Job training and inequality

by

Andrea Benecchi

a dissertation submitted in fulfilment
of the requirements for the Degree of:

Doctor of Philosophy
in Economics

Adam Smith Business School, College of Social Sciences
University of Glasgow

December, 2019

UNIVERSITY
of
GLASGOW

To my family and my friends.

Acknowledgments

THE PHD JOURNEY is a long and tortuous path. It was possible to reach the end only with the help of my family and many fantastic folks. The full list is too long to feature here but surely I could never forget any of you.

The end of the doctorate coincides with the end of my twenties' cycle, and the beginning of my thirties. A lot has happened during all these years, and I can barely imagine what expects me in the coming ones.

I am grateful to my family, my mother Giulia and father Riccardo, my siblings Veronica, Gianluca, and Stefano, and their families, for the support they gave me during all these years. The distance from home has been a painful consequence of my decision to pursue a PhD, but they have always been encouraging me and appreciating my work.

I am very thankful to my supervisors, Jim Malley and Konstantinos Angelopoulos, for their mentoring and guidance. They have been always present and have been warmly counselling me since the very beginning to the end.

The doctorate completes a cycle of education which started at primary school when I barely knew additions or language. It took me back to the start, teaching me there is so much I still don't know. In this sense, I hope I will never stop being a PhD student.

A sincere thank you goes to my friend Elisa who is always on my side whenever I need her; to my friends Eros and Giuseppe, whose connection is not weakened by time nor distance; and to Rami who is one of the most generous and considerate persons I ever met – and I have met countless of them during these years. No less important have been my flatmates, Spyros and Johanna, the officemates, Josue and Petros, but also Antonio, Andres, Giorgia, Laura, Mattia, Max, Simon, and all the amazing comrades I met on the PhD trail. They coloured my days.

I am grateful to the brilliant researchers and colleagues I met during the conferences, with whom I had very inspiring conversations and spent a great time. Among the many experiences, the time spent with Anastasia and Palev is vivid in my memory.

Also, I owe a great thank you to my friends from the water polo club, and all the other fantastic people I have met at the University of Glasgow.

The PhD at Glasgow taught me about the cycle of people. We are all in a continuous transition and that we need to accept that each of us is on his own different path. I am thus indebted to Aleksandra, but also Alan and Moritz.

Finally, I am grateful to the University of Glasgow and the Adam Smith Business Schools, in the form of all its members and staff, for having hosted me warmly and for having financially supported my studies. Oh, and of course I am grateful to my Gaggia which brewed aromatic coffees every single day.

Thank you all,

Andrea

Affidavit

I declare that, except where explicit reference is made to the contribution of others, this dissertation is the result of my own work and has not been submitted for any other degree at the University of Glasgow or any other institution.

Andrea Benecchi

Signed:

Abstract

This thesis is composed of three chapters. After a brief introduction, the first chapter discusses the definition of on-the-job training, reviews the literature, and reports empirical analyses for the specific case of UK. I decompose training participation and study its evolution in the last 20 years for specific sub-groups of workers, providing new compelling evidence.

The second chapter finds empirical evidence in favour of a relation between training and wage inequality between workers with different education level. On this basis, a dynamic general equilibrium (DGE) model with on-the-job training is developed and calibrated to match UK data. I use the framework to study the redistributional effects of training subsidies. The model is intentionally simple, to allow for a better understanding of the dynamics of macroeconomic variables after policy changes.

The third chapter proposes a more articulated general equilibrium model which features training externalities and distortionary income taxes. I present evidence that motivates the use of this framework, and its underlying assumptions. Thus, I calibrate the model to replicate the salient characteristics of the UK economy and I employ it to evaluate the welfare effects of policy reforms on training. The main contributions of my work are summarised in the conclusions.

Contents

Listing of figures

Listing of tables

Listing of appendices

Introduction

The thesis studies the role and the effects of firm-provided training activities in the UK and it produces policy recommendations about the desirability of training subsidies. In the first part, I collect and summarise empirical results dispersed throughout the literature and I provide novel evidence on training activities of UK companies. Then, I develop a theoretical model for firms' training decision and use it to study the effect of fiscal subsidies to training on workers' wages. Finally, I build a more realistic framework and produce policy recommendations for the optimal level of training subsidies.

Training activities are an important element of modern developed economies. Training is associated with higher wages for workers, and higher productivity for companies. The latter capture most of the benefits of training spells and finance most of their costs. Also, these activities, that in the UK amount to approximately 4% of firms' gross value added, have important aggregate effects on economic growth. According to the literature, training contributes to increasing total factor productivity together with other activities, such as research and development (R&D), and the diffusion of formal education.

Integrating this evidence, I show that, after a surge in the 1990s, training provision has fallen since 2002. Despite severe data limitations, it appears that training has declined both in terms of duration and number of workers who receive training. To investigate the drivers of these changes, I study how training opportunities are distributed across workers, bringing new insights on age-dependent trends in the UK data.

In the literature, there is evidence that training provision is unequally distributed among workers with different educational attainments. I contribute to this literature by reporting and analysing the inequality in training participation between University educated and non-University educated employees. Also, I directly link training inequality to wage inequality with an empirical analysis of the UK labour force data. The small but economically relevant relationship between training and wage inequality suggests that there is scope for policy-makers to intervene in this market.

Chapter 2 develops this issue, studying the effect of fiscal policies with a stylised general equilibrium (GE) model. I calibrate the model to data from the UK and ensure that it generates training and wage inequality that are consistent with the data. Then, I evaluate policies that target training for the unskilled workers by subsidising the firms and reducing the relevant financial costs. The model's predictions regarding the magnitude of the effect of training subsidies to training participation and the effect of the reduction in training inequality on wage inequality are consistent with the empirical evidence collect. With respect to both relationships, the model predicts effects just below the lower bound of my estimates.

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%), 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%), 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. Thus, the proposed framework helps explain the new empirical evidence and it builds a solid basis for the evaluation of policies related to firm-provided training.

Chapter 3 brings forth the analysis of wages and inequality that I started addressing in Chapter 2. The main contributions are three: (i) I build a new stylised model which features a range of endogenous channels affecting training provision and wage and productivity outcomes, and I extend it by introducing a more complex fiscal menu at disposition of the policy-maker; (ii) I find novel and compelling evidence that training provision is negatively affected by the separation rate of workers, i.e. poaching externality; and (iii) I provide a first quantitative evaluation of the effects of fiscal policies with respect to training subsidies that the UK government could implement.

The revised framework allows to measure changes of welfare for skilled and unskilled workers. On the empirical side, the analysis suggests that an increase in the job-to-job separation rate leads to a decrease in the training participation rate. This is consistent with the idea that firms curtails training investments if trained workers are likely to be hired by their competitors. Thus, it appears that training provision is sub-optimal.

Under these considerations, training subsidies can be Pareto improving. Yet, for high level of subsidies, the distortions caused by their financing outweigh the benefits. Measuring welfare for *ex-ante* different households, it can be observed that households have very different preferences on training subsidies.

Recent data suggests that UK institutions subsidise about 2.4% of the training monetary costs paid by firms. The analysis concludes that the subsidy rate is too low. Both households would benefit by a moderate increase of the subsidies to skilled training activities, however, such a policy would increase the inequality of outcomes. Conversely, unskilled training subsidies are more controversial. These subsidies increase the welfare of the average worker, but are likely to reduce skilled workers' welfare and to reduce labour productivity. Thus, the choice of the training fiscal policy requires the balancing between equity, efficiency and social preferences.

"Job training empowers people to realize their dreams and improve their lives."

Sylvia M. Burwell, US Secretary of Health and Human Services

1

On the job training: theory and empirical evidence

1.1 Introduction

The chapter provides a comprehensive overview on training activities that firms offer to their workers. To this purpose, I collect and summarise empirical results dispersed throughout a large literature and I provide novel evidence on training activities of companies located in the UK.

After presenting a working definition of training and a theoretical framework that justifies its provision, I discuss why training activities are an important element of modern developed economies. In particular, training is associated with higher wages, for workers, and higher productivity, for companies. The latter capture most of the benefits of training spells and finance most of their costs. Also, these activities, that in the UK amount to approximately 4% of firms' gross value added, have important aggregate effects on economic growth. According to the related literature, training contributes to increasing total factor productivity, together with other activities such as research and development (R&D) and the diffusion of formal education across the workforce.

Next, I quantify training activities in the UK and I provide evidence that, after a surge in the 1990s, training provision has been falling since 2002. Data limitations prevent estimating the intensive margin, but, given the available information, it appears that training has declined both in terms of duration of spells and number of workers who receive training. In relation to this issue, I investigate how training opportunities are distributed.

In the literature, there is evidence that training provision is unequally distributed among workers. I contribute to the literature by reporting and analysing the inequality in training participation between University educated and non-University educated employees. Also, I directly link training inequality to wage inequality with an empirical analysis on the UK labour force data. The small but economically relevant relationship between training and wage inequality suggests that there is scope for policy-makers to intervene in this market.[1]

Then, I investigate the business cycle characteristics of UK training participation rate. Contrary to earlier results in the literature, I show that there is no clear relationship between economic activity and the provision of training, at least for the UK. The results can be reconciled by the fact that other researchers employed different definitions of training, referred to other countries, or employed different proxies for the business cycle.

I also report that the great recession had limited effect on the training provision in the UK. This evidence further strengthens the idea that training activities are unresponsive to their indirect costs, i.e. the opportunity cost of working time, while direct costs may have a stronger influence.[2]

Finally, I employ decomposition techniques to further investigate the long-run

[1] I further develop this issue in Chapter 2, where I study the effect of fiscal policies using a general equilibrium model.

[2] This conclusion is drawn by the evidence that, *ceteris paribus*, companies facing high direct costs are less likely to provide training to their workers.

trend in training participation for the UK. What emerges from this analysis is that (i) workers moved towards training-intensive sectors and occupations; (ii) training inequality has reduced mostly because of a reduction in training of University educated workers, which is not related to their observable characteristics but appears to be residual to these factors; (iii) the probability of training for older workers, who are close to retirement age, has drastically increased and compensates the decline in training observed for all the other age groups.

The changes in training provision that I find have not been observed by the literature and provide a interesting puzzle. While I postulate several explanations to reconcile theory and empirical data, I do not present an ultimate solution the puzzle since more comprehensive data is needed to identify the factors that drove the observed changes in the distribution of training across the UK workforce.

1.1.1 Human capital theory and observed wage profiles

Since the seminal work of Mincer (1958), economists have investigated the relation between the human capital of a worker, i.e. her education, and her productivity. The latter, in absence of better proxies, is generally associated to her wage. It is widely agreed that the human capital accumulated through schooling and University education can explain much of the earning differential among workers with different education levels (see e.g. the Handbook of Education and Economics (2011)).

At the same time, it has been noted that average wages display specific patterns during the work-life of an employee. They generally rise at the start of her career, stabilise later on, and decrease slightly before retirement. Also, workers with different education levels have different wage profiles. In particular, more educated workers have steeper increases in their wage at the beginning of their career. To explain these three stylised facts, researchers have presented a number of competing theories: e.g. on-the-job learning (i.e. sheer experience), on-the-job training (OJT), and contract theory.[3]

In the context of the OJT literature, the main hypothesis is that training is complementary to education as a source of additional human capital for workers, increasing their productivity and, thus, their wage.[4]

As shown below, training is one of the main factors contributing to wage and productivity dispersion across the workforce. Despite its importance, little is known about job-related training for three main reasons: (i) training activities are different in terms of skills they develop and they have different impacts, e.g.

[3]See Rubistein and Weiss (2006) in the Handbook of Economics of Education, volume 2, chapter 1 for a comprehensive review of the literature.

[4]Section 1.2.4.3 presents alternative mechanisms that economists have proposed to justify the link between training spells and productivity.

safety vs technical training; (ii) these activities are poorly observed and surveyed; and (iii) training spells interfere with other job-related events, e.g. promotions and job changes within the same firm, that are often not (or not timely) recorded in survey data. Not controlling for training activity heterogeneity is likely to dilute the effect of training, as some of the training spells are mandatory or less relevant for the company. Conversely, promotions and career advancements that are concurrent to training may bias upward the estimates of training returns. These issues have received a great attention by the literature and are the subject of the next section.

1.2 Empirical evidence on training costs and returns

The section reports the most relevant evidence on job-related training that is available to date with respect to training costs and its returns. Even though this work focuses on the UK case, I expand the discussion to other countries to compensate the lack of UK-based evidence.[5] Although it should not be taken for granted, the strong similarities in training patterns between European (and generally OECD) countries suggest that empirical results from other countries

[5] Detailed information on training is costly to collect, thus countries have very different datasets. Some of those datasets span for long periods, other provides exceptionally detailed information about training activities over a shorter time. Hence, even though a large number of empirical works use UK and US data, important information can be gathered only from the evidence collected in other countries.

may hold true also for the UK.

1.2.1 Definition of training

Training activities are many and diverse, and most surveys fail to capture such heterogeneity. I focus on formal training courses, run either on the employer's premises or off-site. As shown below, these represent a large share of adult learning activities. Thus, I do not consider apprenticeships and formative courses for unemployed workers since, although those may be as important as on-the-job training, they are intrinsically different from it. In fact, apprenticeships are an instrument to facilitate the transition of new hires from education to the workplace, while remedial training should serve to mitigate the effect of industrial restructuring and to support the relocation of dismissed workers.[6]

The best way to discriminate between on-the-job training and other forms of adult education is to consider who pays for it. Indeed, apprenticeships are often state-subsidised and co-financed by employees (through lower salaries), and, generally, the state pays for virtually all training provided to unemployed workers. Conversely, on-the-job training is paid chiefly by employers. For this main reason, the latter is studied and analysed separately from other forms of adult learning.

[6] I use the word 'should' because several researchers, e.g. Rosholm and Skipper (2009) and Heckman and Carneiro (2003), argue that remedial training has weak positive effects in terms of job-market outcomes for its participants. For a more extensive review of this literature, see Heckman *et al.* (1999) and, more recently, McCall *et al.* (2016).

Further, the literature that studies employer-provided training proposes a distinction between formal and informal activities.[7] Several authors (e.g. Frazis and Loewenstein (2005) and Albert *et al.* (2010)) suggest that, even though there is no official data, a large amount of working time is invested into these informal activities. Yet, whenever data about informal training is available, these activities appear to have a very little effect: they don't increase firms' productivity, nor workers' wage. For example, Dostie (2013), employing a dataset of Canadian companies, shows that only formal training has a measurable effect on these variables. Black and Lynch (1996) obtain similar results for the US. For these reasons, I focus my attention on formal activities, which are more consistently reported in the surveys. This is also the prevailing approach in the literature.

1.2.2 Training costs and who pays them

The identification of who bears the cost of training is the first essential step to comprehend the role and nature of on-the-job training in developed economies. This is also the first step before I can evaluate the effects of a possible intervention of the state in the training sector.

[7] Informal training is an activity that takes place on the workplace. It is generally defined as the process of learning by others, e.g. a line manager or more tenured co-worker, or learning by oneself, which could by either learning by doing or learning by watching (see e.g. Black *et al.* (1999) and Destré *et al.* (2008)).

A few stylised facts with respect to the costs of on-the-job training can be identified. In particular, I find that (i) in European countries, the cost of job-related training activities is covered chiefly by employers who, on average, pay 80% of the costs (the share varies among datasets and countries but is never below 50%); (ii) in the UK, taking as an example the year 2004, companies paid 77% of the direct costs related to the training of their employees; and (iii) UK public funds and grants cover for less than 1% of the total training expenditures, after controlling for firms' own contribution to such funds.[8]

From a theoretical standpoint, workers should be paying for their training if and only if they are the beneficiaries of this investment. According to Becker (1962), training is *general* as long as it is equally valuable to any company. In this case, the firm needs to increase the salary it pays up to the level of the worker's marginal productivity if it wants to retain the trained worker. Thus, the latter is the only beneficiary of training. If training is *specific*, the productivity increase is entirely captured by the firm as the accrued human capital does not increase the worker's outside option.

Nonetheless, as Leuven (2004) suggests, market imperfections may render general training equivalent to firm-specific training, e.g. when the prevailing market

[8] Different data sources report a different size of government intervention, however they all suggest that training funds financed by general taxation are few percentage points of the direct training costs. On this matter, see the UK training datasets mentioned below.

wage is below the worker's marginal productivity. The literature names this situation *wage compression*, e.g. see Acemoglu and Pischke (1999). More formally, wage compression occurs whenever the elasticity of wage to productivity is lower than one. In this case, firms have incentives to provide training, although their offer could be sub-optimal. As a consequence, the theory can't predict *a priori* who pays for on-the-job training.

To address this question, Bassanini *et al.* (2007) have collected data from the Continuing Vocational Training Survey, run every five years among a large number of European countries. They report that employer-provided training represents the main component of all training activities, and that workers do not pay for OJT through lower initial salaries or flatter wage-tenure profiles. More precisely, their data reveal that training spells paid by employers represent about 70-80% of the total training expenditures and about 50-60% of the total time spent on training. The difference between cost and time shares can be justified by the fact that remedial training, often financed by the family or the government, lasts longer than employer-provided courses.[9]

To confirm these findings for the UK case, I collect specific evidence from the UK

[9]There are several justifications that can explain why remedial training may cost less than on-the-job training despite lasting longer. For example, this would be the case if fixed costs represented a large share of total training expenditure and if they were larger for firm-provided training.

Quarterly Labour Force Survey (QLFS). In this survey, training is divided into: (i) workplace training, if run within one of the establishments of a company; (ii) away training, if the activity is held in a separate place; and (iii) mixed training, if only part of the training is run within the company. In the period 2005-2015, employees reported that 40% of activities were workplace training, 42% were away training, and 18% were mixed. Workplace training is by definition an activity organised and fully paid by employers. Hence, to understand to what extent companies pay for training their employees, I have to consider who pays for the two other training types.

As reported in Table 1.1, around 60% of the remaining training activities are paid mostly by employers.[10,11] Combining these statistics, I conclude that, taking as reference the year 2014, UK employers have financed about 77% of the total expenditures in job-related training.

Several academics, e.g. Parent (1999) and Acemoglu and Pischke (2007), have advanced the hypothesis that workers may accept lower salaries to receive training, i.e. they finance these activities indirectly. Against this hypothesis, Booth

[10]Unfortunately, the survey does not ask the exact contribution of employer, family or government, but only the order by which they contributed. Thus, a worker reports who paid the largest part of the training costs but she does not specify if e.g. the company paid fully or partly the training costs.

[11]Since the question has not been asked for several years, Table 1.1 reports the older and newer period for which data are available. The excluded years provide similar figures as workers' response to the question is relatively stable.

Table 1.1: Who pays for training, %

	1992-1996	2010-2014
employer or potential employer	63.6	59.9
self, family or relative	15.9	19.2
other government or labour organization	11.1	8.0
no fees	5.3	9.7
employment action or similar program	2.0	0.3
other	1.7	2.2
don't know	0.4	1.7
total	100	100

and Brian (2002) show a positive but insignificant effect of next year's training on current wages for UK households. On the opposite, Connoly and Gottshalk (2008) show that, *ceteris paribus*, initial wages are lower for trainees. Yet, they adopt a broader definition of training which includes apprenticeships. As discussed earlier, apprenticeships have employee-employer dynamics different from that of job-related training activities. Thus, I do not consider their finding relevant to the current analysis. In the context of on-the-job training the main evidence is that training costs are not rebated onto trainees.

Finally, I verify that UK public funds don't cover a large share of these expenses in place of the employers. To do so, I look at the National Employer Skills Survey (NESS) carried out by the UK Commission for Employment and Skills (UKCES) in the years 2011, 2013, and 2015. They decompose training expenditures into multiple items. Their results are summarised in Table 1.2. In most cases, companies contribute to funds managed by the government, local authorities or trade associations, and they receive back credits to support their training needs. As reported in Table 1.2, the net transfer from public funds to employers represents less than 1% of the total expenses, even after considering tax deductions. In other words, the UK government contributes to the training expenses of employers, while UK firms bear most of the training costs for job-related training activities.

Table 1.2: Decomposition of UK job-related training expenditures

	2011		2013		2015	
	£bn	%	£bn	%	£bn	%
all Training, total	£43.8	100	£43.0	100	£45.4	100
off-the-job training, total	21.1	48	21.3	50	22.9	50
on-the-job training, total	22.7	52	21.7	50	22.6	50
levies minus grants	-0.3	-0.7	-0.2	-0.5	-0.2	-0.4

1.2.3 Implications of the empirical evidence

Since empirical evidence indicates that companies pay for most of the training, I advance two hypotheses: either training activities are firm-specific or there are important frictions in the labour market causing wage compression.[12] Since discriminating between the two cases may have important policy implications, many economists have collected evidence on this issue.

[12] As briefly discussed above, wage compression can occur because of e.g. asymmetric information, wage rigidities, search costs, and social norms (see e.g. Akerlof and Yellen (1990)). Frazis and Loewenstein (2006) find strong evidence of wage compression for the US labour market.

Most empirical investigations conclude that training has both a specific component and a general one. Parent (1999) analyses data from the US National Longitudinal Survey of Youth (NFLS), reporting that employers provide training to their workforce, even if, at the same time, their competitors are willingly to offer higher wages to trained employees. Parent maintains that firms benefit from training workers even after taking into account the risk of *poaching*. Labour poaching occurs when a competitor firm hires the trained worker. The new firm can capitalise on her higher productivity without bearing any training cost, hence, it can offer a higher wage. This is more likely to occur when training is general rather than firm-specific.[13,14]

If training were chiefly general, no firm would train its workers. For this reason, researchers exclude the possibility that training is purely general. On the other side, the presence of spillovers in favour of trained workers excludes that training is purely firm-specific. Hence, Garloff and Kuckulenz (2006), who find evidence in favour of firm-specific training for Germany, recognise that employer-provided training must have a general component too.

[13]See Brunello and Gambarotto (2004) for a study on local firm density and training provision in UK. They find these two variables to be negatively correlated, and one of the accredited explanations is the fact that firms in denser areas face higher poaching risk. Thus, they are less likely to train employees.

[14]Industry, or sector, specific training is comparable to general training, as competing companies desire the skills acquired and are willing to pay to poach them. Therefore, the externality will be stronger, the stronger is the competition for these skills in the local labour market.

To summarise, firm-provided training has general and firm-specific components in varying degrees; and, to complicate the picture, institutional factors, such as labour unions and legislation, also affect how rents from training are shared between employers and workers.[15] Provided this empirical evidence, as long as training spells increase productivity, companies should be the main beneficiaries of such investments. I verify this conjecture in the next section.

