, D., and M. St-Pierre (2008). 'Unequal opportunities and human capital formation'. *Journal of Development Economics*, 86, 395–413.

[46] Manuelli, R. E., and A. Seshadri (2014). 'Human capital and the wealth of nations'. *American Economic Review*, 104, 2736–2762.

[47] Méndez, F., and F. Sepúlveda (2016). 'A comparative study of training in the private and public sectors: evidence from the United Kingdom and the United States'. IREF working papers No. 2013-06, Institute for Research in Economic and Fiscal Issues, 1–33.

[48] Merz, M. (1995). 'Search in the labor market and the real business cycle'. *Journal of Monetary Economics*, 36, 269–300.

[49] Mincer, J. (1992), 'Human capital: a review'. Discussion Paper Series No. 591, Columbia University, New York.

[50] Mincer, J., and H. Ofek (1982). 'Interrupted work careers: depreciation and restoration of human capital'. *Journal of Human Resources*, 1, 3–24.

[51] Moreno-Galbis, E. (2012). 'The impact of TFP growth on the unemployment rate: does on-the-job training matter?'. *European Economic Review*, 56, 1692–1713.

[52] O'Mahony, M. (2012). 'Human capital formation and continuous training: evidence for EU countries'. *Review of Income and Wealth*, 58, 531–549.

[53] Parent, D. (1999). 'Wages and mobility: the impact of employer-provided training'. *Journal of Labor Economics,* 17, 298–317.

[54] Pourpourides, P. (2011). 'Implicit contracts and the cyclicality of the skill-premium'. *Journal of Economic Dynamics and Control*, 35, 963–979.

[55] Rogerson, R., and R. Shimer (2011). 'Search in macroeconomic models of the labor market'. Handbook of Labor Economics, 4, 619–700.

[56] Sofer, C. (2004). 'Human capital over the life cycle: a European perspective'. Edward Elgar Publishing, Cheltenham.

[57] Stevens, M. (1999). 'Human capital theory and UK vocational training policy'. *Oxford Review of Economic Policy*, 15, 16–32.

[58] Stokey, N. L., and S. Rebelo (1995). 'Growth effects of flat-rate taxes'. *Journal of Political Economy*, 103, 519–550.

[59] Wößmann, L., and G. Schütz (2006). 'Efficiency and equity in European education and training systems'. Analytical report for the European Commission prepared by the European Expert Network on Economics of Education (EENEE), Bruxelles.

Appendix A: Data Description

variable	source	frequency	description
ave. wage, skilled workers	QLFS	1995.1-2015.4	NSA, nominal
ave. wage, unskilled workers	QLFS	1995.1-2015.4	NSA, nominal
training particip. rate, total	QLFS	1994.3-2015.4	NSA, share of workers who trained in last 13 wks to all workers
training particip. rate, skilled workers	QLFS	1994.3-2015.4	NSA, share of workers who trained in last 13 wks to all workers
training particip. rate, unskilled workers	QLFS	1994.3-2015.4	NSA, share of workers who trained in last 13 wks to all workers
ave. weekly hours, skilled workers	QLFS	1994.1-2015.4	NSA
ave. weekly hours, unskilled workers	QLFS	1994.1-2015.4	NSA
ave. weekly hours, total	QLFS	1994.1-2015.4	NSA
total number of skilled workers	QLFS	1994.1-2015.4	NSA
total number of unskilled workers	QLFS	1994.1-2015.4	NSA
per firm training receipts (subsidies)	CVTS 3 & 4	2005 & 2010	nominal, ave. by SIC sector in 2005 & 2010
per firm training costs	CVTS 3 & 4	2005 & 2010	nominal, ave. by SIC sector in 2005 & 2010
ave. firm size (number of employees)	CVTS 3 & 4	2005 & 2010	ave. number of employees by SIC sector in 2005 & 2010
ave. number of days of training	CVTS 4	2010	per worker; see Table 5.2, p. 92 of the CVTS report (2010)
gross value added	ONS	2010	SA, nominal
gross capital stock	ONS	1997-2015	SA, nominal
GDP deflator	ONS	1992.1-2015.4	1990=100
real consumption expenditure	ONS	1992.1-2015.4	SA, 1990 prices
real gross domestic product	ONS	1992.1-2015.4	SA, 1990 prices
real gross fixed capital formation	ONS	1997.1-2015.4	SA, 1990 prices
real interest rate	BOE	1992.1-2015.4	quarterly real rate of discount, 3 month Treasury bills