1.2.4 Returns to training

A large strand of the literature has investigated the returns to training. Those are measured in various ways, but in most cases training costs are not taken into account as such information is poorly measured, if available at all, in current datasets. Thus, researchers have focused on the difference in wage or productivity between a trained and an untrained worker, or the difference in productivity of companies with different training levels.[16] Below, I first review the evidence with respect to training returns from the point of view of companies, and then from that of their workers.

[15] As an example, take Acemoglu and Pischke (2003), who discuss extensively the consequences of the introduction of the 'minimum wage' legislation on unemployment and training investments.

[16] Productivity is generally defined as the gross value added, per worker, of a company.

1.2.4.1 Firms' returns to training

With respect to the benefits of training that are captured by the firms, I maintain that the estimation of the returns to training is challenging due to endogeneity and unobserved heterogeneity. Nonetheless, recent estimates for UK firms suggest that a 1 percentage point increase in training participation increases value added per worker by 0.6 percentage points (Dearden *et al.* (2006)). Also, returns to training seem to be heterogeneous. Some firms thus do not train their workers because the return rate on these investment is negative. Primarily, these companies face higher training costs, but it is not excluded that they have less to gain from a trained workforce.

As Table 1.2 shows, UK companies spend around 45 billion pounds per year in training their employees. Given the amount of resources invested, I expect these firms to receive sufficiently large benefits from training, unless they use these activities as a fringe benefit that integrates the salary of their employees.

Empirical works report almost unanimously positive effects of training on firms' productivity (see the evidence reported in Table 1.3). The effect of higher training participation is observed both at the firm level and among sectors. However, the effect of training is not easily measurable.

Table 1.3: Firm returns to training

paper	dataset	dependent variable	controls	effects
Dearden *et al.* (2006)	UK; 1983-1996	Value added per worker	capital per worker, R&D intensity, hour per worker, workers' qualification	0.6% for 1% increase in training (GMM)
Stanca, Colombo (2008)	IT; 2002-2005	Value added over employment at firm level	capital per worker, R&D intensity, size, training duration, age, executives' share, workers' share, sector	0.074% (GMM) 0.028-0.045% (FE-OLS) for 1% increase in training
Almeida, Carneiro (2009)	PG; 1995-1999	Marginal return of one more training hour for all employees	employees, workers' qualification, capital stock, executives share	-0.30% full sample (δ=17%) 8.60% for firms who are training
Zwick (2006)	DE; 1997-2001	Value added at the establishment level	collective wages, IT, hiring rules, sector, size, workers' qualification	0.76% for 1% increase in the share of trainees (FE with three IVs)
Dostie (2013)	CA; 1999-2006	Value added per worker at the firm level	worker flows (turnover), industry, year, type of training	11% more productive than untrained workers (GMM)
Vanormelingenm, Konings (2015)	BE; 1997-2006	Value added and wages	capital, firm's sector, $N°$ employees, year, average education, gender	.76% (.44%) value added (wage) increase for one hour of training

The literature has identified two main issues that affect estimation of training spells on productivity: endogeneity of training investments and unobserved heterogeneity that correlates with training. According to Stanca and Colombo (2008), failing to account for the potential endogeneity of training leads to underestimates of the effect of training on productivity. As Bartel (1994) and Black and Lynch (2001) have suggested, it is likely that, when a firm faces adverse market conditions and labour productivity is low, it invests more resources on training. On the other side, failing to account for unobserved firm heterogeneity leads to overestimation of the returns. In fact, when Stanca and Colombo use a fixed effects estimation procedure instead of OLS, returns to training are halved. This difference is not accidental, but consistently observed in many other works cited above.

Some of these empirical works estimate also the effect of training on average wages at firm level. In general, the effect of training on wages is estimated to be about half its effect on firm's productivity. Thus, it seems that either training has a relevant firm-specific component, or that firms' bargaining power is large enough to capture the rents from general human capital. In both cases, underinvestment seems not so severe as the earlier literature envisaged. In last instance, such evidence leaves less scope for policy intervention over concern about efficiency.[17]

[17] These are the conclusions, in a nutshell, that both Leuven (2004) and Pischke (2005) draw

Table 1.3 reports a summary of the most recent works that estimate the marginal effect of training on productivity. For the UK, Dearden *et al.* (2006) show that a 1 percentage point increase in training participation increases by 0.6 percentage points the per worker value added of a sector. Unfortunately, most researchers cannot estimate the rate of return due to lack of information about training costs. They can only estimate gross returns to training.

Almeida and Carneiro (2009), using a unique Portuguese dataset containing information about both training costs and returns, estimate that the average firm has a return on investment of about 8% (which may vary by few percentage points depending on the assumption on human capital depreciation). Yet, if those firms that do not provide training were to provide it, the marginal return of one hour of training would be negative. This evidence, although limited to firms with more than 100 workers, suggests that training is heterogeneous both in its provision and in the returns associated to it.

1.2.4.2 Employees' returns to training

With respect to the effect of training on trainees' wages, a few important empirical results emerge. The early literature claims that employees realise high returns from training. The high range estimates are up to seven times higher than the

from reviewing the past literature.

returns to schooling. However, several authors are sceptical about these estimates, e.g. Leuven (2004) and Pischke (2005). More recent works have found weak evidence of a causal effect of training on wages.

The literature has attempted to identify a non-spurious causal link between training and wage earnings as a first step to derive policy implications. Estimating the effects of training is a challenging task, as it is empirically correlated with many other events in the work-life, such as promotions. Thus, it is difficult to establish a causal relationship in one direction or the other, as well as to control for each factor. Nonetheless, several authors have found evidence in favour of a strong positive effect of training on wages, e.g. Frazis and Loewenstein (2005).[18]

To give a broad idea of the empirical evidences collected so far, in Table 1.4, I report a far from exhaustive list of the estimated effect of training on log-wages for a handful of countries. As emphasised by Pischke (2005), these works, in particular the ones published during the 90s, estimate "huge returns" from training.[19] If converted to annual returns, they are about five to six time greater than the returns from schooling.

[18] Yet, one of their concerns is that they may be not properly controlling for promotions, causing a bias in the estimation of log-wages changes due to training spells. They are also worried about other issues, in particular about: (i) measurement errors, (ii) heterogeneity in wage growth, and (iii) heterogeneity of returns to training with respect to job characteristics.

[19] In Table 1.4, I do not report them to focus on the results of more recent literature. However, summary tables are available in Leuven (2004).

Table 1.4: Worker returns to training

paper	dataset	log wage on	controls	effects
Lillard, Tan (1986)	USA; 1983	Dummy for training in current job	SS unemployment rate, ethnicity, region schooling, tenure, experience	22%
Gerfin (2004)	SW; 1996, 1999	N° of training events	age, tenure, job, residence, skills, schooling, industry	2% per training event (difference in difference)
Bassanini *et al.* (2005)	EU; 1995-2001	Sum of training events over six years time	schooling, gender, country, industry, age, marital status	between 3.7% and 21.6% (depending on country)
Parent (2003)	CA; 1995	Dummy for last year training	GPA, marital status, parents' education schooling, tenure, drop-out (dummy)	10-11% men (FE-OLS) 2-8% women (FE-OLS)
Budria, Pereira (2004)	PG; 1998-2000	Dummy for ever received training	age, experience, schooling, tenure, part-time, firm's size	12% for men 37% for women
Albert et al. (2010)	EU; 1995-2001	Dummy for last year training	gender, education, experience, hours, firm's size, seniority industry affiliation, occupation	positive (OLS) not significant (FE)
Dearden *et al.* (2006)	UK; 1992, 1997	Trained 1, 2, 3 or 4 quarters ago	age, qualification (dummy), occupation, firm's size, industry, gender	0.15%, production sector only
Leuven, Oosterbeek (2002)	HO; 2001	Dummy for last year training	age, schooling, tenure, firm's size, control group: involuntary untrained	not significant
Leuven, Oosterbeek (2004)	HO; 1999	Dummy for last year training	age, schooling, tenure, firm's size, RD method: discontinuity due to law deductions if worker is older than 40	not significant
Sousonis (2009)	UK; 1998-2005	Dummy for last year training	gender, schooling, tenure, firm size, age, job, region, time, children	-3% (OLS) not significant (FE)

Pischke is sceptical about these values, and he suggests two possible explanations. First, he argues that a single training spell could be extremely useful, for example to someone promoted to a managerial position, but any additional investment would be unproductive. In this case, marginal returns decrease fast to zero for any additional training. An alternate hypothesis is that standard methods fail to control for the endogenous selection of workers into training.

Indeed, recent works that employ different empirical techniques to study the relation between training and wages have revised downwards these estimates. Most of the recent studies use instrumental variables (IV) to control for endogeneity between wages, training and skills. They find little or no relationship between training and wages. Yet, their results are criticised over two main concerns. First, compared to the traditional approach, they rely on small samples with low statistical power. Second, as they generally select small groups of similar workers, one as treatment and the other as control, their estimation may not be representative of the whole population.

An example is Leuven and Oosterbeek (2002). They employ a survey based on 3074 interviews with workers from randomly selected Dutch companies to perform the following test. They compare the average wage of trained workers to that of workers who couldn't train due to some random event, such as an illness or family circumstances. The authors fail to reject the null hypothesis of no wage difference between participants and control group.[20]

A serious concern coming from this literature is that some workers may have unobserved characteristics that lead both to steeper wage-experience profiles and to higher training participation. Thus, the relation between wages and training remains somewhat disputed.[21]

1.2.4.3 How training can affect productivity and wages

The hypothesis of positive effects of training on wage and productivity relies on the underlying assumption that, through training, workers gain productivity-enhancing skills. However, there could be other reasons for a causal relationship between training and productivity. The literature has advanced a number of alternative hypotheses with respect to the role of training in advanced economies. I take into consideration four potential effects of firm-sponsored training: (i) employees happiness, (ii) the probability of promotion, (iii) turnover reduction, and

[20] In another work, Leuven and Oosterbeek (2004) report previous estimates that are consistent with their findings: "Examples of studies with low or zero returns are Booth (1993) for the United Kingdom and Pischke (2001) for Germany. Lynch (1992) and Veum (1995) report returns to company training incidence not statistically different from zero for US NLSY data".

[21] Recent works report statistically, and economically, significant wage returns from training for developing countries, for example the work by Almeida and Faria (2014). Yet, in these economies, training could be a substitute for education, especially if the quality of education is particularly low. For that reason, these results cannot be extended to different contexts, e.g. to OECD countries.

(iv) human capital externalities. These theories support the importance of training for companies, sectors, and the economy of developed countries.

HAPPINESS Happiness, or job satisfaction, has been related to a series of positive labour market outcomes, such as reduction in absenteeism, and higher growth of wages and productivity.[22] Consequently, researchers from economics and psychology have studied the main determinants of job satisfaction. Among them, Budria (2012) estimates the effect of training spells on self-reported job satisfaction based on the European Community Household Panel Survey (ECHPS). He argues that the effect of training on happiness can be equated to a 17.7% increase in earnings. Notwithstanding his results, the link remains disputed as all such analyses rely on employees' self-reported mental status, rather than objective measures.

PROBABILITY OF PROMOTION Other works have tried to test whether training can predict promotions. Melero (2010), using the BHPS, estimates a positive affect of training on the probability of promotion for women. The coefficient is smaller and non-significant for men. The author suggests that firms reward women according to the market value of the skills they possess, while for men promotion is mainly a tool to induce effort. Thus, training has much less importance for

[22] See respectively the works of Judge *et al.* (2001) and Wegge *et al.* (2007).

the latter group. Since this differential could be attributed to other uncontrolled factors, such as composition effects (women achieve on average lower education levels), the author considers other plausible explanations (rejecting them). If confirmed, the link between training spells and promotions is additional evidence that job-related training is an important activity for both firms and workers. In particular, for a firm, training could be an efficient alternative to hiring new employees to fill specific vacancies (see e.g. Blatter *et al.* (2012)).

Reducing turnover Several authors argue that providing training to low productivity workers can lead to a significant reduction in the turnover rate.[23,24] This literature assumes that idiosyncratic shocks hit workers productivity and, if large enough, they break the match. More precisely, an employer separates from the worker whenever the rent from the match becomes negative. A firm that provides training to low-productivity workers faces lower firing costs since idiosyncratic shocks are less likely to induce separations.

However, since training is strongly correlated with educational attainment and wages, Budria and Pereira (2007) argue that training is "far from remedial". In other words, firms invest mainly on top employees. If training could reduce the

[23]Labour turnover is defined as the proportion of a firm's workforce that leaves during a given period of time.

[24]Theoretically, such result can be obtained by introducing endogenous separation in a model with employer-employee matching, see for example Lechthaler (2005) and (2009).

turnover rate of low skilled workers, the data should display a high training participation rate for these workers on the marginal productivity threshold; arguably, it doesn't. On the contrary, Dostie (2013) provides the empirical evidence that workplaces with high turnover provide more training, due to the incidence of induction training for new hires. It is possible that researchers overlook informal training activities, since the latter are virtually unreported in available datasets. However, according to Dostie, this type of training does not seem to be productivity-enhancing, as it only provides some basic knowledge to the new hires. Thus, this theoretical framework, based on idiosyncratic shocks, finds little support in empirical data.

Nonetheless turnover could be affected by training through a different mechanism. The investigation of Koster *et al.* (2011), on a small sample of 2833 Dutch pharmacy assistants, suggests that investing in general training contributes to job satisfaction and to the "perceived support in employment development" (PSED), which can reduce sensibly turnover rates. The underlying theory is borrowed from the social exchange theory, as developed by Blau (1964) or Eisenberger *et al.* (1986) among the others. According to the authors, workers' reciprocity and gratitude allow firms to profit from investments in general human capital. Unlike the matching theory reviewed above, reciprocity affects any worker, therefore it is compatible with the fact that more educated and productive workers receive more training.

In a more recent paper, Leuven *et al.* (2005) argue that worker reciprocity has a positive and significant effect on the probability that a firm will provide training. They estimate a 15% increase in the probability of receiving training due to observable differences in the reciprocity index, which they engineered from a Dutch survey. They speculate that, since training is mainly general, a firm knows that training workers who are more likely to feel indebted yields higher expected returns. The reciprocity theory, unfortunately, cannot be tested on the QLFS dataset due to data limitations. Nonetheless, this channel could have a high economic impact for companies. For example, Mattox and Jinkerson (2005), using a private dataset of a large US company, show that training programs meant to retain key managers or experienced employers had a return on investment ranging from \$12 to \$21 per dollar spent.[25]

From a macroeconomic perspective, economists have found evidence of unconditional negative correlation between training and the separation rate of workers,

[25]It must be noted that the rate of return estimated by Mattox and Jinkerson (2005) is incredibly high compared to other works', e.g. Almeida and Carneiro (2009). Yet, their result is not directly comparable with all the others since: (i) it is based on a very specific sub-sample of employees, i.e. key managers; (ii) data comes for a single large US multinational company; and (iii) the training programmed under scrutiny is precisely monitored, whereas, in general, researchers observe more noisy information.

which can explained in two ways.[26] From one side, trained workers are less likely to change job, at least in the short term (consistently with what reported above). On the other side, firms may train workers when the probability of separation is low, as expected returns are higher. In this respect, Brunello and De Paola (2009) fail to find a statistically significant relationship between past training and the probability of changing job the year after the training spell. Hence, they question the rationale for government intervention on the training sector out of efficiency concerns. They use data from 7 waves of the European Community Household Panel, but they employ a linear probability model and, while they control for individual fixed effects, they are unable to control for endogeneity issues, or reverse causality. For these reasons, their results do not appear robust and leave scope for further research. In chapter 3, I show that, at least in the UK, firms reduce training when the rate of job-to-job transitions is higher in the local labour market. This represents evidence of the poaching externality and it calls for government intervention.

Workplace externalities A strand of the literature, among the others Galindo-Rueda and Haskel (2005) and Metcalfe and Sloane (2007), has shown that within-sector and within-firm education spillovers can be strong, especially

[26] I confirm this result for the UK in Chapter 3.

from high educated to less educated workers.[27] Further, O'Mahony and Riley (2012) show that training can boost the relationship between an employee's wages and the education of his peers. Similarly, several authors have argued that human capital can be informally passed on from some workers to others as they work together. For example, De Grip and Sauermann (2011) have carried out a field experiment in a call centre from a Dutch telephone company to estimate return to training for workers and their peers. What emerges is a consistent transmission of know-how from trained to untrained workers. More precisely, they estimate that the performance of an untrained worker is in average 0.51% higher if the share of trained peers is increased by 10 percentage points.

If these spillovers are strong enough, they can bias downward the estimates of wage returns to training. Also, another possible consequence is that these spillovers may induce firms to train only part of their workforce, as the know-how will diffuse among the workers through informal learning.[28]

[27]The cited works refer to spillovers within sectors and within company's workforce, respectively. Also, they are both based on UK datasets.

[28]It must be noted that the experiment is carried out in a very specific context. Thus, it is unclear whether their results can be generalised to other workplaces. De Grip and Sauermann (2011) themselves recognise this limit.

1.3 Training inequality

Since the 1980s, researchers have pointed out that training opportunities are not evenly distributed among the workforce (e.g. Lillard and Tan (1986), Barron *et al.* (1989), and Brown (1989)). This section shows that there is a large variability in the provision of training which can be attributed to many observable and unobservable characteristics, and that training outcomes are persistent with time.

1.3.1 Determinants of training

A worker's probability of getting trained appears to be influenced by both individual characteristics and characteristics of the firm that employs her. I consider each group of factors separately in the next subsections.

1.3.1.1 Training differential among firms

Two main factors affect the probability of training at the firm level, i.e. dimension and innovativeness.[29] According to Bassanini *et al.* (2005), in all EU countries large and/or innovative firms invest similar resources in training. For that reason, the difference in training participation rate among those countries can be explained by the share of small-medium enterprises and their different propensity to train

[29] As an exception, innovative and non-innovative firms are equally likely to train employees in countries with high levels of training participation (e.g. Scandinavian countries).

the workforce. Both lower costs and higher returns to training can justify the evidence that large companies invest more in training.

Also, they observe that higher R&D investment entails more training. However, the effect is weaker on college graduates as they require less training to adapt to new technologies. The schooling system plays an important role as well. On average, after technological innovations, workers coming from vocational schools receive more training than those coming from a general track.[30]

Finally, training differs sensibly among sectors as well. For the UK, the difference in OJT participation between services and production sector is relatively small. According to the results of the fourth continuing vocational training survey (CVTS), training participation is estimated to be 49% in production and 56% in services, which means that in the year of the survey about half of the workforce has received at least some sort of training. Yet, the authors of the report warrant that in many of these cases training is mandatory, enforced by the state through laws, e.g. to ensure workers' security on the workplace. This explains why "mining and quarrying" leads the production sectors with a 65% share of trainees over its workforce.[31] The cases where firms decide how much to train

[30]For example, as Beck *et al.* (2009) and Grund and Martin (2012) discuss, German firms invest in apprenticeships more than other countries', thus they tend to provide less training afterwards as workers have already acquired the skills they need on a given job.

[31]The authors recognise that lacking information on mandatory training is one of the limits of their investigation.

their employees are more interesting, as these companies provide training either to retain key competencies and/or to enhance workforce productivity. Sadly, the distinction between different training activities has not been considered in worker, or household, surveys.

With respect to the UK services sector, "education", "health" and "public administration" have the highest level of training incidence with a share of 71, 77 and 80 percent of trainees over total workers, respectively. The finance and insurance sectors follow them with a training participation rate of 59%, and they represent the private sector with the highest incidence of training.

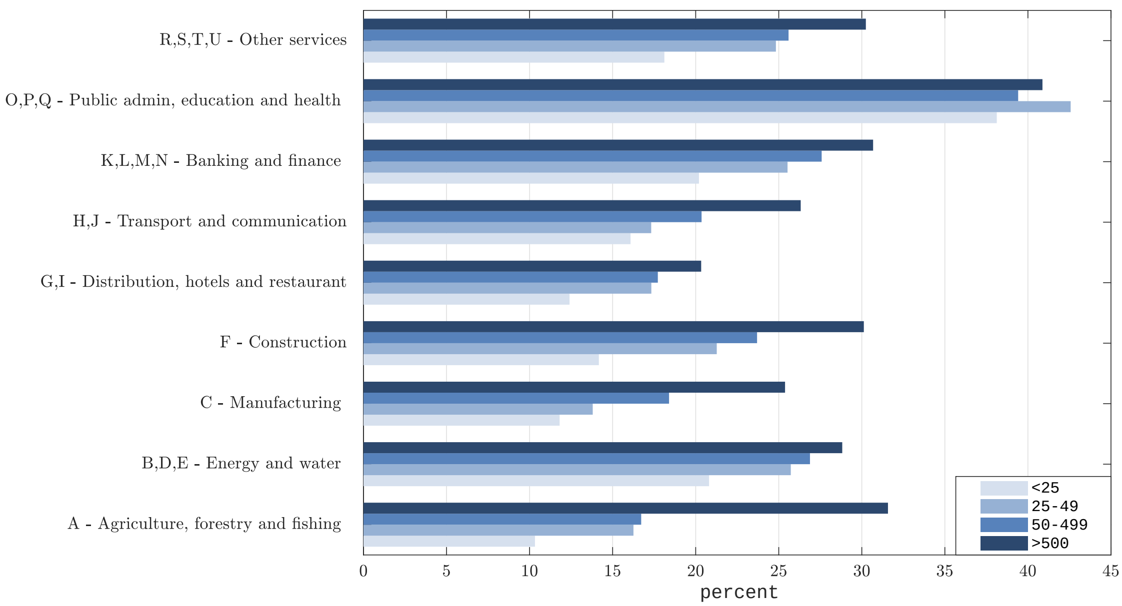


Figure 1.1: Training participation by industry and size, average 2009-2016

The analysis of training participation rates based on the QLFS confirms the qualitative findings of the CVTS report, in particular with respect to differences among sectors. However, a smaller proportion of workers report they are training in the former survey. The difference between the two indexes ranges between 20 and 30 percentage points. So, on average, 30% of workers get trained every quarter, but only half of the working population reports having received any form of training in the last year. This evidence supports the intuition that some workers receive training on a regular basis, whilst a large group of workers have small or no access to training activities. In Figure 1.1, it is possible to see that firm size has a big impact on training probability of workers, for all sectors except ”public admin, education, and healthcare.” For this sector only, about 60% of the employees report that they work for a state-owned, or public, company. Thus, the evidence can be accommodated by assuming that public companies commit to higher training targets, and the size of the company does not influence the training policy that is implemented.