Note that all QLFS data is based on employees and self-employed who are between 25 and 65 years old.

Appendix B: Derivatives Firm's FOCs

The derivatives used in the FOCs of the firm in the main text are defined as follows:

$$\frac{\partial y_t^f}{\partial k_t^f} = \frac{A^\alpha \rho (k_t^f)^\nu (y_t^f)^{1-\alpha} (1-\mu)}{k_t^f [\rho (k_t^f)^\nu + ([l_t^{f,s}(1-t_t^s)]^\omega [h_t^s]^{1-\omega})(1-\rho)]^{1-\frac{\alpha}{\nu}}}, \tag{36}$$

$$\begin{aligned}\tfrac{\partial y_t^f}{\partial l_t^{f,u}} &= \mu\omega A\{\mu([l_t^{f,u}(1-t_t^u)]^\omega [h_t^u]^{1-\omega})^\alpha + (1-\mu)\times \\ &\times[\rho(k_t^f)^\nu + (1-\rho)([l_t^{f,s}(1-t_t^s)]^\omega [h_t^s]^{1-\omega})^\nu]^{\frac{\alpha}{\nu}}\}^{\frac{1}{\alpha}-1}\times \\ &\times([l_t^{f,u}(1-t_t^u)]^\omega [h_t^u]^{1-\omega})^\alpha (l_t^{f,u})^{-1},\end{aligned} \tag{37}$$

$$\frac{\partial h_{t+1}^u}{\partial l_t^{f,u}} = \gamma^u H^u (t_t^u h_t^u)^{\gamma^u} (l_t^{f,u})^{\gamma^u - 1}, \tag{38}$$

$$\begin{aligned}\tfrac{\partial y_t^f}{\partial l_t^{f,s}} &= \omega(1-\mu)(1-\rho)A\{\mu([l_t^{f,u}(1-t_t^u)]^\omega [h_t^u]^{1-\omega})^\alpha + (1-\mu)\times \\ &\times[\rho(k_t^f)^\nu + (1-\rho)([l_t^{f,s}(1-t_t^s)]^\omega [h_t^s]^{1-\omega})^\nu]^{\frac{\alpha}{\nu}}\}^{\frac{1}{\alpha}-1}[\rho(k_t^f)^\nu + (1-\rho)\times \\ &\times([l_t^{f,s}(1-t_t^s)]^\omega [h_t^s]^{1-\omega})^\nu]^{\frac{\alpha}{\nu}-1}([l_t^{f,s}(1-t_t^s)]^\omega [h_t^s]^{1-\omega})^\nu (l_t^{f,s})^{-1},\end{aligned} \tag{39}$$

$$\frac{\partial h_{t+1}^s}{\partial l_t^{f,s}} = \gamma^s H^s (t_t^s h_t^s)^{\gamma^s} (l_t^{f,s})^{\gamma^s - 1}, \tag{40}$$

$$\begin{aligned}\tfrac{\partial y_t^f}{\partial t_t^u} &= -\omega\mu A\{\mu([l_t^{f,u}(1-t_t^u)]^\omega [h_t^u]^{1-\omega})^\alpha + (1-\mu)\times \\ &\times[\rho(k_t)^\nu + (1-\rho)([l_t^{f,s}(1-t_t^s)]^\omega [h_t^s]^{1-\omega})^\nu]^{\frac{\alpha}{\nu}}\}^{\frac{1}{\alpha}-1}\times \\ &\times\tfrac{([l_t^{f,u}(1-t_t^u)]^\omega [h_t^u]^{1-\omega})^\alpha}{1-t_t^u},\end{aligned} \tag{41}$$