In the bar chart, I report the average training participation for the years 2009-2016 for a number of reasons. First, there is very little variation between one year and the other. More importantly, sectors comove closely. Thus, it is not particularly interesting to look at yearly variations in the training participation rate across sectors. To verify the absence of relevant idiosyncratic trends, I use two digit SIC classification and I rank sectors by their training participation rate. Then, for each year, I distinguish the top 30 sectors from the rest with a dummy variable. I observe that only 10 out of 88 sectors move from one group to the other in at least 30% of the last 20 yearly observations. That is to say that very few sectors have experienced a significant change in training (increase or decrease) that deviates from the overall training pattern. I come back to this issue in Section 1.6.2.1. The next section discusses how individual characteristics of workers correlate with training participation rates.

1.3.1.2 Training differential among individuals

Heterogeneity in training participation depends largely on worker characteristics. As Bassanini *et al.* (2007) show, training probability is positively correlated with educational attainment, even after controlling for several factors such as firm's sector. Thus, all else equal, low skilled workers are less likely to be trained. I also expect that a relatively larger share of their training is compulsory. This would be the case, for example, if laws enforced the same requirements on health and safety trainings for all workers.[32]

The same source reports that employers are more likely to train men rather than women, especially at the beginning of their career. At the same time, mar-

[32] White collar jobs should be in general safer than manual works though.

ital status and the type of contract appear to have a significant incidence on training probability.[33] Another factor that correlates with training is trade union membership. Empirical analyses have found that Unions have a positive effect on the probability of training, and in some cases even on returns to training. Booth and Francesconi (2003) provide some evidence in this respect for the UK economy.

Most importantly, training participation is a highly persistent outcome. Workers who receive training are more likely to be offered new training the next time (see e.g. Bassanini *et al.* (2007) on page 70). This feature of training can favour workers' segmentation. Even though educational attainment and other individual characteristics are similar, some workers receive training regularly, while others are largely untrained. For the UK, Sousounis and Bladen-Hovell (2010) report that, for a previously untrained worker who gets trained in one period, the probability of being trained in the next period is raised by 0.401, on average, due to the persistence of training outcomes. The corresponding figure for women is 0.362.

Through this mechanism, the wage differential between trained and untrained workers will increase with time, one training spell after the other. This outcome may be socially undesirable, especially in the case training outcomes do not to reflect differences in ability, skills, or effort of workers. Indeed, for the countries of

[33] Probabilities are estimated for a pooled dataset at European Union level, after controlling for time and country effects. Data come from the European Community Household Panel (ECHP) for the years 1995-2001.

south-western Europe, parents' educational attainment has been found to predict training participation of a worker. According to Bassanini and his colleagues, who report these results, in such countries family and social relationships play a great role in finding a good job, thus in receiving company provided training.

A more detailed evaluation of all the determinants of training participation is offered in the next section.

Novel evidence on training determinants Although the literature has pointed out that UK training has changed through time both qualitatively and quantitatively, little has been done to analyse these changes. Thus, to integrate the evidence coming from the literature, I estimate the probability of training with a Probit model, splitting the last twenty years of (quarterly) data into two periods, i.e. before and after the training peak of 2002.[34]

As shown in Section 1.6.2, the average training participation rate is almost the same between the two periods, at least at the endpoints. Thus, the analysis focuses on how the distribution of training has changed from the past to the present. More precisely, I estimate over different sub-periods the simple Probit

[34]The choice, although based on the peak observed in the data, may appear arbitrary. Thus, as robustness check, I take the oldest 8 years and the 8 most recent years of data, and perform the same exercise. Results are consistent to the one reported here, but the changes in training patterns are relatively larger. This is to be expected since, in the latter case, I drop the years 2002-2008 to compare past and present training patterns.

model:

$$y_{i,j,t} = 1\left[X_{i,j,t}\beta + \beta_j + \gamma_t + \varepsilon_{i,t} \geq 0\right] \tag{1.1}$$

where $y_{i,j,t}$, a binary outcome which denotes whether worker i, who is employed in cell j,[35] has participated in work-related training in the period t; $1[\cdot]$ is an indicator function that can assume only the values 0 or 1, depending on the realizations of the latent variable within the squared brackets; $X_{i,t}$ is a set of exogenous individual and job-specific characteristics, such as age, gender, firm's size; β_j are sector and occupation specific fixed effects; γ_t is a set of quarterly time dummies (remember that I use a pooled cross-sectional dataset); and $\varepsilon_{i,t}$ are normally distributed error terms, $\sim N\left(0, \sigma_\varepsilon^2\right)$, which may be heteroskedastic and correlated within industry-sectors.

I use the Probit estimates to compute the predicted probability of training with respect to a number of worker's characteristics. In particular, I interacted all observables with the skilled-unskilled worker dummy, and the female-male dummy since these two characteristics are the most relevant.[36]

Table 1.5 reports the probability of training conditional on several variables, under the assumption that all other characteristics are *as balanced*, which means

[35] A cell is defined by the interaction between the sector and the occupation of a worker.

[36] There is a large and growing literature that studies the differences in pay, career opportunities, and training between male and female employees, e.g. Don and Sheryl (2002), Kunze (2005), Arulampalam *et al.* (2007), and Blau and Kahn (2017), among the many.

Table 1.5: Training probability by worker's characteristics,%

	1995-2001		2009-2016		1995-2001		2009-2016	
	unskilled	skilled	unskilled	skilled	male	female	male	female
non married	22.2	27.7	21.7	26.5	23.6	26.2	23.2	24.9
married	18.8	24.9	19.8	23.4	21.3	22.3	21.1	22.0
male	20.4	24.5	20.4	24.0	-	-	-	-
female	20.5	28.2	21.1	25.9	-	-	-	-
temporary	20.1	23.6	20.1	23.0	20.4	23.3	20.7	22.4
permanent	20.8	29.2	21.4	26.9	24.5	25.1	23.6	24.5
part-time	18.7	24.1	18.8	23.3	21.5	21.1	20.8	21.1
full-time	22.4	28.7	22.8	26.6	23.4	27.5	23.5	25.9
small firm	17.9	25.7	18.9	24.2	20.0	23.3	20.3	22.7
large firm	23.2	26.9	22.8	25.6	25.0	25.0	24.1	24.2
no-white	19.9	25.5	20.8	24.5	22.1	23.2	22.3	23.0
white	21.0	27.1	20.7	25.3	22.8	25.2	22.0	23.9
private	18.5	23.4	18.9	22.2	20.3	21.4	20.2	20.8
public	22.5	29.4	22.7	27.8	24.7	27.1	24.2	26.2

that they are all equally likely. In this sense, each predicted probability refers to a hypothetical individual who is equally likely to be e.g. part-time and full-time when controlling for e.g. skills and private sector. By doing so, it is possible to compare these probabilities to each other and discover which worker is more likely to be trained, all else being constant.[37]

As a key result, the analysis suggests both a large skill-premium, in favour of educated, and a large gender-premium, in favour of female employees. Thus, skilled workers are more likely to be trained, both in the old and in the new period, and so are women. The result that, *ceteris paribus*, female workers are more likely

[37]The Probit model includes as regressors workers' age, time dummies, and SIC industry. Yet, those are not shown here as they are not the focus of this section.

to be trained stands in contrast with earlier evidence, but is in line with the findings of Dearden *et al.* (2006). The evidence from Greenhalgh (2002) helps explaining the divergence of results. According to his findings, women's training provision rose faster than men's in the 1980s and surpassed the latter since 1989.

The highest training participation rate is observed for skilled married female, skilled workers in full time employment and workers working in the public sector. In the more recent period, training differentials between male and female, and between skilled and unskilled workers, shrank but they are still large and significant.[38]

Table 1.6: Training probability by occupation,%

	1995-2001		2009-2016		1995-2001		2009-2016	
	unskilled	skilled	unskilled	skilled	male	female	male	female
managers, directors, senior officers	25.5	32.1	23.4	27.9	27.7	29.7	24.9	26.3
professional occupations	30.0	37.1	28.0	32.8	32.4	34.6	29.6	31.1
professional and technical associates	28.8	35.7	26.6	31.3	31.1	33.2	28.1	29.6
administrative and secretarial	20.0	25.8	17.3	21.1	21.9	23.7	18.6	19.7
skilled trades	16.9	22.1	19.0	23.0	18.7	20.2	20.3	21.6
caring, leisure and other	21.3	27.2	26.7	31.4	23.3	25.1	28.3	29.7
sales, customer service	22.8	28.9	20.2	24.3	24.8	26.7	21.6	22.9
process, plant, machine operators	13.7	18.3	16.4	20.0	15.2	16.6	17.6	18.7
elementary occupations	10.6	14.5	12.6	15.7	11.9	13.0	13.6	14.5

Table 1.6 focuses on training differences between job occupations. The results are in line with previous research, and confirm that more specialised jobs require

[38]According to Bassanini *et al.* (2007) the higher training participation rate of female employees is related to their higher willingness to pay for their training, however I cannot test this hypothesis with the UK QLFS data.

more training investments. It is worth noting that there is a sizeable increase in training probability from the old to the new period for those jobs which train less, e.g. elementary occupations, whereas training participation has decreased for those jobs where training was more frequent, e.g. professional occupations. This result suggests that recent technical progress has demanded a skill upgrade to those at the bottom of the skill distribution.

Table 1.7: Training probability by tenure,%

	1995-2001		2009-2016		1995-2001		2009-2016	
	unskilled	skilled	unskilled	skilled	male	female	male	female
less than 3 months	25.7	27.3	26.7	26.3	26.3	26.8	26.2	26.8
3 months but less than 6	24.0	29.4	25.9	28.2	25.3	28.1	26.0	28.1
6 months but less than 12	21.8	28.2	23.0	27.3	23.7	26.1	24.0	26.2
1 year but less than 2	19.6	27.3	20.6	25.6	21.8	24.8	22.0	24.1
2 years but less than 5	18.7	25.0	18.7	23.8	20.8	22.7	20.5	21.9
5 years but less than 10	18.2	23.4	17.5	22.6	19.8	21.7	19.5	20.4
10 years but less than 20	18.0	24.7	17.3	22.5	20.5	21.9	19.4	20.2
20 years or more	18.5	25.4	17.8	23.4	21.7	22.0	20.4	20.6

Finally, Table 1.7 presents the relationship between tenure and probability of training. In this case, the most striking (but not unexpected) result is that tenure has a non-monotonic effect on training provision after controlling for skill, sex, and other covariates. The highest training rates are observed within one year from the hiring. Thereafter, training spells are progressively reduced. However, as tenure is sufficiently high, i.e. 10 (20) years for skilled (unskilled) workers, firms invest once again in their formation. The result is even more remarkable considering that I control for age which is highly correlated to workers' tenure within a company. It appears that firms are compelled to train senior workers to compensate for the depreciation of skills as they have completed their education much earlier than other workers.

So far, I omitted any reference to age. However, the latter has a strong and important effect on the probability of training. Furthermore, the UK working population in the last twenty years have been changing significantly in terms of age structure, and this, I argue, has largely affected the training policies of firms. For these reasons, the next section describes and discusses this issue.

1.3.2 Does education affect lifetime training?

Scholars, e.g. Hansen and İmrohoroğlu (2009), show that a simple overlapping generation (OLG) model with on-the-job training can reproduce the worker's tendency to accumulate high initial skill capital early in the career and reduce her stock by curtailing on-the-job training later in life. Indeed, as workers have less residual work time, the present value of further training investments, up to the point that individuals close to retirement choose not to invest time and/or physical resources to accumulate new human capital.

Figure 1.2 shows that, for unskilled workers the training participation rate de-

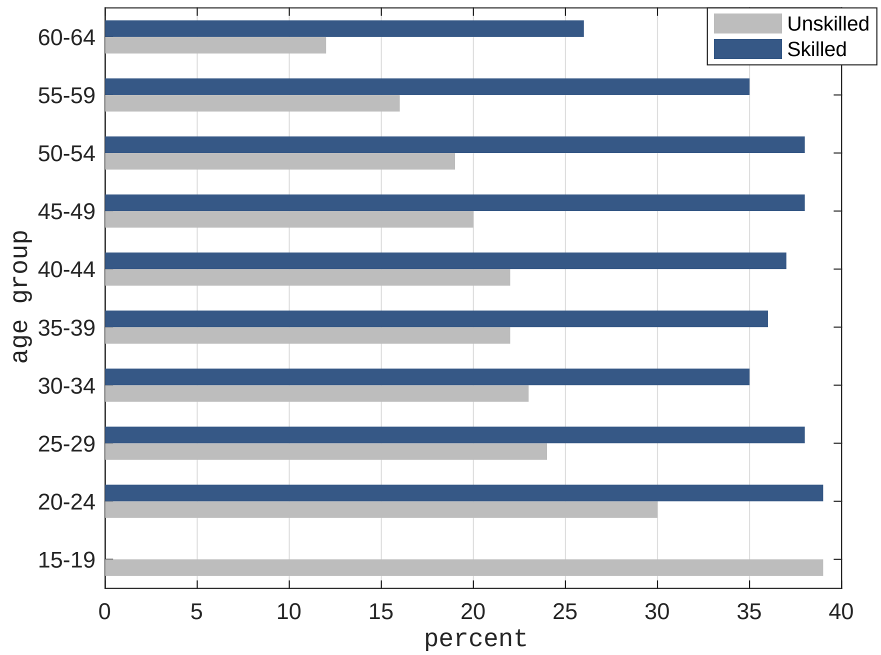


Figure 1.2: Training participation by age and skill

creases monotonically as they get older, whereas for skilled workers the training participation rate does not decline significantly until very late in their life.[39] For the latter group, the largest drop in training participation occurs between the age groups 55-59 and 60-64. Thus, old skilled workers are less likely to train than young ones, but they receive more training than most unskilled workers.[40]

Computing the average training level by age cohort based on about 20 years of data available from the QLFS hides significant changes in the relationship

[39]Keep in mind that I consider ”skilled” a worker who has at least a bachelor’s degree and ”unskilled” a worker who does not have any University degree or equivalent qualification.

[40]Only young unskilled workers receive more training than old skilled ones. This proves how uneven are training opportunities between the two groups.

between training and a worker's age. Later, in Section 1.6.2, I come back with a more robust analysis of this relationship, highlighting the changes observed in the training-age profile over the last twenty years. The next section focuses on one of the key issues of my thesis, i.e. the relationship between training and wage inequality.

1.4 Training and wage inequality

The section contributes to the literature by generating evidence that exposes the link between training inequality and wage inequality. As Sousounis and Bladen-Hovell (2010) suggests, the persistence of training outcomes and the unequal distribution of training opportunities, together with the positive effects of training on wages, pave the way to a causal link between training and wage inequality.

I test this relationship for the UK with the data collected by the Office of National Statistics for the Quarterly Labour Force Survey. Since at the individual level education is one of the main factors influencing training participation, I divide the workforce into non-University educated and University educated workers, whom, from now on, I call unskilled and skilled workers.

I compute the training participation rate of each group as the share of workers who got trained in a given quarter (the unit reference of the survey). Thus, I use

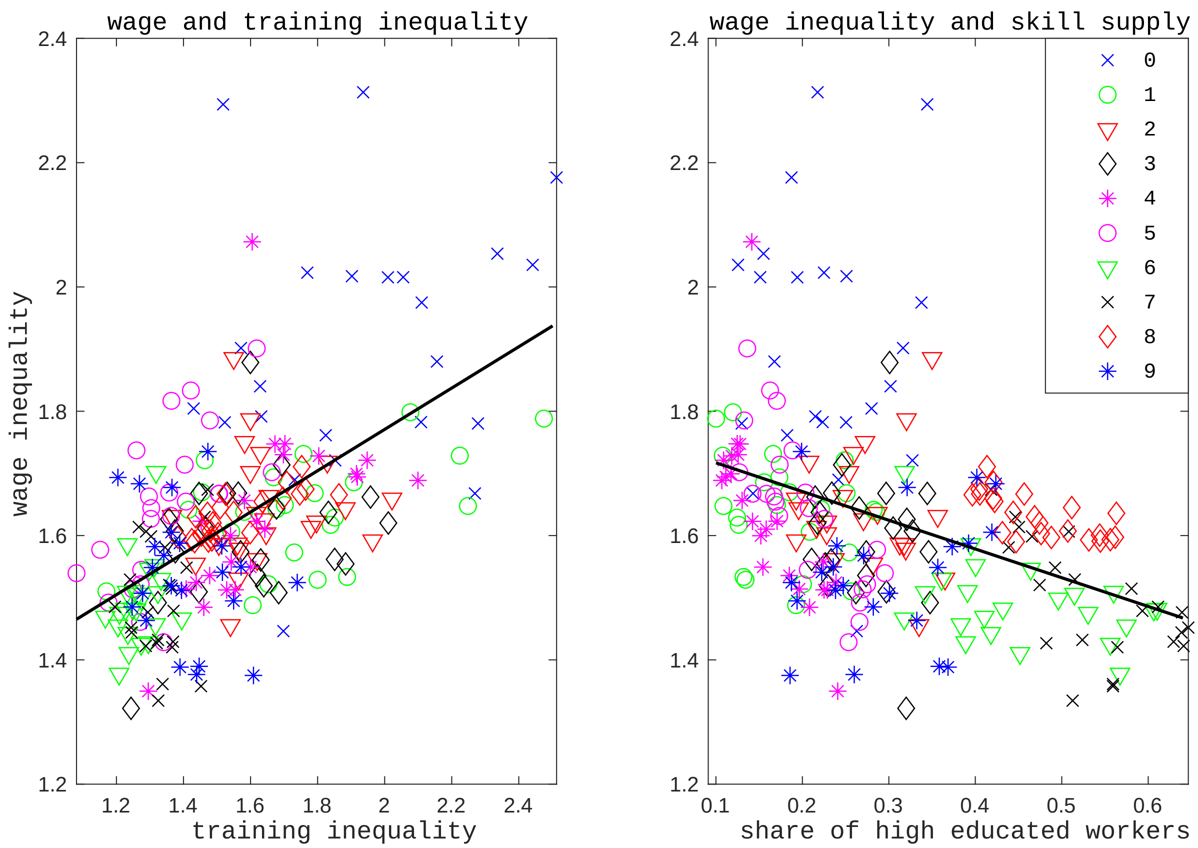


Figure 1.3: Training inequality, skill supply, and wage inequality

the ratio of training participation of skilled workers over that of unskilled workers as the index for training inequality. Next, I compute wage inequality as the ratio of skilled workers' average wage over unskilled workers' average wage. Also, I use the ratio of skilled workers over the total working population to control for skill supply.[41]

[41]Numerous works, e.g. see the literature review by Acemoglu and Autor (2011), have identified the relative supply of high educated workers as one of the main drivers of the wage inequality

I build a panel dataset with quarterly observations and (2 digit) SIC industry averages for the above variables. The panel is composed of 79 industries and, for each of them, it contains an average of 70 observations. I trim the data, dropping observations with less than 25 skilled workers and 25 unskilled workers, to have more robust inference.[42,43] To extend the analysis back in time as much as possible, I employ the conversion code from SIC92 to SIC07 available on Jennifer Smith's web page for the sector classification during the years 1994-2008.[44]

In Figure 1.3, I present two scatter-plots using sector-level averages. On the left hand side, I plot wage inequality over training inequality, and on the other side I plot wage inequality over the relative supply of educated workers.[45] As can be seen, the skill supply is correlated with both wage and training inequality. Also, the data exhibit a clear tendency for sectors that have higher training inequality to have higher wage inequality. To test more formally the relationship between

in the data.

[42]Dearden *et al.* (2006) drop observations based on less than 40 workers, in total I require 50 workers, i.e. 25 per group. For robustness, in Table 1.8, I report estimates based on trimming observations with less than 70 unskilled and 70 skilled workers along with the baseline estimates.

[43]For robustness, I also use yearly averages and 2 digit SIC classification, obtaining similar estimates for the effect of training inequality on wage inequality. Further information, and the code, is available on request. Notice that with the baseline classification, the panel is unbalanced as many sectors do not have enough observations for inference in some periods.

[44]The conversion code is available on *http://www2.warwick.ac.uk/fac/soc/economics/staff/jcsmith/sicmapping/resources/direct/*. I have accessed it on October 22, 2016.

[45]To draw the figure, the data have been collapsed to 1 digit instead of 2 digit SIC sectors.

training and wage inequality, I estimate the following model:

$$\frac{w^s_{j,t}}{w^u_{j,t}} = \alpha_1 + \alpha_2 \frac{p^s_{j,t}}{p^u_{j,t}} + \alpha_3 \frac{n^s_{j,t}}{n^u_{j,t}} + q_t + \gamma_j + \varepsilon_{j,t}, \tag{1.2}$$

where $\frac{w^s_{j,t}}{w^u_{j,t}}$ is the ratio of wages for skilled, $w^s_{j,t}$, to unskilled,$w^u_{j,t}$, employees in period t for sector j; $\frac{p^s_{j,t}}{p^u_{j,t}}$ is the ratio of training participation for skilled, $p^s_{j,t}$, to unskilled, $p^u_{j,t}$, employees; and $\frac{n^s_{j,t}}{n^u_{j,t}}$ is the ratio of skilled, $n^s_{j,t}$, to unskilled, $n^u_{j,t}$, employees in period t and sector j. I always include time dummies, q_t, as control, and, for the random effects estimation, I assume that sector-specific differences, γ_j, are equal to zero. Random effects model is reported for completeness, as the visual inspection of data suggests strong sector-level differences in both training and wage inequality that are addressed only in the fixed effects regression. As expected, fixed effects are the more conservative results.

Table 1.8: Effect of training inequality and skill supply on wage inequality

	baseline			minimum of 70 obs.		
	α_1	α_2	α_3	α_1	α_2	α_3
FE	1.550	0.0578	-0.2289	1.390	0.0851	-0.0257
p-value	0.000	0.059	0.188	0.000	0.148	0.893
RE	1.475	0.0775	-0.0495	1.353	0.1058	-0.0180
p-value	0.000	0.004	0.633	0.000	0.046	0.873

The results of my investigation are reported in Table 1.8. I find that, *ceteris paribus*, a decrease by 1 percentage point in the industry average differential between skilled and unskilled training participation is associated with a decrease of wage inequality by 0.05 percentage points. Arguably, the effect is small but statistically significant across specifications. In chapter two, I provide a theoretical justification for this finding. In particular, the general equilibrium model I build in Chapter 2 suggests that the small size of the coefficient may due to the spillovers of training benefits from one group to the other. I proceed to analyse the business cycle properties of training activities in the UK.

1.5 The training business cycle

This section investigates the business cycle properties of UK on-the-job training.[46] The literature generally assumes training to be countercyclical, see e.g. DeJong and Ingram (2001), Kim and Lee (2007), and Brunello (2009). Empirical evidence suggests that companies invest more on training when they are less competitive than their peers and when the cost of foregone output is relatively smaller. Since about half of training costs are represented by opportunity costs (i.e. trainees' time allocated to training rather than producing), when consumer demand is low, training costs are lower. Thus the theory predicts that firms should be more inclined to provide training. Yet, the theory also provides reasons why

[46]The literature identifies the business cycle as the fluctuations in economic series that have a periodicity of less than 33 quarters (see e.g. Kydland and Prescott (1982) King and Rebelo (1999)).

training should be cyclical. In a recession, the short-term benefits of training are expected to decrease as consumer demand weakens. Whether the first positive effect outweighs the negative one is an open empirical question.

On a practical level, the first issue is to define precisely which training activities are under investigation. Brunello (2009) argues that apprenticeships, or any initial workplace training of long duration, are drastically different from on-the-job training, the latter generally being a short-term activity. He maintains that an economic slowdown affects apprenticeships less intensely than training of senior employees.

In his view, apprenticeships is less influenced by the business cycle because new hires are relatively cheaper and their formation will be completed in about two years. However, the expected duration of the downturn plays a crucial role to the validity of this argument. The counter-cyclicality thesis holds as long as the economy recovers rapidly. If the recession is long-lasting, the returns from both apprenticeships and training activities will be diminished and businesses may cut drastically on training investments.

Evidence for the cyclical behaviour of training is available mostly for the US. Among the many, Einarsson and Marquis (1998) report that their empirical proxy for skill accumulation is negatively related to aggregate output, with a correlation coefficient of -0.187.[47]

With respect to the European countries, Bassanini *et al.* (2007) is again an important source of information. Using a multi-country panel with annual data on training, output gap, and unemployment, they show that training has an elasticity with the output gap which lies between -2.8 and -7% depending on the specification they employ.[48] As part of my research, I test whether on-the-job training is procyclical in the UK economy.

1.5.1 Empirical evidence for the UK

To study the business cycle properties of training for the UK, I employ the information on training reported in the QLFS and macroeconomic time-series available from the ONS website. I seasonally adjust all the series with presumed or evident seasonality with the X-13 ARIMA-SEATS toolkit.[49] For a deeper understanding of UK business cycle I consider also the behaviour of other relevant economic indicators, such as consumption, investments and labour market indicators.

Figure 1.4 shows the cyclical and trend component of training participation and

[47] They also notice that a RBC model, carefully calibrated, produce a much negative correlation between output and human capital investment (very close to -1). The only solution they find is to introduce a random independent process for human capital depreciation. This assumption drastically reduces the correlation between output and training investments.

[48] They obtain similar results when they use the unemployment rate as a proxy for the business cycle. In this case, the coefficient of interest is positive.

[49] Training and labour series show a clear seasonal pattern, together with GDP and gross fixed capital formation.

GDP. It can be noted that training is less volatile than GDP, although the great recession is the main phase when GDP fluctuations are particularly larger than training ones, and that there is no clear correlation between the two series.

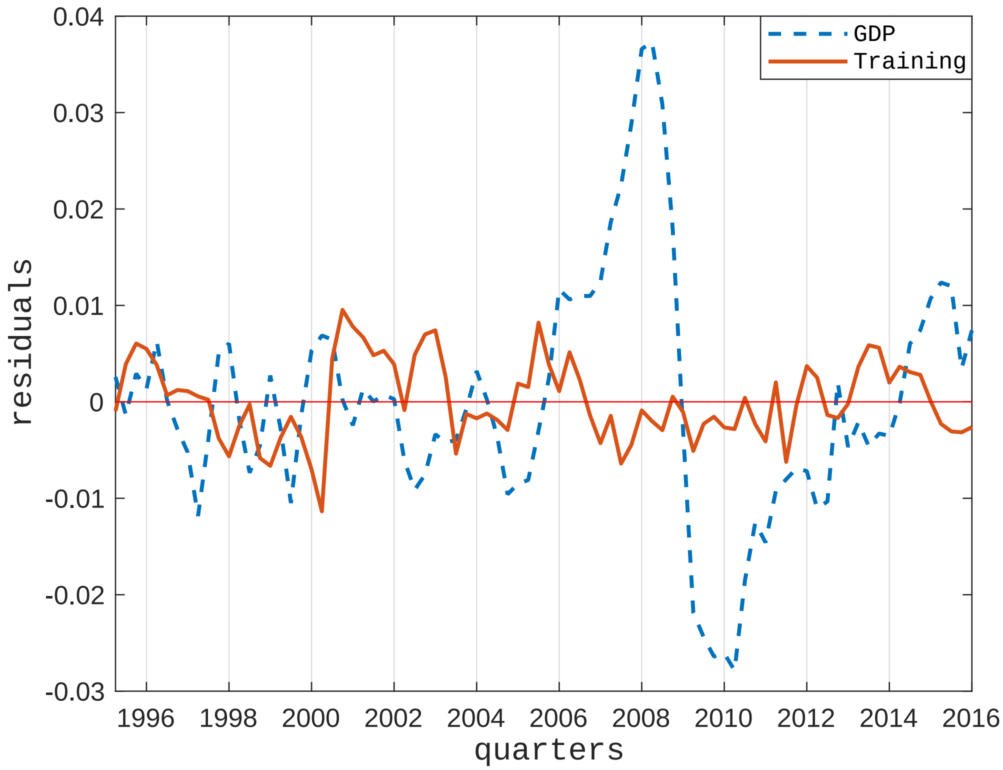


Figure 1.4: Hodrick-Prescott residuals of training and GDP series

I complement the graphical analysis with Table 1.9 which reports the business cycle moments of the UK economy. To this purpose, all data have been detrended with a HP-filter with $\lambda = 1600$ as commonly employed in the RBC literature. The most striking results are that: (i) training is negatively correlated with output, but the correlation coefficient is small; (ii) after seasonally adjusting and detrending data, training has a low standard deviation; and (iii) the skill premium is also almost uncorrelated with output, yet it has a very high variability.

Table 1.9: UK business cycle moments

	std(x_i),%	std(x_i)/$std(y)$,%	corr(y, x_i)	$\rho(x_i)$	N° obs.
output, (y)	1.16	1.00	1.00	0.89	96
consumption	1.06	0.91	0.77	0.86	96
investments	3.89	3.35	0.83	0.73	76
skill premium	3.66	3.15	-0.02	-0.07	84
wage	0.88	0.76	0.08	0.56	83
skilled training, 4w	0.34	0.29	-0.01	0.25	85
unskilled training, 4w	0.22	0.19	-0.05	0.50	85
skilled training, 3m	0.59	0.51	-0.06	0.59	86
unskilled training, 3m	0.39	0.33	-0.10	0.63	86
weekly hours	0.43	0.37	0.44	0.37	95
employment	0.38	0.32	0.69	0.85	96

Interestingly, the business cycle properties of training look unstable over time. To highlight this, I split the sample for estimation in five year intervals and computed the correlation of training and GDP for each sub-sample. As can be observed in Table 1.10, the correlation appears to be stronger and negative during the decade 1995-2005. The correlation is weaker in between 2005 and 2015. This holds true when I consider the training participation rate computed over the last quarter and over the last month. Moreover, using growth rates instead of HP-filtered residuals affects these results only marginally, and not systematically.[50]

[50]The standard deviation, almost constant throughout the estimates, reinforces the idea that the cyclicality of training activities cannot be identified. The formula for the standard deviation

Table 1.10: Correlation between training participation and GDP

	growth rates		HP-residuals		N° obs.
	1995-2005	2005-2015	1995-2005	2005-2015	
training last 4 weeks	-0.01	0.09	-0.24	0.03	42
std	0.158	0.157	0.153	0.158	
training last 3 months	-0.18	0.07	-0.15	-0.05	42
std	0.156	0.158	0.156	0.158	

When considering the business cycle properties of training, one may ask what was the effect of the great recession on training. As discussed above, several authors, e.g. Blatter *et al.* (2009), claim that the length of a crisis has important consequences on the cost-benefit analysis of training provision. A long period of low demand and high financing costs can force firms to reduce training investments. In the next subsection, I verify if this has happened in the UK economy during the great recession (2007-2012).

1.5.2 Training participation during the great recession

At the outset of the great recession, several UK institutions feared a dramatic fall in training participation rates, and several business confidence indexes showed that managers expected a broad reduction in training activities.[51] Yet, the UK

is $\sqrt{\frac{(1-\rho^2)}{(n-2)}}$, where n stands for the number of observations and ρ for the correlation coefficient.

[51] For example, in 2008, the Confederation of British Industry, the Trade Union Congress and other institutions have cosigned an open letter pleading with UK business managers not to cut training during the recession. Both the CBI Industrial Trends Survey and the Quarterly Economic Survey, run by the British Chamber of Commerce, have a section dedicated to training activity expectations. These surveys foretold a drastic reduction of training investments.

National Employer Skill Survey (NESS) figures suggested that most of the establishments intended to maintain pre-crisis training participation levels. Indeed, QLFS data shows that the great recession did not have a significant effect on training rates. As reported by the UKCES report (2013) and by Felstead *et al.* (2013), the overall training participation rate shows a marked long-run declining trend. However, the fall in training has begun in 2002, and it has not been influenced by the great recession.

The literature cited above suggests that the recession has influenced how training is organised and provided to workers rather than affecting the overall training provision. Felstead's interviews of employers and HR managers reveal that they switched towards on-line, or digital, courses, and they organised on-the-job training rather than paying for external off-the-job formation. Many respondents have emphasised the importance of managing more efficiently the resources. According to Felstead and his colleagues, enterprises have training floors, i.e. must-have training activities that, for a reason or another, cannot be remitted. In particular, they argue that six constraints sustained the firms' demand for training: (i) compliance with legal requirements, (ii) operational needs, (iii) skills shortages, (iv) market competition, (v) managerial commitments, and (vi) customer demand needs. Due to these constraints, the great recession did not affect the training participation rate nor the total training expenditure.[52]

1.6 Long run training trends

The long-run trend of training participation features a clear inverse-U pattern. Figure 1.5 shows the time series for the whole UK economy derived from QLFS data. Whilst in the 1990s training participation has been expanding, since the beginning of the new century employers have cut training provision. This trend emerges not only from the data reported here, but from any dataset on UK training. The Centre for Learning and Life Chances in Knowledge Economies and Societies (LLAKES) in 2013 has presented a report showing that as many as 11 different surveys reveal a decrease in training demand. Some of these indexes refer to participation rates, others to training duration. As shown below, the length of training spells has shortened too. Furthermore, the drop is more intense after controlling for sector relocation.[53] This makes the decrease in training even more remarkable.

Green and his colleagues suggest several competing explanations for this decline, each of them only partly supported by the data. An optimistic hypothesis is that

[52] Although the total expenditure hasn't changed during the great recession, it can be observed a decrease in the average cost of training spells. This is in line with the evidence of firms adopting cheaper training types, e.g. growth of on-line training at the expenses of classroom training.

[53] Indeed, workers' relocation has mitigated the drop, as workers moved towards sectors with higher training participation.

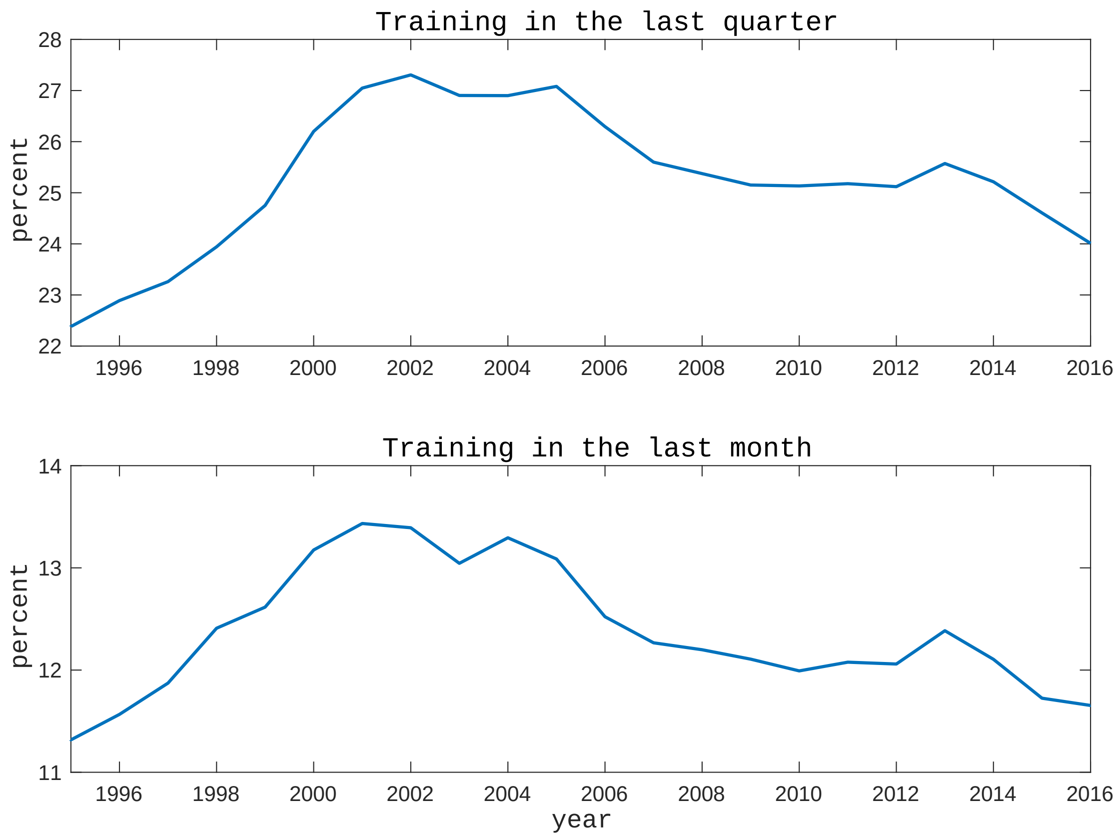


Figure 1.5: Aggregate annual training participation rate

better educated workers require less training as they learn faster. Yet, I have shown earlier that training is complementary to education, so this explanation is weak.[54] Another hypothesis is that in recent times companies provide training more efficiently, and, in particular, managers make better use of informal know-how transfers among workers. In this sense, training is quantitatively less, but is qualitatively unchanged. I have virtually no means to test this hypothesis.[55]

[54]Also, UK PISA scores (a test measuring the ability of 15 yo UK pupils, run every three years since 2000) have remained relatively constant through time (see e.g. Mo (2016)).

[55]A possible test would be to verify whether training has now a higher return to investment.

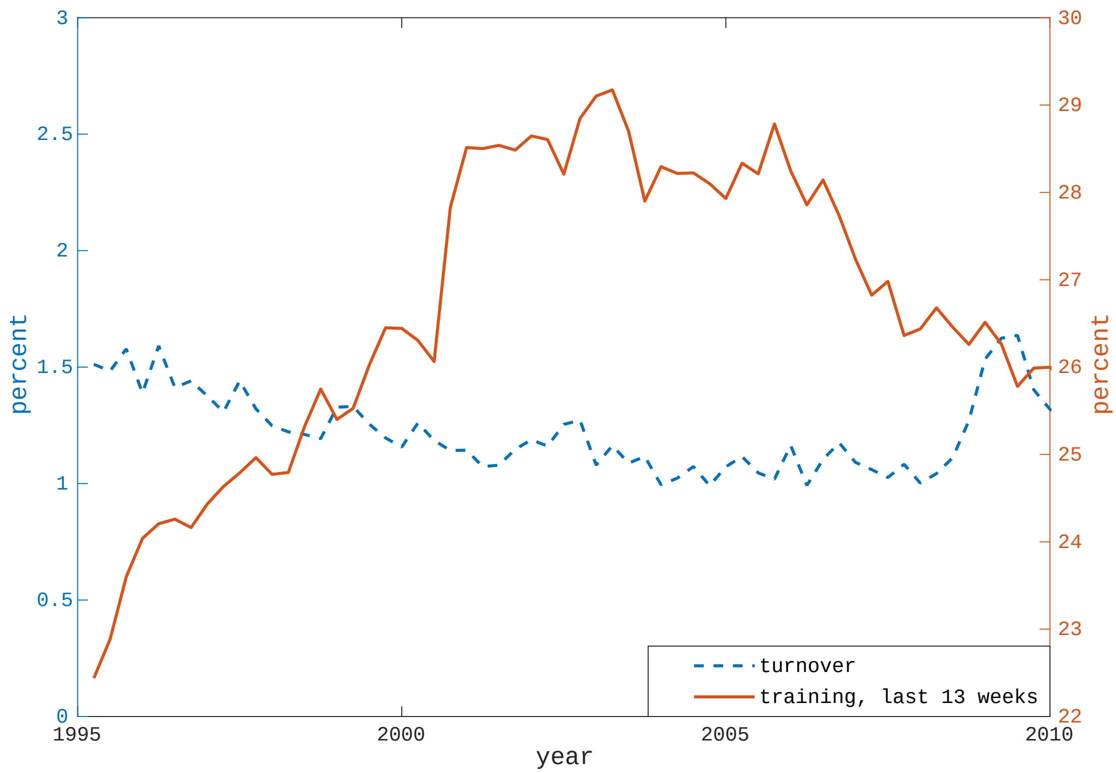


Figure 1.6: Training participation rate and turnover

More pessimistically, managers may have lower expectations about training returns than in the past. They may have revised downward their training targets in reaction to the change in gross returns to training. However, Dearden *et al.* (2006) find no evidence of a decrease of training returns during the previous ten years. Another possibility is that human capital depreciation, due to relocation of workers from one sector to the other, or from an employer to the other, has decreased thanks to a reduction of the turnover rate of UK workforce and an

If, presumably, training is qualitatively the same but quantitatively less, returns to training should have increased in the last decade.

increase in average tenure (see figure 1.6).[56] Against this hypothesis, it must be noted that the decrease of turnover has anticipated the reduction in training participation, moreover, the latter has continued to drop while turnover of workers is relatively constant.[57] In conclusion, the reasons for the reduction in training participation rates are still not fully understood. What is evident is that UK has moved to a lower training participation equilibrium, and this may have important implications in terms of long run economic growth.[58]

1.6.1 The decline in training length

Another relevant indicator of training is the length of training spells. The literature and, in particular, government reports e.g. Green *et al.* (2013), have reported that the reduction in training participation levels has been accompanied by a reduction in the duration of training. I utilise the QLFS data to verify this empirical evidence. The QLFS does not have information about hours, but interviewers ask about the length of a training spell; this is divided into several intervals: (i) less than a week; (ii) up to a month; (iii) up to three months; (iv) less than a year; (v) up to three years; and (vi) ongoing training.

[56] To draw the figure, I employ the flow from employment to unemployment as reported in Smith (2011). The author has published the series on her website.

[57] Cointegration tests between turnover and training have failed to provide evidence in favour of a long run relationship between the two series.

[58] See De la Fuente and Ciccone (2003) for a commentary of the importance of human capital, both from education and life-time training on economic growth.

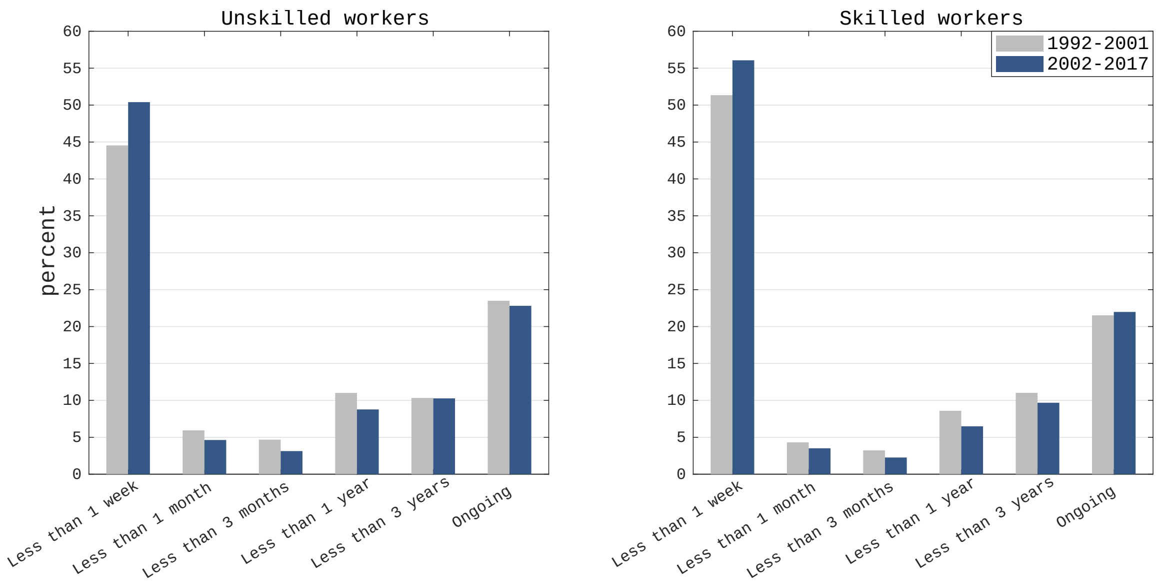


Figure 1.7: Training length by skills

In Figure 1.7, I plot the histograms of training rates by spell duration for the periods 1992-2001 and 2002-2017, and for skilled and unskilled workers. According to the data, training duration has decreased for both skilled and unskilled workers. However, the share of ongoing training has not reduced; it is spells of medium duration that are less frequently observed, while the proportion of workers who report the shortest training interval (less than a week) has increased. Also, it is evident that unskilled training has longer duration than skilled training, both in the past and in more recent times. This may be due to the strong complementarity of formal education and job-related training which the literature has discussed thoroughly, see e.g. Cunha and Heckman (2007). In conclusion, the decline in training length reinforces the evidence of a reduction in the overall provision of training by UK companies that has begun in 2002.

1.6.2 Has training participation declined for all workers?

I contribute to this literature providing empirical evidence that can help understand the trends in training participation by considering the evolution of training for several sub-classes of workers. I classified the QLFS workers by age and education level. With respect to the education, I divide workers into skilled and unskilled ones, while for the age, I define the four following bands: (i) young, who are between 25 and 34 years old; (ii) mid-young, who are between 35 and 44 years old; (iii) mid-old workers, who are between 45 and 54 years old; and (iv) old workers, who are over 55 years old. My intuition is that these statistics should be able to detect relevant changes in the quality of education. In fact, if education has improved, I should observe an intense drop of training participation for the youngest cohort, that anticipates changes for the other groups. The figure shows that this is not the case, as young, mid-young and mid-old workers' training rates comove closely.