$$\frac{\partial h_{t+1}^u}{\partial t_t^u} = \gamma^u H^u (L_t^u h_t^u)^{\gamma^u} (t_t^u)^{\gamma^u - 1}, \tag{42}$$

$$\begin{aligned}\tfrac{\partial y_t^f}{\partial t_t^s} &= -(1-\rho)(1-\mu)\omega A\{\mu([l_t^{f,u}(1-t_t^u)]^\omega [h_t^u]^{1-\omega})^\alpha + (1-\mu)\times \\ &\times[\rho(k_t)^\nu + (1-\rho)([l_t^{f,s}(1-t_t^s)]^\omega [h_t^s]^{1-\omega})^\nu]^{\frac{\alpha}{\nu}}\}^{\frac{1}{\alpha}-1}[\rho(k_t)^\nu + (1-\rho)\times \\ &\times([l_t^{f,s}(1-t_t^s)]^\omega [h_t^s]^{1-\omega})^\nu]^{\frac{\alpha}{\nu}-1}([l_t^{f,s}(1-t_t^s)]^\omega [h_t^s]^{1-\omega})^\nu (l_t^{f,s})^{-1},\end{aligned} \tag{43}$$

$$\frac{\partial h_{t+1}^s}{\partial t_t^s} = \gamma^s H^s (l_t^{f,s} h_t^s)^{\gamma^s} (t_t^s)^{\gamma^s - 1}, \tag{44}$$

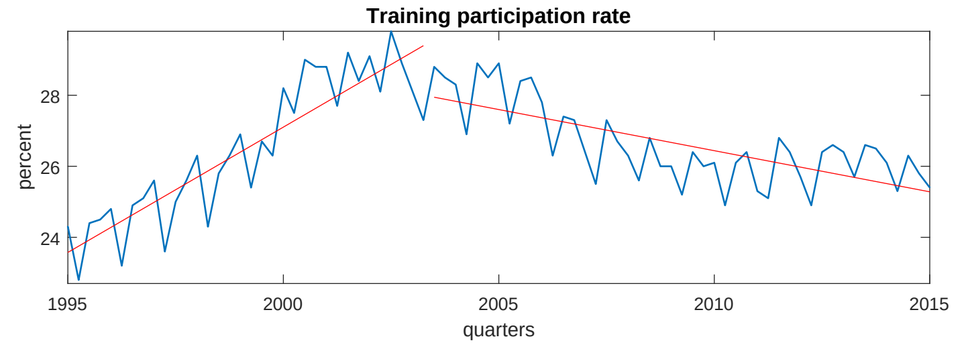
$$\begin{aligned}\tfrac{\partial y_{t+1}^f}{\partial h_{t+1}^u} &= A\mu(1-\omega)\{\mu([l_{t+1}^{f,u}(1-t_{t+1}^u)]^\omega [h_{t+1}^u]^{1-\omega})^\alpha + (1-\mu)\times \\ &\times[\rho(k_{t+1}^f)^\nu + (1-\rho)([l_{t+1}^{f,s}(1-t_{t+1}^s)]^\omega [h_{t+1}^s]^{1-\omega})^\nu]^{\frac{\alpha}{\nu}}\}^{\frac{1}{\alpha}}\times \\ &\times([l_{t+1}^{f,u}(1-t_{t+1}^u)]^\omega [h_{t+1}^u]^{1-\omega})^\alpha (h_{t+1}^u)^{-1},\end{aligned} \tag{45}$$