Figure 1.8 reports the evolution in time of the training participation rate of skilled workers (to the right) and unskilled workers (to the left). I plot the training participation in the last quarter in the first row and the training participation in the last month in the second row.[59] What emerges from this analysis is that there has been a convergence of training levels between younger and older workers. Around the year 1993, young skilled workers were much more likely to train than older ones, whereas nowadays they are as likely as older workers. With respect to unskilled workers, young individuals are still more likely to train. However, the difference in training participation rates has shrunk remarkably. By comparing unskilled and skilled workers, it can be noted that the training participation rate of unskilled workers, after a temporary increase between the years 1997-2002, has maintained the levels of the 1990s. For skilled workers the decrease of training participation have been more pronounced and it has affected all individuals but those older than 55 years old.

To fully understand the impact of these changes, it must be noted that the average age of UK workers has been slowly growing from about 38 to 41 years old in the last twenty years. Table 1.11 summarises the variation in the age composition of the workforce. The table reports that the share of old workers, who receive much less training than young ones, has increased by 6 percentage

[59] Researches are worried about double-counting training spells when using the information about the last three months. Conversely, they may under-measure training activities when using the information about the training activities in the last month before the interview.

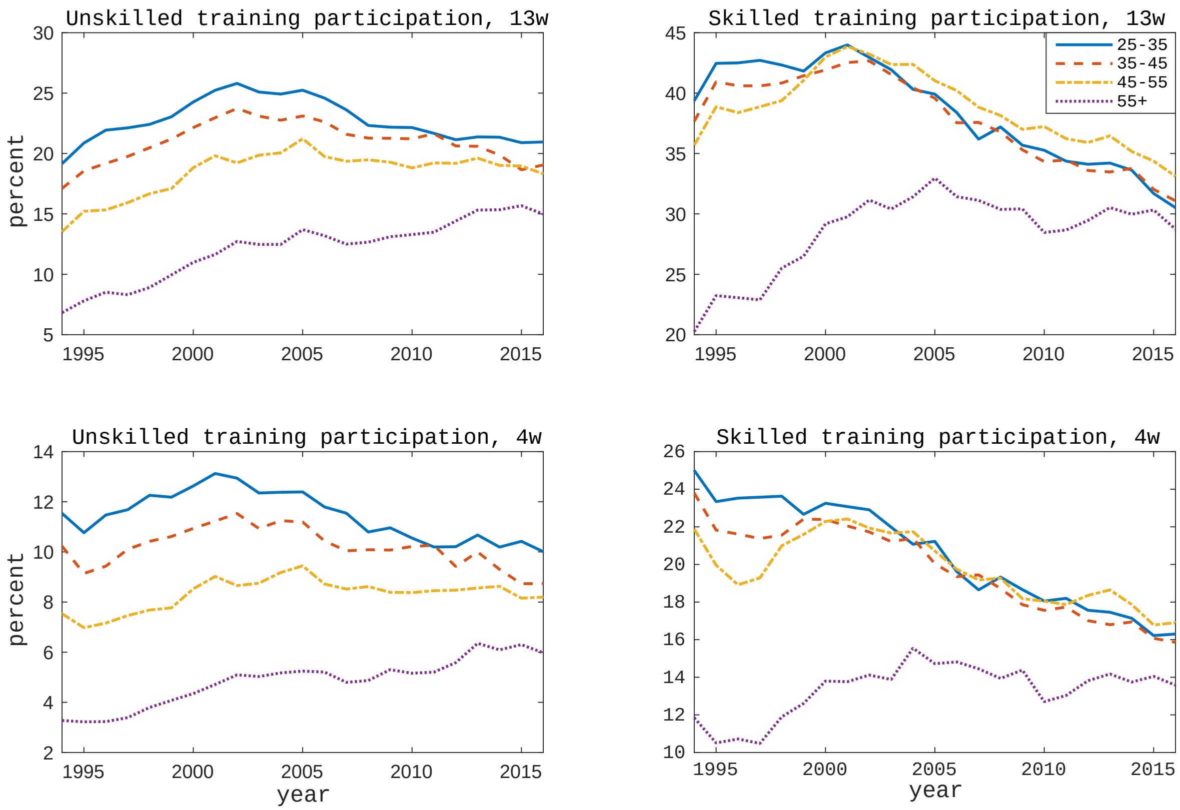


Figure 1.8: Training participation rates by skill and age group

points in the last 20 years. The share of youngest workers has decreased by 16%.[60]

According to these findings, it is plausible that workers' age has played a relevant role on the overall trend in training provision. The demographic trend observed in Table 1.11 suggests that the UK will face a large increase in the share of older workers (age 55+) in the next decade, and it will require sizeable investments to

[60] I exclude workers who are younger than 25 years old to make comparison between unskilled and skilled workers possible, and to avoid including apprenticeships into the on-the-job training statistics. This threshold is commonly employed by the literature I refer to in my thesis.

Table 1.11: Workforce share by age, %

period	age 25-35	age 35-45	age 45-55	age 55+
1994-1996	31	28	26	15
1997-2013	27	29	26	17
2014-2016	26	25	28	21

train these workers. In the next section, I use the Oaxaca-Blinder decomposition method to understand how much of the changes in training participation can be attributed to age and other factors, such as sector relocation.

1.6.2.1 Decomposition of training participation changes

In this section, I analyse more analytically the observed changes in training participation rates. In the QLFS dataset, I generate a dummy variable that distinguishes the observations for the years 1994-2001 from those of the years 2002-2016, i.e before and after the peak of the training participation rate.[61]

This allows to estimate, with *Stata* built-in functions, the Probit model for each period, and to perform an Oaxaca-Blinder decomposition of the changes into endowment and coefficients. The endowment component apprises how much of the difference in training participation can be attributed to changes in the observed characteristic of workers. The coefficient component provides an estimate of how much the change of each coefficient explains the change in participation rates

[61] As a robustness check, I compared and decomposed the years 1994-2000 and 2010-2016, obtaining similar results to what reported in this chapter.

between the old and new group of workers. Any residual, or unexplained, variation is attributed to the interaction term.

More formally, the methodology seeks to decompose the difference between the 2002-2016 and the 1996-2001 average training participation rate:

$$\text{diff} = E\left(t_{02,16}\right) - E\left(t_{96,01}\right) \tag{1.3}$$

based on the assumption that the training participation rate is a linear function of a given set of variables. The latter simply entails that $t_i = X_i'\beta_i + \epsilon_i$, where t_i is training participation, X_i is a vector of firm-specific and worker's characteristics, β_i is a vector of coefficients (including the intercept), and ϵ a normally distributed random error; subscript i indicates each one of the time periods analysed.[62]

Since by assumption $E(\epsilon_i) = 0$ and $E(\beta_i) = \beta_i$, if I substitute the linear functions into equation 1.3, the latter becomes:

$$\text{diff} = \left(X_{02,16}'\beta_{02,16}\right) - \left(X_{96,01}'\beta_{96,01}\right). \tag{1.4}$$

As shown by e.g. Jones and Kelley (1984), it is possible to rearrange the right

[62] Although the variables are the same as in equation 1.1, here the model used is linear. In this context, the linearity assumption holds because I am interested in observing the mean participation rate and not not the individual outcome. Among others, Wooldridge (2016) validates the use of a linear model for dependent variables that are aggregates of individual binary outcomes, such as the employment status.

hand side of equation in a various way.[63] Here, the difference between $E(t_{02,16})$ and $E(t_{96,01})$ is decomposed into three main components:

$$\begin{aligned} \text{diff} = \; & \left[E\left(X_{02,16}\right) - E\left(X_{96,01}\right)\right]' \beta_{96,01} + \\ & + E\left(X_{96,01}\right)' \left(\beta_{02,16} - \beta_{96,01}\right) + \\ & \left[E\left(X_{02,16}\right) - E\left(X_{96,01}\right)\right]' \left(\beta_{02,16} - \beta_{96,01}\right). \end{aligned} \tag{1.5}$$

where the first term relates to differences in observables, i.e. endowments, the second term refers to differences in the coefficients of the linear model, and the last term is the interaction between the first two.[64]

As well known, one of the key assumptions necessary for the decomposition is the conditional independence assumption, also called ignorability of the treatment. This requires that the errors have the same conditional expectation across groups. Clearly, it is very hard to refute or support such assumption when comparing the training probability of people employed in two different periods. Nonetheless, the exercise is still informative and helps summarising the main changes in the provision of training.

In Table 1.12, I report the summary results from the decomposition exercise. In particular, I observe a small difference between the average training participation

[63] For example, many works in the discrimination literature, Cotton (1988) among them, employ a twofold decomposition to distinguish explained from unexplained (or residual) differences between two groups.

[64] For a longer treatment of Oaxaca-Blinder decomposition see e.g. Jann (2008).

Table 1.12: Summary of training participation decomposition

variable	percentage points
training rate, 2002-2016	27.9
training rate, 1996-2001	27.2
difference	0.74
endowments	2.12
coefficients	-1.04
interaction	-0.34
N° of obs.	2,956,909

rate of workers in the period 1994-2001 and that of workers in the period 2002-2016, the latter being slightly greater than the former. Yet, this is not due to training being provided the same way as in the past, but it is the consequence of how endowment, coefficients, and other unobservable factors have changed.

The positive value reported in 'endowments' suggests that workers have moved towards jobs that require more training and that they have developed characteristics associated with higher training participation, e.g. more people are skilled or more of them work in large companies. Similarly, the positive contribution of 'coefficients' entails that characteristics leading to higher training in the past have now a stronger effect on the probability to train, e.g. being skilled, or working for a large firm, implies higher training probability than in the past.[65]

To better understand these trends, I report further information in Table 1.13,

[65] Alternatively, it can be the case that the characteristics leading to lower training now have a smaller effect than in the past.

Table 1.13: Decomposition of training participation before and after 2002

variables	(1) endowments	(2) coefficients	(3) interaction
skilled	1.0	-1.8	-0.5
married	0.0	-1.4	0.0
female	0.0	-0.9	0.0
temporary	0.0	0.1	0.0
part-time	0.0	1.1	0.0
region	0.1	0.4	-0.4
quarter	0.0	0.0	0.0
tenure	-0.1	-2.9	0.0
occupation	0.9	-0.9	-0.5
size	0.0	-1.8	0.0
age	-0.8	30.8	0.7
industry	1.0	1.3	0.3
constant	-	-25.0	-

where the detailed results of the decomposition are recorded. From Table 1.13, first column (endowments), it can be observed that age contributes negatively to the difference in training participation rates, as average age has increased, and older workers train less. On the other side, relocation to training intensive industries and occupations raised training participation among the workforce, as the increase in University education did. The third column, where the effects of the interaction terms are reported, is marginally relevant.

The fact that the set of quarterly dummies does not contribute to the differences in training participation between the two periods (as all entries are close to zero), does not mean that they do not have an effect on the probability of training. The correct interpretation is that the effect of seasonality on training has not changed from one period to the other.

The most interesting pattern emerges from the second column, and in particular from the coefficients of the constant term and age. Both coefficients have changed drastically, and their contribution to difference in training participation rate is very large. The overall effect is positive; however, this hides the fact that workers start with a much lower training probability, and the latter does not decrease with age (or it does so but much more slowly than before). As a consequence, the distribution of training has changed in favour of senior workers and at the expenses of the youngest cohort of workers.

Despite being relevant, so far the literature has totally neglected this aspect. My findings entail that any explanation of the aggregate training participation rate needs to be consistent with age-dependent changes in training. Education quality is unlikely to explain these trends. My hypothesis is that (i) recent technological changes demanded more training investments on older workers, and that (ii) institutional changes, e.g. the increase in retirement age and the higher share of old workers in the active population, have driven up the demand for training activities of this age cohort.

The question that remains unanswered is whether the decline of training for the workers who are less than 40 years old reflects a genuine decline in training needs or it reflects negative developments in the economy. O'Mahony (2012), using industry data for 15 different European countries, argues that training positively contributes to yearly GDP growth for about 6 basis points, thus answering this question can have important economic implications.[66] If training is being under-provided, it is crucial to understand what factors have driven training participation down. To this end, it is necessary a clear definition of training activities which discerns them by type, duration, and quality.

As a complement of the previous analysis, the next section performs the same decomposition exercise with respect to skilled and unskilled workers to verify the presence of patterns in the data that may be hidden by aggregation.

1.6.2.2 Decomposition of training participation by skills

In this section, I use the Oaxaca-Blinder decomposition to assess the sources of differences in training participation between unskilled and skilled workers. I run the decomposition for each of the two periods considered in the previous section, i.e. before and after the observed peak in training participation.[67] By comparing

[66] During the period 2001-2007, the average GDP growth was about 1.9% for his sample of countries. Thus, training activities contributed to 3% of the observed GDP growth. Most of the observed economic growth can be attributed to the increase in physical capital stock and hours worked.

[67] As robustness check, I run the decomposition with the observations from the years 2010-2017 only. Selecting a shorter and more recent period as a comparison to the decomposition from the years 1994-2001 provides a more nuanced picture (e.g. training inequality has reduced in the last 7 years but it has not reduced if I look at the whole period 2002-2017). Nonetheless,

the output of the two decompositions, it is possible to highlight what factors were important for training inequality in the past, and what matter most nowadays.

Table 1.14: Summary of training participation decomposition by skills

	1994-2001	2002-2017
unskilled training participation	20.93	21.69
skilled training participation	42.74	37.38
difference	-21.80	-15.68
endowments	-11.04	-9.30
coefficients	-8.04	-5.96
interaction	-2.73	-0.42
N° of obs.	732,041	2,238,618

The decompositions reported in Table 1.14 suggest that the training participation rate of unskilled workers was 21.8 percentage points lower than that of skilled workers during the period 1995-2001. This gap decreased to 15.7 percentage points in 2002-2017. A finer decomposition, i.e. over smaller sub-periods, would show a non-monotonic evolution of training inequality. Indeed, looking back at Figure 1.5, the peak of training inequality matches broadly with the period when the training participation rate was the highest, i.e. around the year 2002.[68] Nonetheless, some clear trends can be observed in the data through the decomposition exercise. Table 1.15 attempts to disentangle the main (observable) characteristics

the main conclusions drawn here are not affected by the choice of how to break the dataset into sub-periods.

[68] The time series for training inequality can be found in Chapter 2, Figure 1.3.

of workers, and jobs, that contribute to training inequality.

Table 1.15: Decomposition of training participation by skills

	- - - - 1994-2001 - - - -			- - - - 2002-2017 - - - -		
	(1)	(2)	(3)	(4)	(5)	(6)
variables	endowments	coefficients	interaction	endowments	coefficients	interaction
married	-0.1	1.5	0.0	0.0	1.6	0.0
female	0.1	-2.4	0.0	-0.1	-1.6	0.0
temporary	0.2	0.4	-0.1	0.1	0.1	0.0
part-time	-0.8	0.2	0.1	-0.4	-0.1	0.0
white	0.0	0.2	0.0	0.0	-0.5	0.0
pub	-1.6	-2.6	0.2	-1.3	-2.0	0.1
quarter	0.0	0.0	0.0	0.0	0.0	0.0
region	0.3	0.0	0.0	0.1	0.0	0.0
tenure	0.0	-1.4	0.0	-0.1	-1.6	0.0
occupation	-4.7	4.1	-2.2	-4.2	1.0	-0.2
size	-0.1	0.8	-0.3	-0.2	0.4	-0.1
age	-0.8	6.2	-0.2	-0.3	-6.4	-0.1
industry	-3.5	-1.0	-0.1	-3.1	-0.8	0.0
constant	-	-14.0	-	-	3.9	-

Before the year 2002, differences in occupation and industry between skilled and unskilled workers (column one) explain as much as 8 percentage points of the training gap. This entails that workers were sorted according to their education to different occupations and industries. Unskilled employees worked in sectors with less training needs, while skilled employees worked in sectors with more training needs. In the period 2002-2017, this sorting has become slightly less effective (column four).[69]

[69] The reduction of inequality due to sorting is more pronounced over the period 2010-2017, as observed in the robustness checks available on request.

Comparing column two with column five, it is possible to observe that sectors training needs of skilled and unskilled workers have become more similar within the same industry and sector. In fact, the contribution of these coefficients is closer to zero (in absolute value).[70]

Other differences in either endowments or coefficients are small, and generally close to zero, except for the constant and for the age coefficients. In fact, the difference between the constant term between skilled and unskilled training contributes by 14 percentage points to the training gap in the period 1994-2001. This is two thirds of the whole training gap reported in Table 1.14. The change in this component from -14 to 3.9 in the period 2002-2017 (column five) entails that nowadays unskilled workers are more likely to train than skilled ones. However, the change is largely compensated by the variation in the age coefficients. As a consequence, I only observe a modest reduction in training inequality.

For both unskilled and skilled workers training is a monotonic decreasing function of age Therefore, the fact that the difference in age coefficients contributes to reducing training inequality during the period 1994-2001 implies that skilled workers have a steeper age-training profile than unskilled workers. The opposite is

[70] Thanks to the normalization of categorical variables (region, occupation, and industry), the choice of the baseline category does not affect the interpretation of the coefficients. Thus, for example, sector's coefficients do not represent the difference from the participation rate of the baseline sector, but from the average training participation rate among sectors.

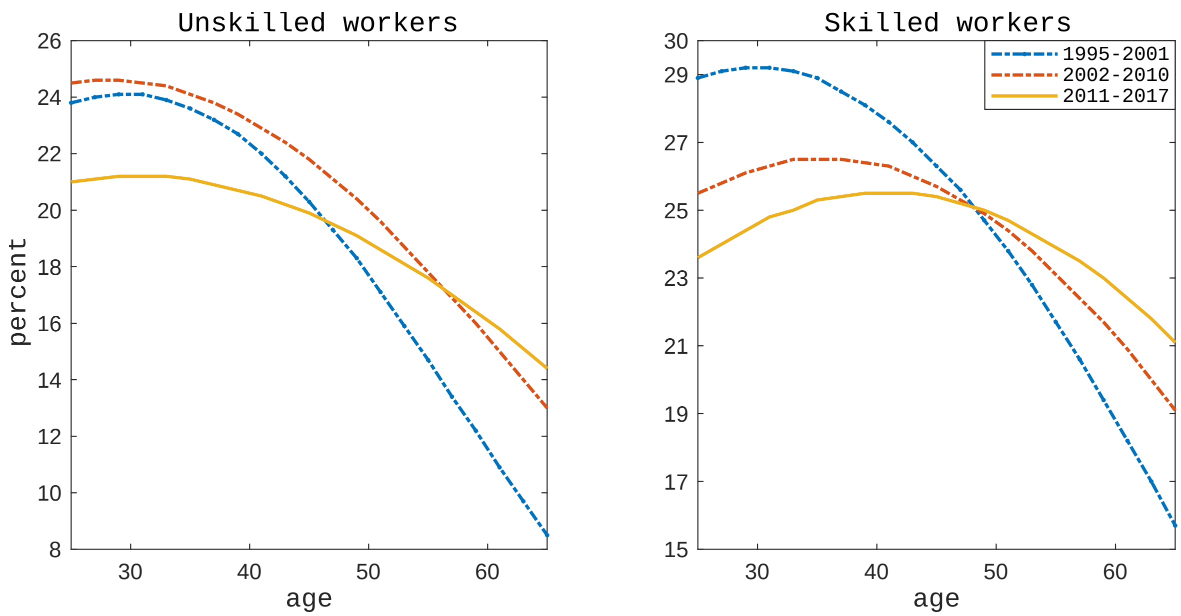


Figure 1.9: Training probability over the work life

true for the period 2002-2016, where the training probability of unskilled workers decreases faster with age than that of skilled workers. The best way to appreciate the drastic change in training provision is to look at the predicted training probabilities, conditional on the worker's age.

Figure 1.9 shows the age-training profile for the two skill groups and three different sub-periods. For unskilled workers, the transition from the first to the middle period looks like a upward parallel movement, while in the most recent period training has been redistributed from young generations to older ones. For skilled workers, the transition looks more gradual and it features a large reduction of training and a redistribution towards older workers.

To conclude, this exercise provides further evidence that: (i) assortative matching explain parts of the observed difference in training between skilled and unskilled workers; (ii) age-related changes have affected skilled and unskilled workers unevenly; and, finally, (iii) there are factors affecting training inequality (at least with respect to educational attainment) that are uncorrelated with observable characteristics of workers and standard job characteristics such as firm's size and sector.

Perhaps other factors, such as technological progress in the supply of training and shifts in the type of training demanded by firms, are responsible of the patterns observed with available data. Unfortunately, the lack of detailed information about training activities and their costs represents a strong impediment to identifying the drivers of the evolution of UK training provision.

1.7 Conclusion

This chapter combines empirical evidence job-related training and theoretical contributions to provide an accurate representation of on-the-job training activities in the UK. I demonstrate the importance of training activities in supporting economic activity, and, to this end, I review the literature estimating returns to training for workers and companies and link these results to stylised facts about

UK labour market.

I provide evidence that training provision is acyclical for the UK. This result suggests that indirect costs of training spells do not influence firms willingness to invest on these activities, at least in the short run. It is however more challenging to understand the long run trend of UK training participation rate. The literature has not provided a convincing explanation for trend observed in UK data. In this regard, my contribution is that I provide new evidence about training patterns within the workforce. By looking at specific subgroups of workers, I exclude some of the proposed explanations. For example, I rule out that changes in the education system or changes in the technology may be the main cause for the observed decline in training.

Another key contribution of Chapter 1 is the findings about the distribution of training across age groups. In the last ten years, training participation has declined for all age groups except for workers in the age band 55 to 65 years old. A puzzle arises from the evidence I collect, since economic theory predicts no human capital investments for workers who are close to retirement. A potential research venue for the future is to investigate what led companies to shift training towards this category of workers.

I find that University educated workers were twice as likely to get trained as non-

University educated workers in the 1990s and that training inequality is correlated to wage inequality. Although both wage and training inequality have reduced in the last ten years, I maintain that governments should consider fiscal subsidies to training as a policy that supports and integrates other interventions aimed at reducing labour income and wage inequality between University educated and non-University educated workers.

The rest of the thesis focuses on the relationship between training and wage inequality and the scope for fiscal policy.

”Labor looks different in the 21st century. And so should our job training programs.”

Leila Janah, businesswoman

2

A general equilibrium model with training

The work presented here draws from and enriches the working paper Angelopoulos *et al.* (2017).

2.1 Introduction

As Chapter 1 shows, firm-provided training is an economically significant activity in the UK. According to data from the Quarterly Labour Force Survey (QLFS), every quarter, about one-fourth of workers receives some training. Empirical literature, including both academic and policy-related research, has examined the importance and the characteristics of employee training. The reviewed 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), Konings and Vanormelingen (2015) and other works cited earlier). 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, I speculate that, *ceteris paribus*, the latter could contribute to increasing earnings for low-skill workers and to reducing labour income inequality between groups of workers with different skills. A key observation relating to the UK labour market since the

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

1980s is the existence of pronounced earnings and wage inequality accompanied by a stagnation of wages for the lower income groups since mid-2000s. Among others, Blundell and Etheridge (2010), Brewer and Wren-Lewis (2016), Belfield *et al.* (2017) and Angelopoulos *et al.* (2017a)) have reported such trends in inequality measures. An important dimension of inequality in many developed 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] Thus, persistently lower labour market returns for workers with lower skills characterise the UK economy.

Enhancing the skills and productivity of those with lower education is one of the possible interventions to boost earnings for the low skilled and to reduce inequality by closing the gap from the lower end. One way to achieve this result is by improving the quality of basic education, see e.g. Machin and Vignoles (2005), Wößmann and Schütz (2006), and Autor (2014). Academics and policy-makers have considered complementing such efforts by interventions to improve the skills and productivity of workers already in the labour market. These have been applied

[2] In the next section, data for the UK are presented and discussed in more detail.

to individuals with high school degrees through ongoing vocational training and lifelong learning schemes (see e.g. Stevens (1999), Sofer (2004), Brunello *et al.* (2007) and the European 2020 Strategy).[3] Thus, more intensive training of the unskilled, or 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), Brunello *et al.* (2007) and Busemeyer (2014)).[4] However, extensive evidence suggests that there is inequality in training, i.e. more skilled workers are offered more training opportunities. 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 backgrounds. Similarly, in Chapter 1, data from the UK QLFS 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. The analysis in

[3]Note that here I refer to training of employees only, as opposed to training of unemployed workers or individuals leaving the labour force to study. The literature reviewed in Chapter 1 has found that this second type of training activities is generally not effective.

[4]In contrast, policy interventions to encourage training for efficiency reasons are more controversial as part of the literature argues that training under-provision as a result of market failures is hard to estimate (see e.g. Bassanini *et al.* (2007) and Brunello and de Paola (2009)).

Section 2.2.1, based on 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 expected to have a small impact on wage inequality.

It is hard to fully understand how job-related training determines its subsequent effects on wage growth and wage inequality. This complexity is mainly due to the fact that job-related training takes place at the expense of work time, thus it is largely dependent on firm's choices and is affected by the structure of production and by changes in other inputs. In particular, a firm's decision to train its employees depends 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 workers' skills, which in turn depends on the structure of production (e.g. capital-skill complementarity and skill-unskilled substitutability).

In such a context, the government can 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 my data analysis below, I show that UK

[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 still employing this scheme.

firms that receive a higher financial training subsidy, train a higher proportion of their employees.

In light of the above, I aim to evaluate the quantitative implications of policies that increase employer's incentives to train workers. I construct a dynamic general equilibrium (GDE) model that is consistent with the main features of job-related training and wage inequality outlined above and allows for the relevant policy interventions. I focus on the effects of such policies on inequality in training, skill and wages. To model job-related training and skill creation, I 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. Some of these works are 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 the worker time that is not diverted to training are combined to create the effective labour input.

What defines this form of skill acquisition as job-related training is that, in my 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, consistently with the evidence discussed in the previous chapter. While this approach adds complexity to the problem of the firm by making it intertemporal, it is more in line with the empirical evidence. According to the evidence, firms, and not workers, cover the training costs. 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.

I 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

[6]Since I focus on wage (and not wealth or income) inequality, I 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 same household.

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 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.

I calibrate the model to data from the UK and ensure that it generates training and wage inequality that are consistent with the data. I then evaluate policies that target training for the unskilled workers by subsidising the firms and reducing the relevant financial cost. 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, the model predictions are consistent with the empirical evidence collected. With respect to both relationships, the model predicts effects just below the lower bound of my estimates. Yet, 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. My work is thus 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 fiscal multipliers that are greater than unity. What underlies these significant effects is: (i) the strong impact of training on returns to labour, and (ii) the spillover effects that work, in general equilibrium, to enhance the positive effect on any labour input.[7]

From one side, subsidies to increase job-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. Hence, the policy-maker faces a trade-off to be addressed by weighting the different objectives.

The result of this exercise is far from being negative. Even though subsidising job-related training does not have a big impact in reducing "inequality", lower skilled are benefited from the intervention. Furthermore, the positive spillovers entail that unskilled training subsidies support the productivity of both skilled and unskilled workers. Thus, it is only because skilled workers benefit indirectly from the policy that the income (and wage) gap is reduced by a small amount. In other words, the conclusion that training subsidies do not significantly reduce inequality should be viewed as a welcoming consequence of the positive spillovers that they create.

[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).

Chapter 2 presents a model that is a specific case of the more general case developed in Chapter 3. This may raise a question about the value added of the former. However, under a direct comparison of the results, it is possible to observe that Chapter 2 allows to better appreciate the channels through which a subsidy to training affects wages of both unskilled and skilled workers and the whole economy. In fact, the added complexity of Chapter 3 model will hide some of the results highlighted here e.g. the spillovers effects.

The rest of this chapter is organised as follows. In Section 2.2, I 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 2.3, I develop the model used for the quantitative evaluation of the nexus between training, inequality and policy, and I discuss its calibration and quantitative relevance. Section 2.4 shows the effects of a positive innovation to total factor productivity. Then, in Section 2.5, I evaluate the effects of policy aimed at mitigating inequality of training outcomes by incentivising job-related training. Section 2.6 contains the conclusions.

2.2 Training costs, returns and inequality

This section reviews and adds to the empirical evidence on the extent of job-related training, its importance for employers and employees as well as its effect on wage inequality. Job-related training refers to all training activities which individuals who are in employment, i.e. workers, participate in.

The sub-plot (1,1) of Figure 2.1 shows workers' participation in this type of training in the UK, using quarterly data from the QLFS from 1995.1 to 2015.4.[8] Training participation is measured as 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]

As extensively discussed in the first chapter, 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),

[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 similarly high (see, e.g. Markowitsch *et al.* (2013) who use the CVTS dataset from Eurostat).

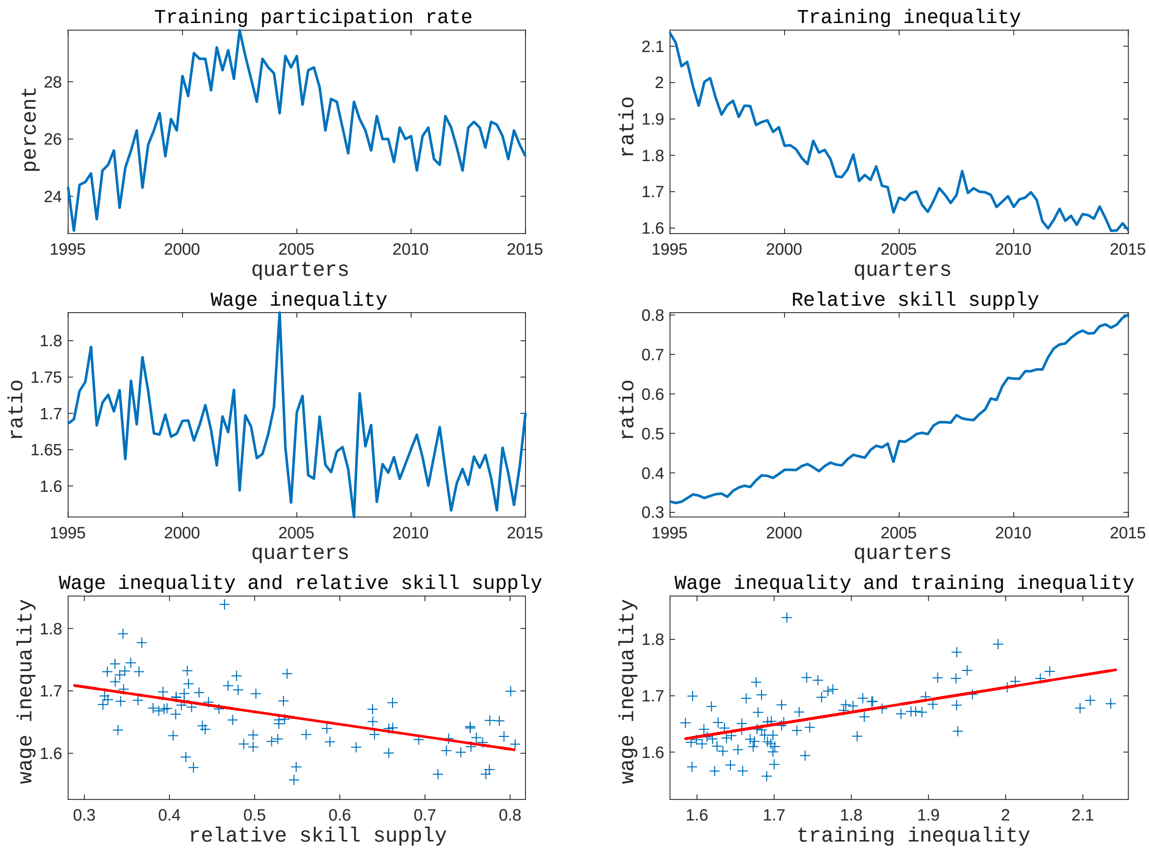


Figure 2.1: Stylised facts

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 they are more robustly significant (see, e.g. Hansson (2008)). A positive effect of training on productivity is also confirmed in studies for the UK, such as Dearden *et al.* (2006).[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, workers are expected to finance their training activities (see e.g. Becker (1962)), especially if firms engage in poaching of employees from competitors. However, there is empirical evidence, at least for the US, that firms support employee training despite poaching (see e.g. Parent (1999)).[11] As reported earlier, the data for the UK shows significant firm-sponsored training activities, which are accompanied by high returns for firms from training their employees; this suggests that training has both a general and firm-specific component, and that firms can capture a great share of the returns from training.

Data from the QLFS, CVTS, and ESS (see Chapter 1) 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. In recent years, UK Government subsidies have covered approximately 2.4% of total training costs according to CVTS data (2005 and 2010)).[12]

[10]See again Chapter 1, Section 1.2.4, for more detailed information.

[11]The role of poaching will be discussed more extensively in Chapter 3.

[12]Note that this ratio is based on gross receipts over total training costs as reported by surveyed firms, averaged across the two years available. It might be 1-2 percentage points larger if I were to include tax deduction of training expenses, but, since deductions are possible only

The importance of firms' contribution to training expenses is also confirmed using European-level data. In particular, Bassanini *et al.* (2007), analysing European data from the European Community Household Panel, find 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.

As shown earlier, 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. For example, Bassanini *et al.* (2007)) report that, in most EU countries, there is a gap in training participation between workers of different education levels and of different family background (see Chapter 1, section 1.3.1). Also, they find that training increases wages more for better educated workers. I further elaborate on training inequality and the relationship between training and wage inequality in the UK in the following sub-sections.

---

if the firm makes profits, for simplicity I did not consider this component.

2.2.1 Inequality in training and wage inequality

To analyse training and wages inequality, I use data from the UK QLFS on training, wages, employment and hours of work by education groups between 1995.1-2015.4. I split the sample into the group of workers who have at least a bachelor's degree or high level qualification (University educated) and those without these qualifications (non-University educated). For simplicity I call them skilled and unskilled, respectively. For each education group, I 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, I calculate the ratio of the University educated to non-University educated training participation rates (see sub-plot (1,2) in Figure 2.1). 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, I expect training inequality to contribute to wage inequality. Although this is a plausible speculation, I am 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 sub-plot (2,1) of Figure 2.1 (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).[13] This can be linked to increased University education, which implies that the relative supply of skilled labour over unskilled has grown, as shown in sub-plot (2,2) of Figure 2.1.[14,15] Indeed, as the scatterplot in sub-plot (3,1) of the same figure 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 sub-plot (3,2) of Figure 2.1 shows.

It is interesting to note that the trend in wage inequality is more consistent with the trend in training inequality. In particular, the biggest reduction in wage inequality took place between 1995-2004, the period where training inequality also reduced significantly, whereas after 2005 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

[13]The skill premium is the ratio of the average skilled to the unskilled wage over the period 1995.1-2015.4. All working individuals who are between 25 and 65 years old are considered workers. Hence, this definition include part-timers and public sector employees. The wage is computed by dividing weekly labour income by the number of hours worked per week from the main job.

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

[15]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.

inequality, I regress the former on the latter and on the relative supply of skilled to unskilled labour. In particular, I consider the following relationship:

$$\frac{w_t^s}{w_t^u} = \alpha_1 + \alpha_2 \frac{p_t^s}{p_t^u} + \alpha_3 \frac{n_t^s}{n_t^u} + \sum_{i=1}^{3} \gamma_{it} Q_{it} + \varepsilon_t, \tag{2.1}$$

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 in the UK economy.[16] Given that training exhibits quarterly regular variation (see e.g. Felstead *et al.* (2013)), I include a set of quarterly time dummies, i.e. $\sum_{i=1}^{3} \gamma_{it} Q_{it}$. Finally, $\varepsilon_t \sim Niid(0, \sigma_\varepsilon^2)$ is the error term.

Table 2.1: Wage and training inequality in the UK economy

	$\widehat{\alpha}_1$	$\widehat{\alpha}_2$	$\widehat{\alpha}_3$	$\widehat{\gamma}_{1t} = \widehat{\gamma}_{2t} = \widehat{\gamma}_{3t} = 0$		serial correlation	
estimates	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

The results for the coefficients of interest are reported in Table 2.1 and support my hypothesis that training inequality is correlated with wage inequality.[17] I

[16]For the present analysis I use the time series for the whole UK. Thus, the methodology differs from the panel data analysis reported in Section 1.4. Yet, the two approaches lead to consistent results.

[17]As already remarked, this relationship has been tested on a panel dataset composed of yearly observations for 1-digit SIC UK sectors in Section 1.4. The results of the analysis, reported in

also report an F-statistic for the joint significance of the three quarterly time dummies. Finally, I 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, the coefficients for p_t^s/p_t^u and for n_t^s/n_t^u are significant at the 2% and 7.8% levels respectively, and they both have the expected signs. The results suggest that training inequality is positively related to wage inequality, even after controlling for improvements in the education composition of the labour force. Furthermore, the 95% confidence interval for $\widehat{\alpha}_2$ ranges from 0.022 to 0.249.

2.2.2 Cost-incentives matter

As discussed above, UK firms take on most of the training costs, and the government contributions to the monetary costs are small. At this point, it is relevant to know whether the decision to train employees is sensitive to subsidies to the direct monetary costs that job-related training entails. I am not aware of existing evidence on the link between training subsidies and training participation. 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

---

Table 1.8, are consistent with the results of the time-series regression performed on the whole UK economy and reported here. The coefficient $\widehat{\alpha}_2$ in that case was twice as large, suggesting that the aggregation employed here partly hides the relationship between training and wage inequality as sector level differences are netted out.

Gambarotto (2004)), market power (e.g. Bilanakos *et al.* (2017)), and size (e.g. Almeida and Aterido (2015)). To the best of my knowledge, none of these works have been able to control or study the effect of fiscal incentives, due to lack of data. Thus, I 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 UK companies.

I 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.[18] Due to changes in the classification, I can only match 17 sectors between the two waves. Thus, to make best use of the available data, I merge them into an unbalanced panel dataset. I 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, I 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

[18]Note that discontinuities in the sampling and classification prevent the use of the earlier editions of the CVTS for the present analysis.

regression below.

I finally combine the sectoral QLFS training participation rate data with the corresponding sectors in the CVTS database. In some cases, I 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.

Exploiting the cross-sectional and time-series dimensions in the sample discussed above, I undertake a panel data random effects analysis, estimating the following model:

$$s_{i,t} = \beta_1 + \beta_2 sub_{i,t} + \beta_3 size_{i,t} + \mu_{i,t}, \tag{2.2}$$

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 $t = 2005, 2010$. Given that sectors with bigger firms may train a higher share of their employees to exploit economies of scale in training provision, I also include the average number of employees per firm, $size_{i,t}$, in the model. Finally, I allow for further unobserved sector heterogeneity captured by the error term and consistent with a random effects specification.

The results from estimating (2.2) are shown in Table 2.2. As can be seen, both $\hat{\beta}_2$ and $\hat{\beta}_3$ are positive and significant. Moreover, the results of the Hausman test indicate that the random effects assumption cannot be rejected in favour of the fixed effects one. 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.[19]

Table 2.2: 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	0.000	0.000	0.000	p-value	0.381

Unfortunately, having at disposition only two waves of data prevents me from using more complicated econometric models. In particular, I am unable to make claims about causality between subsidies and training. Yet, I can exclude the hypothesis that subsidies crowd out private investments, i.e. firms' expenditures in training activities.[20] In the next section, I introduce my theoretical model that

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

[20]The main confounding factor would be strong externalities, or spillovers, between one group and the other. As I show with my model's simulation, spillovers are likely to dampen the link from training to wage inequality. A simple example may provide the basic intuition: once skilled employees get trained, overall productivity of the training company increases, and this leads to

replicates quantitatively the empirical results just presented.

2.3 The model

To evaluate the quantitative implications of policies that raise firms' incentives to train low (high) skilled workers, I 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 I analyse the quantitative implications of policies, I focus on their effects on inequality in training, skill accumulation and wages. In particular, I 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, I first solve the model at the steady-state, choosing the parameters so that the steady-state is similar to the actual UK economy. I then take this as the initial position of the economy and evaluate the effects of one-off, permanent,

relatively higher wages for the group of unskilled employees. Revisit Section 1.2.4.3 for more empirical evidence of human capital (and training) spillovers.

change in policy by simulating the economy as it converges to its new long-run equilibrium.

2.3.1 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. I assume that household members share the household labour and asset income and have equal consumption irrespective of their labour market status as assumed in large part of the literature on unemployment since Merz (1995). This allows me to focus on between-groups wage inequality without additional modelling assumptions required to enrich the setup with wealth and consumption inequality. In this context, the head of the household makes all choices on behalf of its members, differentiated in this case, by labour market skills. In particular, the head of the household maximises discounted lifetime utility:

$$U = \sum_{t=0}^{\infty} \beta^t \frac{\left\{c_t^{\psi_1} \left[n^u \left(1 - l_t^u\right)\right]^{\psi_2} \left[n^s \left(1 - l_t^s\right)\right]^{\psi_3}\right\}^{(1-\sigma)}}{1-\sigma}, \tag{2.3}$$

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, $\left(1 - l_t^u\right)$, and skilled leisure, $\left(1 - l_t^s\right)$ in utility respectively.

The household's budget constraint is:

$$c_t + k_{t+1} - \left(1 - \delta^k\right) k_t = n^u w_t^u l_t^u + n^s w_t^s l_t^s + r_t k_t + \pi_t - T_t, \tag{2.4}$$

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. 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 clearer when I examine the production side of the model economy. The Lagrangian for the household is given by:

$$\begin{aligned} L = \textstyle\sum_{t=0}^{\infty}\{\beta^t \frac{\left\{c_t^{\psi_1}\left[n^u\left(1-l_t^u\right)\right]^{\psi_2}\left[n^s\left(1-l_t^s\right)\right]^{\psi_3}\right\}^{(1-\sigma)}}{1-\sigma} - \beta^t \lambda_t^k [c_t + k_{t+1} - \left(1 - \delta^k\right) k_t - \\ - n^u w_t^u l_t^u - n^s w_t^s l_t^s - r_t k_t - \pi_t + T_t]\}, \end{aligned} \tag{2.5}$$

where $\lambda_t^k > 0$, is the Lagrange multiplier. The household chooses $\left\{c_t, l_t^s, l_t^u, k_{t+1}\right\}_{t=0}^{\infty}$

taking the initial condition, k_0, the policy variable, $\{T_t\}_{t=0}^{\infty}$, prices, $\left\{w_t^u, w_t^s, r_t\right\}_{t=0}^{\infty}$ and profits $\{\pi_t\}_{t=0}^{\infty}$ as given. The first-order condition (FOC) with respect to consumption:

$$\lambda_t^k = \frac{\psi_1 \left[\left(n^u \left[1 - l_t^u\right]\right)^{\psi_2} \left(n^s \left[1 - l_t^s\right]\right)^{\psi_3}\right]^{1-\sigma}}{c_t^{1-\psi_1(1-\sigma)}}, \tag{2.6}$$

states that the shadow price of the budget constraint (2.4) is equal to the marginal utility of consumption, $\frac{\partial U}{\partial c_t}$, at time t.

The *intratemporal* FOCs with respect to unskilled and skilled labour supply:

$$\frac{\psi_2}{\psi_1} \frac{c_t}{n^u \left(1 - l_t^u\right)} = w_t^u, \tag{2.7}$$

and

$$\frac{\psi_3}{\psi_1} \frac{c_t}{n^s \left(1 - l_t^s\right)} = w_t^s, \tag{2.8}$$

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

Finally, the Euler equation for capital:

$$\frac{1}{\beta} \left[\frac{c_t^{\psi_1} \left(n^u \left[1 - l_t^u\right]\right)^{\psi_2} \left(n^s \left[1 - l_t^s\right]\right)^{\psi_3}}{c_{t+1}^{\psi_1} \left(n^u \left[1 - l_{t+1}^u\right]\right)^{\psi_2} \left(n^s \left[1 - l_{t+1}^s\right]\right)^{\psi_3}}\right]^{1-\sigma} \frac{c_{t+1}}{c_t} = 1 + r_{t+1} - \delta_k \tag{2.9}$$

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.[21]

2.3.2 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:

$$\widetilde{y}_t^f = A\left\{\mu \left(q_t^u\right)^{\alpha} + (1-\mu)\left[\rho\left(k_t^f\right)^{\nu} + (1-\rho)\left(q_t^s\right)^{\nu}\right]^{\frac{\alpha}{\nu}}\right\}^{\frac{1}{\alpha}}, \tag{2.10}$$

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; q_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. $1/(1-\alpha)$ is the elasticity of substitution between capital and effective unskilled labour as well as between effective skilled and effective unskilled labour; whereas $1/(1-\nu)$ is the elasticity of substitution between capital and effective skilled labour. Capital-skill complementarity in production, which is

[21] Expectations are not taken into account here since the absence of uncertainty does not play a relevant role in the policy exercise I conduct.