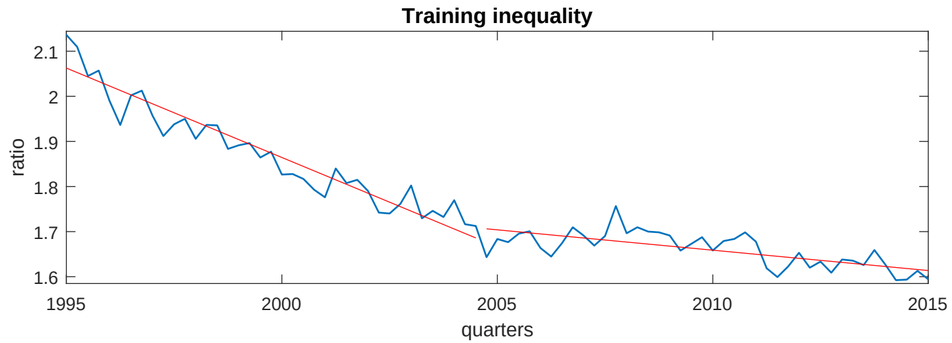
$$\frac{\partial h_{t+2}^{u}}{\partial h_{t+1}^{u}} = 1 - \delta^{u} + \gamma^{u} H^{u} \left(t_{t+1}^{u} l_{t+1}^{f,u} \right)^{\gamma^{u}} \left(h_{t+1}^{u} \right)^{\gamma^{u}-1}, \tag{46}$$

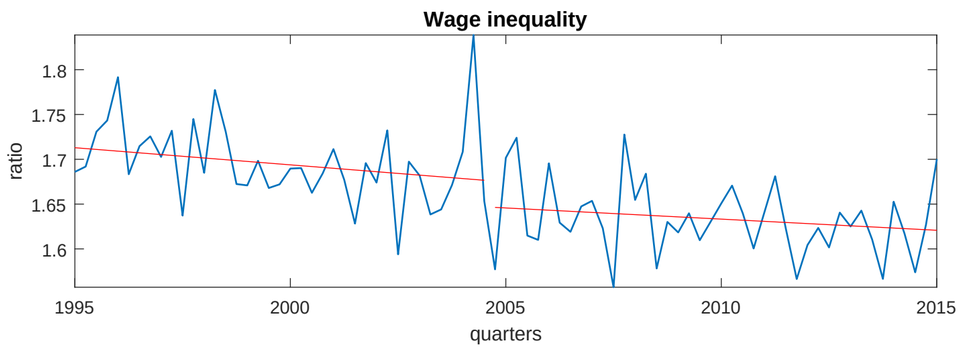
$$\begin{array}{l} \frac{\partial y_{t+1}^{f}}{\partial h_{t+1}^{s}} = A\left(1-\mu\right)\left(1-\rho\right)^{\nu}\left(1-\omega\right)\{\mu([l_{t+1}^{f,u}\left(1-t_{t+1}^{u}\right)]^{\omega}\left[h_{t+1}^{u}\right]^{1-\omega})^{\alpha}+ \\ +\left(1-\mu\right)[\rho(k_{t+1}^{f})^{\nu}+\left(1-\rho\right)\left([l_{t+1}^{f,s}\left(1-t_{t+1}^{s}\right)]^{\omega}\left[h_{t+1}^{s}\right]^{1-\omega})^{\nu}\right]^{\frac{\alpha}{\nu}}\}^{\frac{1}{\alpha}}\times \\ \times[\rho\left(k_{t+1}^{f}\right)^{\nu}+\left(1-\rho\right)\left([l_{t+1}^{f,s}\left(1-t_{t+1}^{s}\right)]^{\omega}\left[h_{t+1}^{s}\right]^{1-\omega})^{\nu}\right]^{\frac{\alpha}{\nu}-1}\times \\ \times([l_{t+1}^{f,s}\left(1-t_{t+1}^{s}\right)]^{\omega}\left[h_{t+1}^{s}\right]^{1-\omega})\left[h_{t+1}^{s}\right]^{-1}, \end{array} \tag{47}$$