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 *et al.* (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 $l_t^{f,i}\left(1 - t_t^i\right)$, whereas $l_t^{f,i} t_t^i$ is the actual time devoted to job-related training. Training increases next period's labour productivity. Building on the human capital tradition since Ben Porath (1967), and following e.g. Huggett *et al.* (2006), I assume that labour productivity, or else the stock of skills accumulated via job-related training evolves according to the following laws of motion:

$$h_{t+1}^u = (1 - \delta^u)\, h_t^u + H^u \left(l_t^{f,u} t_t^u h_t^u\right)^{\gamma^u}, \tag{2.11}$$

$$h_{t+1}^s = (1 - \delta^s)\, h_t^s + H^s \left(l_t^{f,s} t_t^s h_t^s\right)^{\gamma^s}, \tag{2.12}$$

where $0 < \delta^u, \delta^s < 1$ are the depreciation rates for the skills accumulated by unskilled and skilled workers respectively; the stock of skills for each group is measured by the variable h_t^s and h_t^u, respectively; $H^i\left(l_t^{f,i}t_t^i h_t^i\right)^{\gamma^i}$ $(i = s, u)$ represents the 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. On-the-job training skills are firm-specific as it is assumed that workers cannot change employer to free-ride on the training they received. Also, note that both H^i and γ^i are related to workers' learning ability, i.e. the ability of the workers to use existing skills and their training time to create new labour skills (see e.g. Huggett *et al.* (2006)). 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, I 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), I 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, I 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 q_t^i is a function of workers' time and of labour productivity:

$$q_t^s = \left[l_t^{f,s}\left(1 - t_t^s\right)\right]^{\omega}\left[h_t^s\right]^{1-\omega}, \tag{2.13}$$

$$q_t^u = \left[l_t^{f,u}\left(1 - t_t^u\right)\right]^{\omega}\left[h_t^u\right]^{1-\omega}, \tag{2.14}$$

where $0 < \omega < 1$ measures the elasticity of effective labour with respect to production time. Note that the constant returns to scale restriction in (2.13)-(2.14) implies that the production function (2.10) is also constant returns to scale in its five inputs $\left\{l_t^{f,i}\left(1 - t_t^i\right), h_t^i\right\}_{i=s,u}$ and k_t^f.

This setup implies that it is the firm, and not the worker, which assumes the costs of training and owns job-related skills associated with h_t^i, thus capturing firm-specific skills that are augmented by job-related training.[22] As explained in Section 2, this is consistent with empirical evidence which suggests that (i) firms

[22]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).

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 rents for the firms associated with job-related training.

Indeed, in this specification, and given that the production function in (2.10) 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 maximise the present discounted value of lifetime profits (e.g. Chen and Lai (2015)):[23]

$$\Pi^f = \sum_{t=0}^{\infty} Q_t \pi_t, \tag{2.15}$$

where

$$Q_t = \prod_{j=0}^{t-1} \left(1 + r_{j+1} - \delta^k\right)^{-1}, \tag{2.16}$$

[23]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.

defines the discount factor[24] and

$$\pi_t^f = y_t^f - r_t k_t^f - w_t^u l_t^{f,u} - w_t^s l_t^{f,s} - \phi^u t_t^u l_t^{f,u} (1 - \tau^u) - \phi^s t_t^s l_t^{f,s} (1 - \tau^s), \qquad (2.17)$$

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 monetary 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 (2.10)-(2.14) and (2.17). 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, and 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 I observe in the UK data (see Section 2.2). In particular, since the employees have different marginal products of effective labour, there is

[24]It holds that: $Q_0 = \prod_{j=0}^{-1} \left(1 + r_{j+1} - \delta^k\right) = 1.$

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 size of training investments, of skilled and unskilled workers, depends on the quantitative evaluation of these trade-offs.

Taking the initial conditions, $\left\{k_0^f, h_0^s, h_0^u\right\}$, prices, $\left\{w_t^s, w_t^u, r_t\right\}_{t=0}^{\infty}$, policy rates $\{\tau^s, \tau^u\}_{t=0}^{\infty}$ 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 (2.15), subject to (2.11) and (2.12).[25] The Lagrangian for the firm is given by:

$$\begin{aligned}
\Lambda = \textstyle\sum_{t=0}^{\infty}\{Q_t\{y_t^f - r_t k_t^f - w_t^u l_t^{f,u} - w_t^s l_t^{f,s} - \phi^u t_t^u l_t^{f,u}(1-\tau^u) - \phi^s t_t^s l_t^{f,s}(1-\tau^s)\} - \\
-Q_t \lambda_t^u [h_{t+1}^u - (1-\delta^u) h_t^u - H^u \left(l_t^{f,u} t_t^u h_t^u\right)^{\gamma^u}] - \\
-Q_t \lambda_t^s [h_{t+1}^s - (1-\delta^s) h_t^s - H^s \left(l_t^{f,s} t_t^s h_t^s\right)^{\gamma^s}]\},
\end{aligned} \tag{2.18}$$

[25] 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 (2.17).

where λ_t^i are the shadow prices associated the skill accumulation constraints (2.11) and (2.12); and where y_t^f, substituting out q_t^s and q_t^u , is defined as:

$$
\begin{aligned}
y_t^f &= A\left\{\mu\left(\left[l_t^{f,u}\left(1-t_t^u\right)\right]^{\omega}\left[h_t^u\right]^{1-\omega}\right)^{\alpha}+\right.\\
&\left.+(1-\mu)\left[\rho\left(k_t^f\right)^{\nu}+(1-\rho)\left(\left[l_t^{f,s}\left(1-t_t^s\right)\right]^{\omega}\left[h_t^s\right]^{1-\omega}\right)^{\nu}\right]^{\frac{\alpha}{\nu}}\right\}^{\frac{1}{\alpha}}.
\end{aligned}
\tag{2.19}
$$

The *intratemporal* FOCs with respect to capital, unskilled and skilled labour:[26]

$$
r_t = \frac{\partial y_t^f}{\partial k_t^f},
\tag{2.20}
$$

$$
w_t^u + \phi^u t_t^u \left(1-\tau^u\right) = \frac{\partial y_t^f}{\partial l_t^{f,u}} + \lambda_t^u \frac{\partial h_{t+1}^u}{\partial l_t^{f,u}},
\tag{2.21}
$$

and

$$
w_t^s + \phi^s t_t^s \left(1-\tau^s\right) = \frac{\partial y_t^f}{\partial l_t^{f,s}} + \lambda_t^s \frac{\partial h_{t+1}^s}{\partial l_t^{f,s}},
\tag{2.22}
$$

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

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

in skill accumulation, $\frac{\partial h^i_{t+1}}{\partial l^{f,i}_t}$, valued by its corresponding shadow price, λ^i_t. 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 higher future labour productivity.

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

$$\phi^u l^{f,u}_t (1-\tau^u) - \frac{\partial y^f_t}{\partial t^u_t} = \lambda^u_t \frac{\partial h^u_{t+1}}{\partial t^u_t}, \tag{2.23}$$

and

$$\phi^s l^{f,s}_t (1-\tau^s) - \frac{\partial y^f_t}{\partial t^s_t} = \lambda^s_t \frac{\partial h^s_{t+1}}{\partial t^s_t}, \tag{2.24}$$

equate their respective marginal costs to their marginal products. Marginal costs are equal to the opportunity cost of foregone output, $\frac{\partial y^f_t}{\partial t^i_t}$, 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, $\frac{\partial h^i_{t+1}}{\partial t^i_t}$, valued by its corresponding shadow price, λ^i_t.

Finally the Euler equations for unskilled and skilled skills acquisition:

$$\lambda^u_t = \frac{Q_{t+1}}{Q_t}\left(\frac{\partial y^f_{t+1}}{\partial h^u_{t+1}} + \lambda^u_{t+1}\frac{\partial h^u_{t+2}}{\partial h^u_{t+1}}\right), \tag{2.25}$$

$$\lambda_t^s = \frac{Q_{t+1}}{Q_t}\left\{\frac{\partial y_{t+1}^f}{\partial h_{t+1}^s} + \lambda_{t+1}^s \frac{\partial h_{t+2}^s}{\partial h_{t+1}^s}\right\}, \tag{2.26}$$

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.2

2.3.3 Government budget

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

$$T_t = \tau^u \left(\phi^u t_t^u l_t^{f,u}\right) + \tau^s \left(\phi^s t_t^s l_t^{f,s}\right), \tag{2.27}$$

which equates the lump-sum transfer/tax, T_t, with the expenditure to subsidise 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. These assumptions will be relaxed in the next Chapter.[27]

[27] As already discussed, the current aim is to evaluate training subsidies and understand their transmission mechanism without confounding factors, e.g. distortionary taxation effects.

2.3.4 Market clearing conditions

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

$$k_t^f = k_t, \tag{2.28}$$

$$l_t^{f,u} = n^u l_t^u, \tag{2.29}$$

$$l_t^{f,s} = n^s l_t^s, \tag{2.30}$$

$$\pi_t^f = \pi_t, \tag{2.31}$$

and

$$y_t^f = c_t + k_{t+1} - (1-\delta_k) k_t + \phi_u t_t^u l_t^{f,u} + \phi_s t_t^s l_t^{f,s}. \tag{2.32}$$

2.3.5 Initial and transversality conditions

To ensure the existence of a solution, I assume that:

$$k_0 = \bar{k} > 0, \tag{2.33}$$

$$h_0^s = \bar{h}^s > 0, \tag{2.34}$$

and

$$h_0^u = \bar{h}^u > 0, \tag{2.35}$$

where $\bar{k}$, $\bar{h}^s$, and $\bar{h}^u$ could take any strictly positive number, as production cannot take place if any of these inputs is null. Further, since agents are infinitely-living, to ensure a finite and unique solution to the model exists, I impose the following transversality conditions:

$$\lim_{t\to\infty} E_t \beta^t \lambda_t^h k_{t+1} = 0, \tag{2.36}$$

$$\lim_{t\to\infty} E_t \beta^t \lambda_t^h h_{t+1}^s = 0, \tag{2.37}$$

and

$$\lim_{t\to\infty} E_t \beta^t \lambda_t^h h_{t+1}^u = 0 \tag{2.38}$$

which in this context are bound to hold, since e.g. my assumptions on skill-stock accumulation parameters entail that (2.37) and (2.36) cannot be violated.

2.3.6 Decentralised Equilibrium

Given initial conditions and policy rates $\{\tau_s, \tau^u\}$, the decentralised 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 $\left\{r_t, w_t^u, w_t^s\right\}_{t=0}^{\infty}$, 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' law I discard the household's budget constraint (redundant), thus the DE consists of the following 19 equations: (i) the household's 4-FOCs, equations (2.6)-(2.9); (ii) the firm's 2-skill accumulation equations (2.11)-(2.12); (iii) the firm's 7-FOCs, equations (2.20)-(2.26); (iv) the government's budget constraint, equation (2.27); and (v) the 5-market clearing conditions, equations (2.28)-(2.32).

2.3.7 Model calibration and steady-state

The parameters appearing in the DE equations are set with the overall aim that the model generates a steady-state solution implying model generated quantities similar to the actual data for the UK. The calibrated parameters are summarised in Table 2.3. 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, I set the coefficient of relative risk aversion $\sigma = 2$.[28] Similarly, I 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).

[28] For example, Browning *et al.* (1999), Ionescu (2009) and Bakış *et al.* (2015).

Given that the depreciation of job-related skills is hard to measure, I assume $\delta^s = \delta^u = \delta^k$. The literature on work-related human capital, e.g. Blundell *et al.* (1999), suggests that skills depreciate within a decade or so, which implies a yearly depreciation rate of about 10%. Indeed, Mincer and Ofek (1982) estimate 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, my assumption of 2.5% is in-between these estimates.

Next, I set the quarterly discount factor of $\beta = 0.995$ to ensure that the annualised risk-free interest rate net of depreciation is equal to 2.9 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 1995q1-2008q4. 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 about training subsidies to firms is available from CVTS 3 & 4. I divide these subsidies by the training costs, the for the firms whose data is available, and find that the subsidies amount, on average, to about 2.4% 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. I thus set $\tau^u = \tau^s = 2.4\%$, assuming that training is equally subsidised.

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. I normalise 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 $l^u = 0.29$ at equilibrium.

I next move to the group of parameters relating to training and production $\{\nu, \alpha, \mu, \rho, \phi^u, \phi^s, \gamma^u, \gamma^s, \omega\}$. I start with the parameter ω, which, as previously discussed, is directly linked to firms' profitability and to the returns associated with job-related training, i.e. firm-specific rents. To the best of my knowledge, data for the UK on firms' returns, in terms of profitability, associated with firms' expenses

Table 2.3: 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	0.024	public subsidy for training activities (unskilled)
τ^s	0.024	public subsidy for training activities (skilled)

on job-related training do not exist. 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. I 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, I choose ω so that, in conjunction with the remaining parameters, firms' returns on investment in training, defined as above, are about 10%.

I 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) my own calculations from the UK QLFS data;[29] (iv) ONS data on gross value added (GVA) and the estimates of the total training costs reported in the 2011 ESS; (v) my own calculations, based on ESS estimates of total training time per employee, the population shares, and the average training participation rate of non-University

[29]The skill premium is obtained by averaging the ratio of the hourly wage of University educated workers and that of non-University educated workers over the period 1995q1-215q4

educated workers.[30]; (vi) same as (v), but with respect to University educated workers. Together, 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)), I set the elasticities of substitution $\nu = -0.495$ and $\alpha = 0.401$, based on the estimates by Krusell *et al.* (2000). I 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, I choose $\{\mu, \rho, \phi^u, \phi^s, \gamma^u, \gamma^s\}$ so that the model's steady-state predictions regarding the targets $\left\{\frac{n^s l^s w^s + n^u l^u w^u}{y}, \frac{k}{y}, \frac{w^s}{w^u}, \frac{\phi_s t^s l^s n^s + \phi_u t^u l^u n^u}{y}, t^u, t^s\right\}$ are similar to the data.

Table 2.4: 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	0.026
T/y	public spending on training costs-to-output	0.0006	0.0006
rk/y	capital income-to-output	0.306	0.299
$\frac{n^s l^s w^s + n^u l^u w^u}{y}$	labour income-to-output	0.665	0.671

[30] The population shares and the training participation rates are derived from the QLFS with the waves from 1995q1 to 2015q4.

The steady-state solution implied by the parametrisation in Table 2.3 is summarised in Table 2.4. As can be seen, the model's predictions for the long-run quantities are close to the data. Moreover, I can simulate the model to evaluate its predictions 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 2.2 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 0.14% to 0.39%.[31] Given that I cannot differentiate between skilled and unskilled workers in the data, this estimate refers to an average effect, across worker types. Thus, I examine the response of the model to increasing both τ^u and τ^s by 1%, starting from the state in Table 2.4.

The model simulation reveals that, on average, across skilled and unskilled workers, training increases by 0.03%. Thus, the elasticity of training time with respect to training subsidies implied by the model is on the conservative side with respect to the empirical evidence reported in Section 2.2. Before focusing on the fiscal policy, the next section reports the effect of a positive innovation to the total factor productivity.

[31]Note that under the assumption of homoskedasticity, the interval is larger and it extends from 0.05% to 0.47%

2.4 Total factor productivity innovations

Chapter 1 emphasises that, in the UK, the training participation rate is acyclical, at least when the business cycle is represented by GDP fluctuations from its HP-filtered trend. A large strand of the literature, started with Lucas (1977) and revamped by King and Rebelo (1999), has tried to identify stylised facts with respect to business cycle properties of macroeconomic aggregates, e.g. GDP and employment. A main feature of the standard RBC model is that it focuses on total factor productivity as the driver of economic fluctuations.[32]

The present work does not pursue matching second, or higher, moments as it focuses on policy analysis and the time-frame adopted is the medium or long run.[33] For this reason, this section aims to show that the model's behaviour does not differ from the typical behaviour of RBC models.

To do this, I assume that total factor productivity, A_t, is time-varying. This variable follows an the first order autoregressive AR(1) process:

---

[32] This rigid assumption prevents the model from replicating some of the stylised facts observed in the data (see e.g. Hansen (1985) and more recently DeJong and Ingram (2001)). Nonetheless, the framework is still used extensively for research purposes.

[33] The literature has already used a model with training to match data moments. One example is Einarsson and Marquis (1998). As their attempt partly fails, they argue that a model with training investments can improve its fit to training data (at business cycle frequencies) only introducing additional shocks to the human capital law of accumulation.

$$log(A_t) = \rho_a log(A_{t-1}) + \epsilon_{a,t}, \tag{2.39}$$

where $\epsilon_{a,t}$ is an *i.i.d.* innovation that has zero mean and constant variance, and the autocorrelation coefficient, ρ_a, is set to 0.89 following the estimation performed by Harrison and Oomen (2010).[34] As usual, I simulate a 1% shock to total factor productivity and report in Figure 2.2 the impulse response functions (IRFs).

Figure 2.2 shows that the increase in productivity leads to larger investments into physical capital (sub-plot (1,3)). To meet the excess final good demand, producers increase supply by using more labour input. Firms offer higher wages so that workers will supply the desired quantity of labour. Moreover, firms reduce the share of time dedicated to training to boost output. Higher wages entails that consumption rises, increasing further the demand of output.

Notice that the increase in labour is larger than the decrease in training time share, therefore the net effect on skill accumulation is still positive as confirmed by the dynamics of human capital stock (sub-plots (1,4) and (2,4)). After all, a higher human capital stock is beneficial to firms since they can generate larger streams of profits even after the productivity boost dissipates.

Wage inequality falls initially as the unskilled labour supply is less elastic than

[34] This will differ from the autocorrelation coefficient used to compare my model to the literature in Figure 2.3.

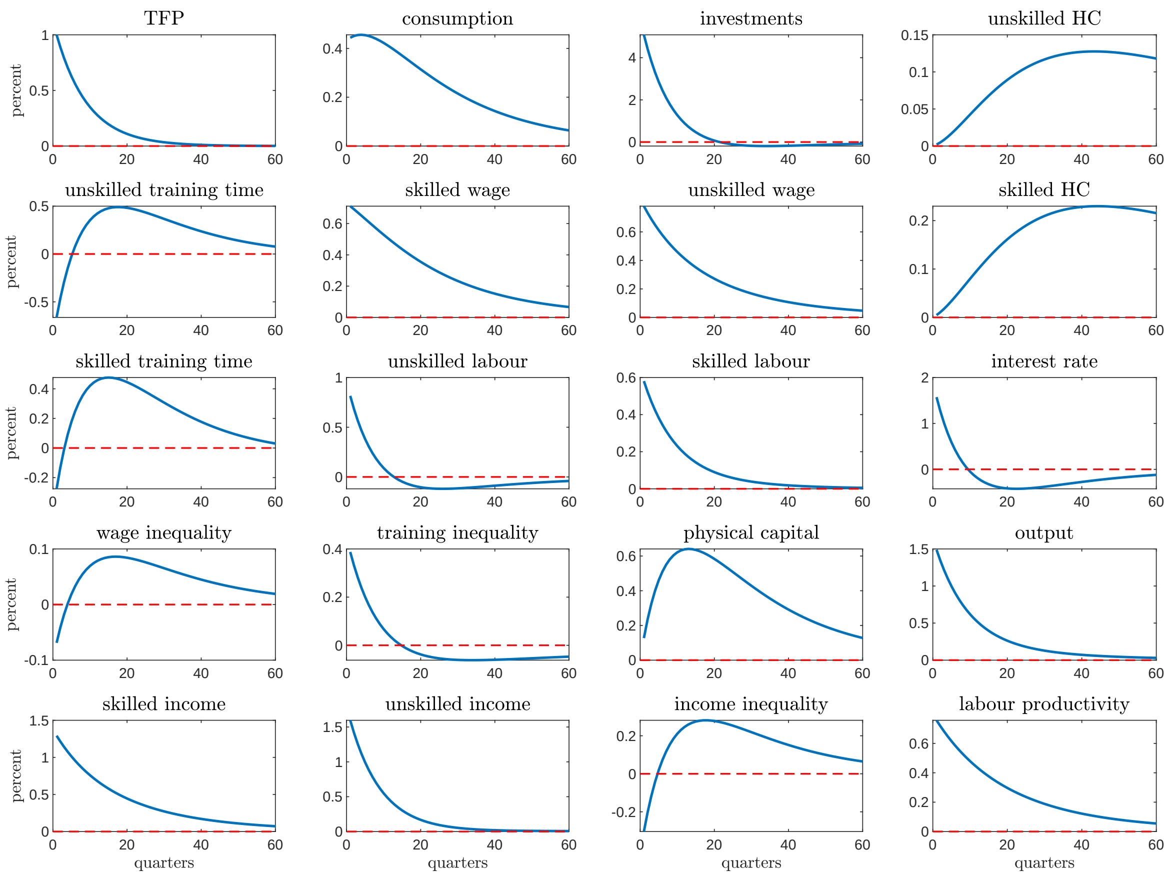


Figure 2.2: Temporary increase in total factor productivity by 1%

that of skilled workers. Thus, firms offer relatively higher wages to unskilled workers. Unfortunately for this group, the complementarity between physical capital and skilled labour entails that firms have greater incentives to accumulate skilled human capital, and in turn this leads to skilled workers facing better salaries in relative terms than unskilled ones. As a consequence, wage inequality rises in the medium run (sub-plot (4,1)).

Figure 2.3 directly compares the impulse response functions of my model and

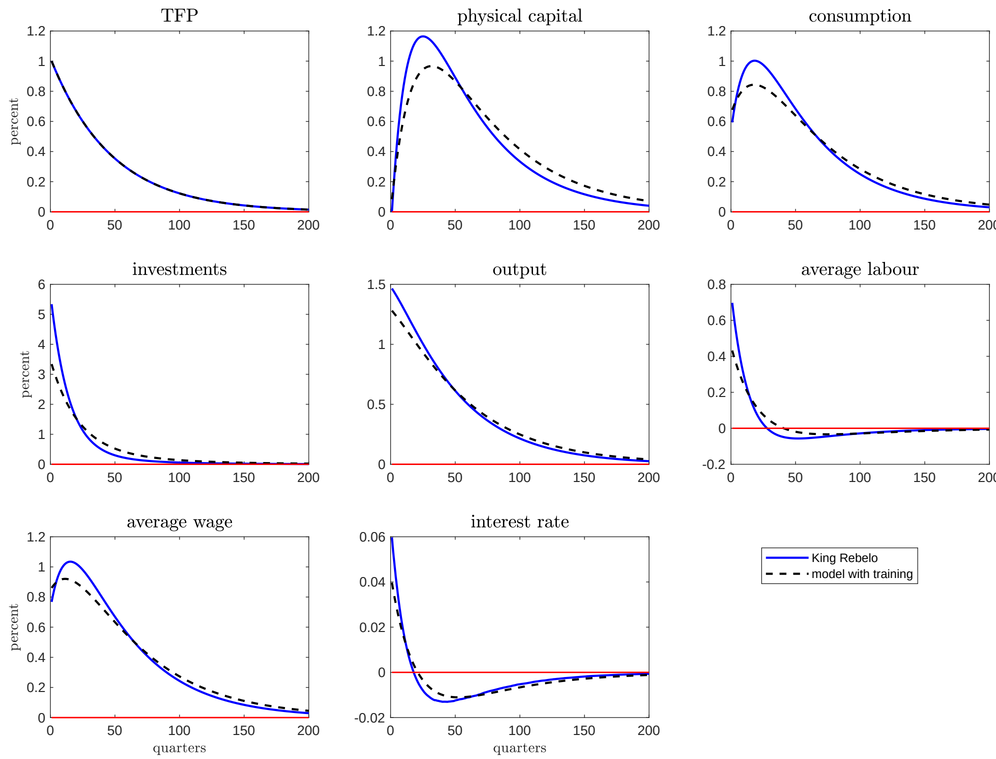


Figure 2.3: Comparison between current and the classical RBC model

that of King and Rebelo's model after a 1% temporary shock to the TFP.[35] Even though the model with training looks very different from the model in King and Rebelo (1999), the response of main aggregate variables is very alike.