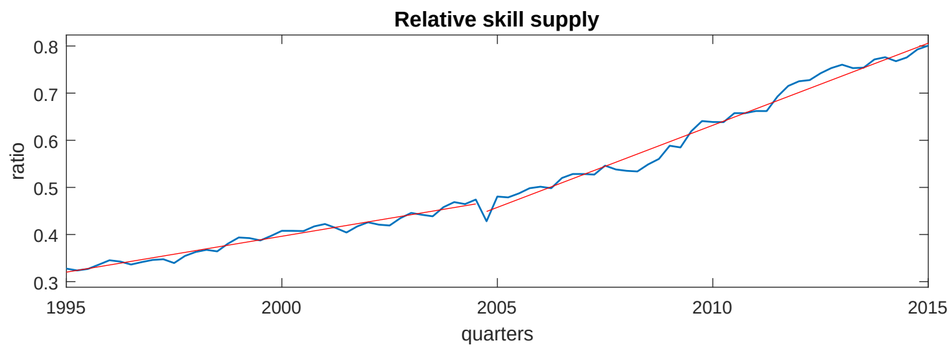
$$\frac{\partial h_{t+2}^{s}}{\partial h_{t+1}^{s}} = 1 - \delta^{s} + \gamma^{s} H^{s} \left(t_{t+1}^{s} l_{t+1}^{f,s} \right)^{\gamma^{s}} \left(h_{t+1}^{s} \right)^{\gamma^{s}-1}. \tag{48}$$

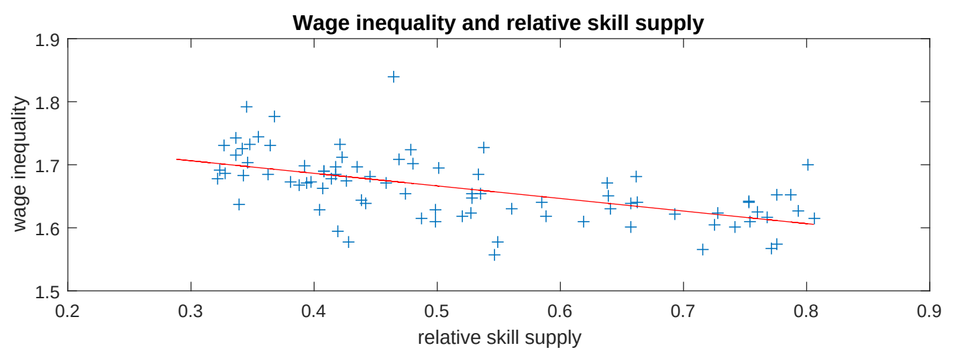
Figure 1: Stylised Facts (1995-2015)











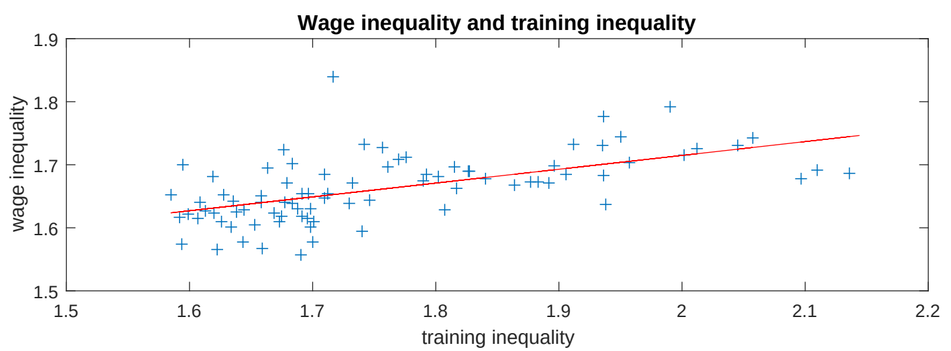


Figure 2: Permanent increase in τ^u from 0.042 to 0.5

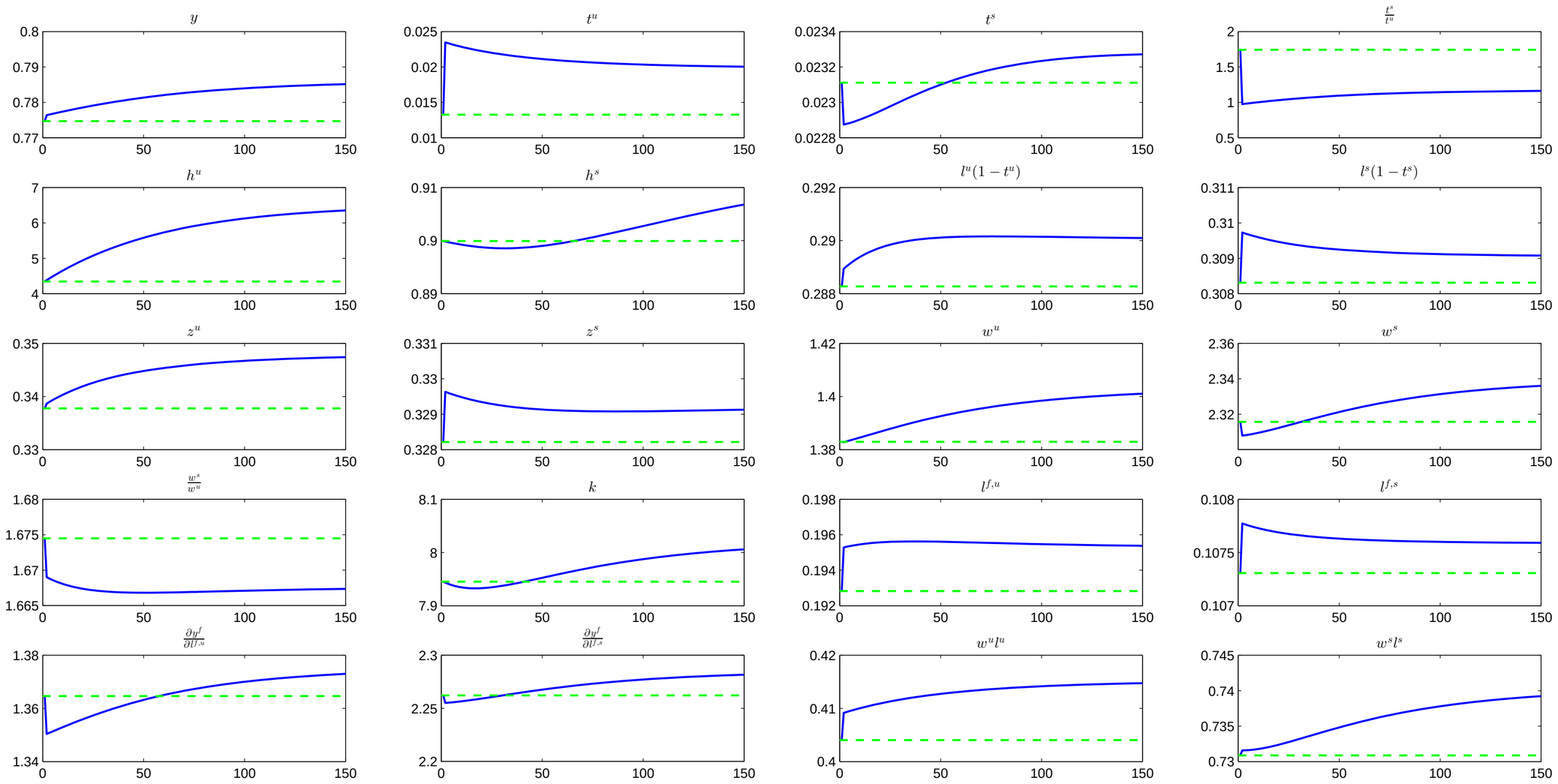


Figure 3: Permanent increase in τ^s from 0.042 to 0.5