Wage, the interest rate, and output all react very similarly after the TFP shock. It seems that investments react less in the training model at impact and the accumulation of physical capital is lower overall. Yet, investments decline faster

[35] The data for the plots are my own, obtained replicating the original work of King and Rebelo (1999). The code is available on request. Since they used a higher autocorrelation coefficient, i.e. $\rho_a = 0.979$, the IRFs in figure 2.3 vary marginally from those reported in Figure 2.2 as in the latter I follow King and Rebelo's calibration for comparison purposes.

in King and Rebelo's model. Thus, after about 70 periods, the physical capital percent deviation is higher in the training model. As a consequence, the same shift can be observed in the paths of consumption and aggregate output.

With respect to the average labour supply, the dynamics is similar between the two models: a first period where supply is above steady state, followed by a long period of low labour supply. Even the timing of this transition is similar. It is only slightly delayed by the presence of training and skills capital in the model proposed in this chapter.

After showing that, despite the innovations, the model does not differ substantially from the classical real business cycle model as presented by King and Rebelo (1999), I perform policy exercises to evaluate the desirability of training subsidies using the proposed framework.

2.5 Policy analysis

This section examines the dynamic effects of training subsidies on training, wages and earnings. To solve for the transition paths, I assume that the economy is at its steady-state, when a one-off, permanent, change takes place in either τ^u or τ^s. I 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

DGE system of equations for S periods. The rational-expectation path is solved non-linearly using standard numeric methods with *Dynare* (see Adjemian *et al.* (2011)). I set $S = 1000$ to ensure that convergence is achieved.[36]

2.5.1 Some preliminary results

The effect of training subsidies on training time can be derived analytically. The conclusion I draw is expected to hold for most (reasonable) values of training subsidies, but will not hold for all values because very large subsidies can have undesirable general equilibrium repercussions. To see the effect of increases in τ^i on training time for skill group i, recall that the first-order condition with respect to training time can be re-written as:

$$\phi^i l_t^{f,i}\left(1-\tau^i\right) = \lambda_t^i \frac{\partial h_{t+1}^i}{\partial t_t^i} + \frac{\partial y_t^f}{\partial t_t^i}, \tag{2.40}$$

and implies that a reduction in the training costs requires an increase of training time. In fact, since the left-hand side has fallen, the right-hand side of (2.40) must also fall. The lever for the firm is to increase the training time t_t^i, since both $\frac{\partial h_{t+1}^i}{\partial t_t^i}$

[36]For most levels of subsidies, convergence is achieved in a modest number of periods.

and $\frac{\partial y_t^f}{\partial t_t^i}$ are a decreasing function of t_t^i:

$$\frac{\partial\left\{\lambda_t^i \frac{\partial h_{t+1}^i}{\partial t_t^i} + \frac{\partial y_t^f}{\partial t_t^i}\right\}}{\partial t_t^i} < 0. \tag{2.41}$$

The first result is a direct consequence of the calibration of each γ_i, which entails that the skill creation function is concave in training time, labour, and human capital stock. The second result can be proven by using the production function (2.10) where $\frac{\partial y_t^f}{\partial q_t^i} > 0$, and Equation (2.14) where $\frac{\partial q_t^i}{\partial t_t^i} < 0$ and $\frac{\partial^2 q_t^i}{\partial^2 t_t^i} < 0$. It follows that $\frac{\partial y_t^f}{\partial t_t^i} = \frac{\partial y_t^f}{\partial q_t^i}\frac{\partial q_t^i}{\partial t_t^i} < 0$ and $\frac{\partial^2 y_t^f}{\partial^2 t_t^i} = \frac{\partial\left[\frac{\partial y_t^f}{\partial q_t^i}\frac{\partial q_t^i}{\partial t_t^i}\right]}{\partial t_t^i} = \frac{\partial y_t^f}{\partial q_t^i}\frac{\partial^2 q_t^i}{\partial^2 t_t^i} < 0$.

Therefore, a rise in subsidies τ^i tends to create a rise in t_t^i via the substitution effect – cheaper training leads to more training efforts. However, since $\frac{\partial y_t^f}{\partial t_t^i}$ is also a decreasing function of t_t^i, this means that the increase in indirect costs of training mitigates the impact of the policy. In other words, a higher share of training time increases the opportunity costs of further marginal increases in training time and dampens the reaction of t_t^i to changes in τ_i.

All this holds as a partial equilibrium analysis, and it is evident that indirect channels may affect the final allocation of training time. Yet, as simulations have confirmed, the general equilibrium effects are of second order importance and do not spoil this result, unless the training subsidy is so large that distortions alter drastically the allocation of all factors and resources leading to contradictory results (e.g. training subsidy crowding out training investment). Given the model features, and the current calibration, crowding out is not meant to occur.

With respect to changes in labour for skill group i, there is no general result that I can present. The generic first order condition with respect to labour input for skill group i can be re-written as labour demand:

$$w_t^i = \frac{\partial y_t^f}{\partial l_t^{f,i}} + \lambda_t^i \frac{\partial h_{t+1}^i}{\partial l_t^{f,i}} - \phi^i t_t^i \left(1 - \tau^i\right), \tag{2.42}$$

and suggests that as training time increases, $\frac{\partial y_t^f}{\partial l_t^{f,i}}$ and $\lambda_t^i \frac{\partial h_{t+1}^i}{\partial l_t^{f,i}}$, while $\lambda_t^i \frac{\partial h_{t+1}^i}{\partial l_t^{f,i}}$ is likely to decrease (since the elasticity of training time to training subsidies shouldn't be larger than one). Thus, the labour demand is shifting to the right. Yet, as I show below, the supply of labour also shifts to the right. Thus, while equilibrium labour will increase, it is impossible to predict whether wage will rise or decrease.

2.5.2 Income and inequality effects

The effects of a permanent increase in τ^u from 0.024 to 0.5 are shown in Figure 2.4. The policy implies that the government subsidises half of the training cost for unskilled workers. A permanent rise in training subsidies, τ^u, increases training time for unskilled workers (see sub-plot (1,2) for t_t^u) and this influences the allocation of all other resources.

The effect of τ^u on training time has been discussed in Subsection 2.5.1. As implied by the first-order condition (2.23), the (predominating) substitution effect entails that the representative firm increases training time since this is now cheaper, *ceteris paribus*. The increase in t_t^u will lead to the accumulation of higher worker skills (see sub-plot (2,1) for h_t^u).

For $\tau_u = 0.5$, the incentives offered by the policy are large enough to entail an increase of the unskilled labour demand which puts upward pressure on unskilled wage. The additional labour input more than compensates the reduction of time due to increased training efforts, as can be observed from sub-plot (2,3) for $l_t^u\left(1 - t_t^u\right)$. With time, the accumulated human capital stock increases labour productivity, and this works to further increase the aggregate labour demand.

On the other hand, the increase in the unskilled labour input tends to decrease the marginal product of unskilled labour (see sub-plot (5,1)), due to decreasing marginal returns. This counterbalances the increase of value of investing labour input to train unskilled workers. At the same time, the households predict higher future output and expand the unskilled labour supply to increase current income (thus consumption). At impact, the general equilibrium effect of these forces is resolved into an increase in $l_t^{f,u}$ (see sub-plot(4,3)) and a very modest – and

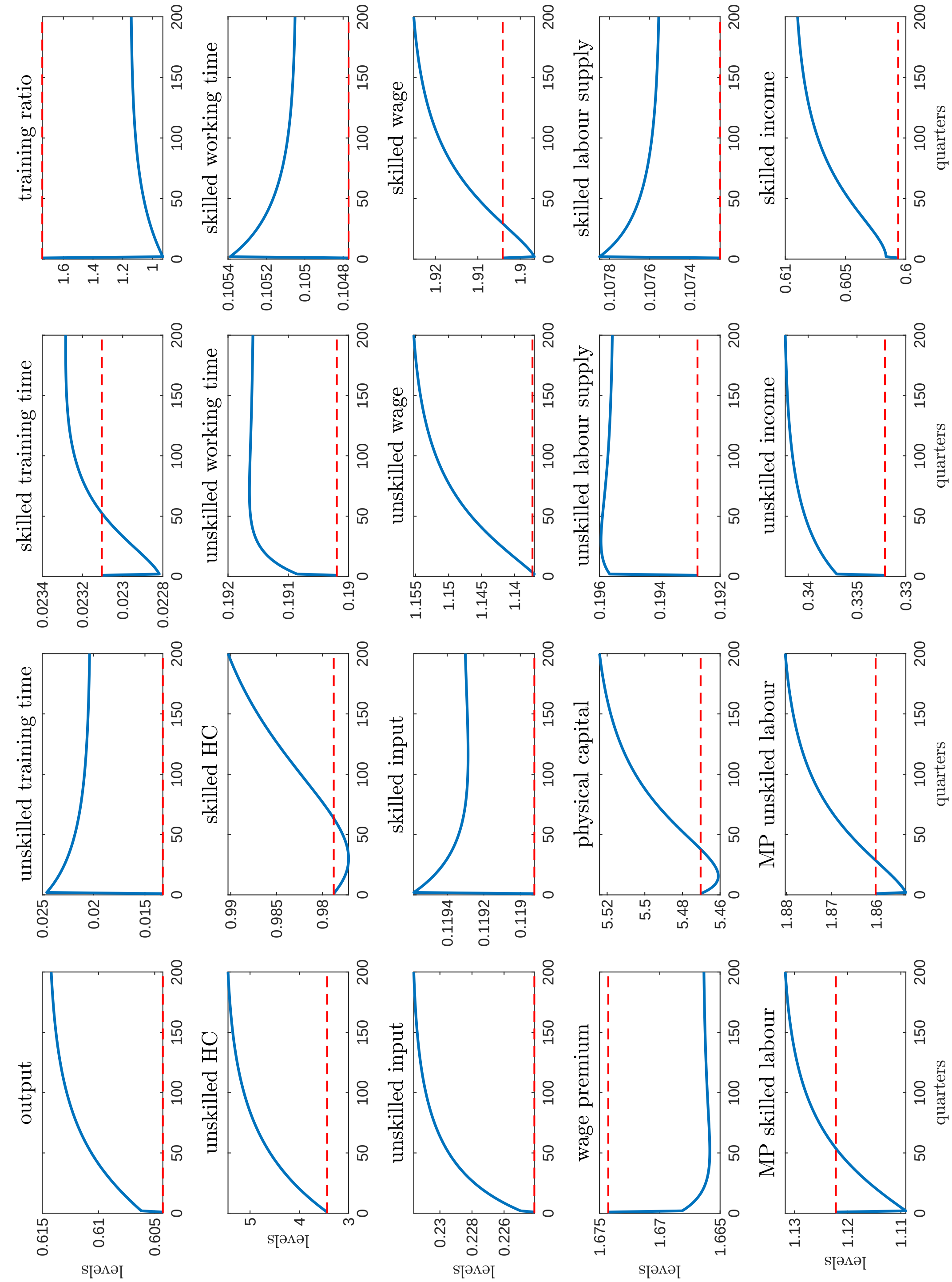


Figure 2.4: Permanent increase of unskilled training subsidy to 0.5

temporary – fall of unskilled wage (see sub-plot (3,3)).[37]

For skilled workers, the labour demand is not supported by the increase in the value of training whereas the labour supply expands as for unskilled workers.[38] As a consequence, the equilibrium skilled wage falls more intensely (see sub-plot (3,4)) than unskilled wage at impact. Later on, as the additional training efforts pay their dividends in terms of increased skill capital, the labour productivity picks up and firms are willing to offer increasingly higher skilled (and unskilled) wages.

The positive developments in the labour market for unskilled labour and the increase in the unskilled effective labour input q_t^u (see sub-plot (3,1)), have positive spillover effects on the productivity and returns to skilled labour. In particular, after the initial decline, the marginal product of skilled labour and thus skilled wages increase (see sub-plots (5,2) and (3,4) respectively). Following these dynamics, capital stock is also increasing (see sub-plot (4,2)). Hence, the increased labour productivity and employment for unskilled workers initially crowds out capital, skilled training and skilled labour productivity (see sub-plots (1,3) and (2,2) for t_t^s and h_t^s, respectively). However, as more resources are diverted to-

[37]As discussed earlier, the impact reaction of unskilled wage would have been different if the subsidy increase were smaller.

[38]Remember that, for simplicity, I assume a single household reunites unskilled and skilled members.

wards 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 sub-plot (4,4) for $l_t^{f,s}$). The changes in wages and hours imply that earnings are also increased (see sub-plot (5,4) for $w_t^s l_t^s$).

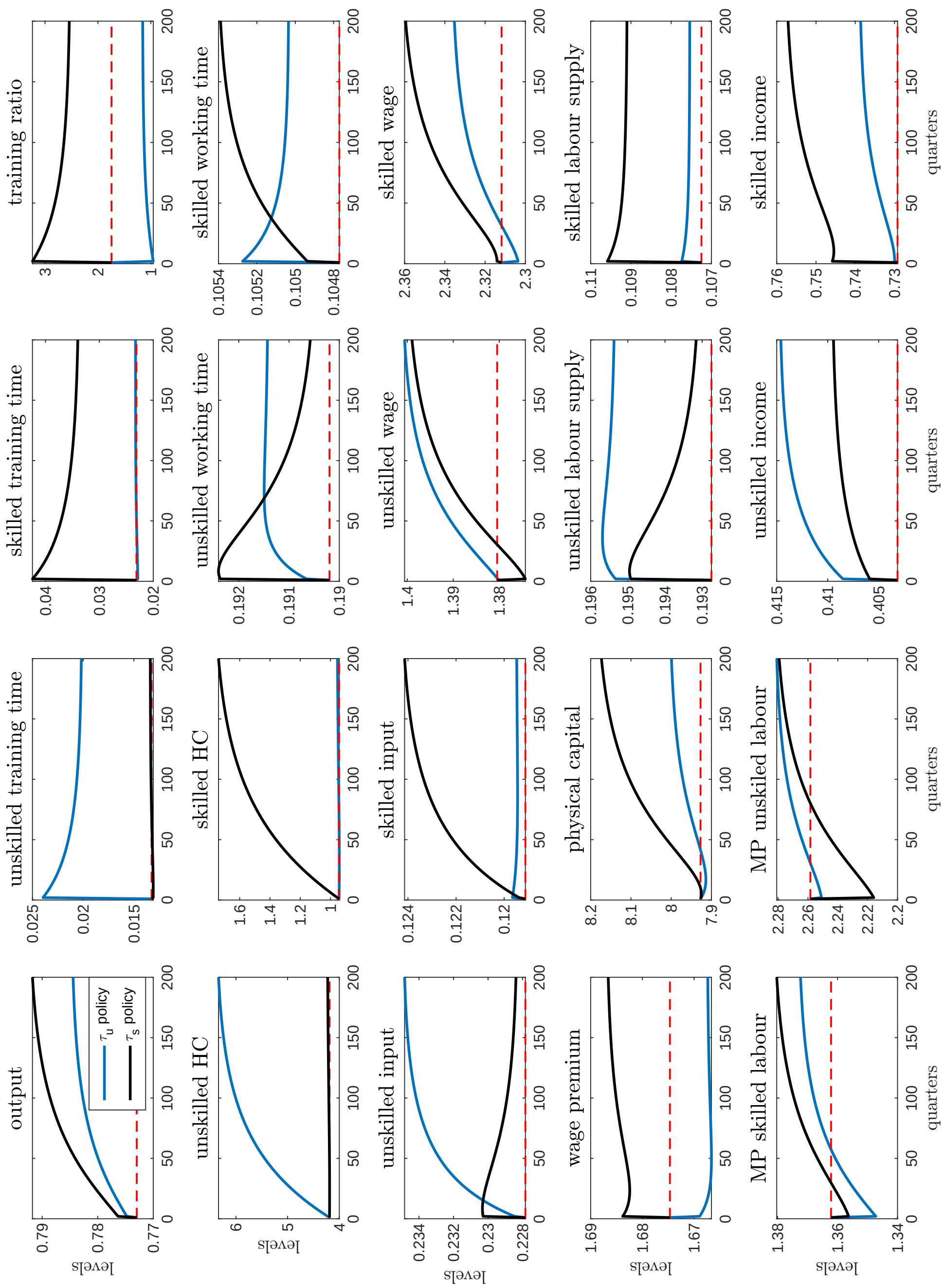


Figure 2.5: Permanent increase of skilled training subsidy to 0.5

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.[39]

Very similar 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 2.5 shows the effects of a permanent increase in τ^s from 0.024 to 0.5, which implies that the government subsidises half of the training costs for skilled workers. This can be compared to previously discussed τ_u policy as this is also reported in the same figure. In this case, although increasing τ^s improves both households' labour income and overall output increases more, it increases inequality.

Table 2.5 summarises 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, I consider three different magnitudes of training subsidies, in particular $\tau^i = 0.25$, $\tau^i = 0.5$ (which was analysed in Figures 2.4 and 2.5), and $\tau^i = 1$ for $i = u, s$. Therefore, each column contains information

[39]Conditional on the assumptions about the financing of the fiscal subsidies.

about a subsidy of different size, and the rows take into consideration the effect on different variables, e.g. unskilled training in the first row or skilled income in the eighth row. The effect is measured in terms of percentage change of each variable from the original steady state to the new one where the fiscal policy is permanently implemented.

The increase in τ^u to 0.25, increases training for unskilled workers by nearly 18% (see cell (1,1) of Table 2.5), implying an increase from 3.4 days of average training to 4 days. Similarly, the increase in τ^u to 0.5 implies an increase from 3.4 days of average training to 5 days.[40] In terms of implied elasticities, these effects suggest that an increase by 1% in τ^u increases training for the unskilled workers by 0.02%, which is at the lower bound of the estimates in Table 2.3.[41] 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, yet quantitatively significant, increases in wages for the unskilled. The effect of the increase in training on wages is also consistent

[40]Using ESS and QLFS data, I 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.

[41]The model-simulated elasticity I mention is obtained as the ratio between the percentage change in unskilled training over the percentage change in unskilled training subsidy, i.e. the value of cell(1,1) divided by $100(0.25/0.024 - 1)\%$.

with previous econometric evidence (see e.g. Table 2 in Blundell *et al.* (1999)). In particular, I find in Table 2.5 (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 training 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 wage of the average worker.

In terms of inequality, training subsidies for unskilled workers reduce slightly wage inequality, because of the concurrent increase in wages for the skilled. Earnings inequality is reduced by more, due to the positive effects of increased training on unskilled hours. The relationship between wage inequality and training inequality, as shown in Table 2.5, is also consistent with the empirical estimates reported in Table 2.2. The results in Table 2.5 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 2.2.[42]

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

[42] The model-simulated elasticity varies with the size of the intervention. For simplicity, I take the ($\tau_u = 0.25$) case as reference, i.e. column one of Table 2.5, and I compute the elasticity as the ratio between the percentage change in wage inequality over the percentage change in training inequality, i.e. sixth row divided by the third row. This equals to 0.011% and it can be compared to the empirical estimates (taking the exercise with a pinch of salt).

Table 2.5: 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

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, I observe important spillovers especially with respect to the unskilled wage.

The message from Table 2.5 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, I explore further the resource effectiveness of these income gains using fiscal multipliers.

2.5.3 Multiplier analysis

I 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, I compute fiscal multipliers based on the simulation exercise described above. In particular, I 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)), I compute the lifetime multiplier of e.g. the variable x according to the formula:

$$\text{lifetime multiplier} = \frac{\sum_{t=0}^{S} Q_t \left(x_t - x\right)}{\sum_{t=0}^{S} Q_t \left(T_t - T\right)} \tag{2.43}$$

where Q_t is the household discount factor introduced in (2.16). I simulate $S = 2000$ periods after the shock to compute (2.43). The multipliers for the training subsidy policies are reported in Table 2.6.

Focusing on the top half of the table (unskilled policies), it can be observed that

Table 2.6: Multiplier effects of increasing the unskilled and skilled training subsidies

		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
		impact multiplier				lifetime multiplier		
	$\tau^s =$	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

all multipliers (except on impact for capital income) are positive and most of the lifetime multipliers are larger than unity. Therefore, over the lifetime, the increase in benefits is greater than the increase in resources required to finance the policy. As expected, given the dynamics observed in Figure 2.4, 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. Also, the multipliers are decreasing as τ_u gets larger, which implies decreasing returns to policy interventions in the training sector.

In the second half of Table 2.6, I report the multiplier effects for permanent increases of the training subsidies to skilled training, τ^s. The results are comparable to those of the first half, 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.

2.6 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, I 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.

The quantitative policy analysis suggests 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.[43]

Training subsidies to skilled workers, while increasing skilled and unskilled earnings, raise the former by more, worsening 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.

For this analysis, I assumed that the market provision of training is efficient. Although this is a rather extreme assumption, it helps focusing on the redistributive aspect of the fiscal policy. Moreover, I can test the model fitness and its consistency with UK empirical evidence, and I can draw clearer conclusions about the channels and the primary effect of training subsidies.

Yet, to fully understand the consequences of fiscal policies in the training sector, it is important to consider the role of externalities. This allows to perform a policy

[43]This issue will be addressed in Chapter 3.

evaluation based on welfare measures, and to reach stronger conclusions about the effectiveness of this fiscal policy tool. For this reason, my third chapter will present an extended and more realistic version of the core model, where channels such as distortionary taxation and externalities are included.

"Developing skills is as important as training. A larger effort is needed to create a skilled workforce with employment potential."

Pallam Raju, politician

3

The welfare effects of training subsidies

3.1 Introduction

The empirical evidence of the previous chapter shows that training provision has a moderate impact on wage inequality. I provide a potential explanation of this evidence with a general equilibrium model calibrated on UK data. The proposed framework has the advantage of being simple enough to allow a clear understanding of the channels through which fiscal subsidies affect training provision and wages. Yet the model has several limits since it abstracts from many real-economy features. The chapter addresses this issue by providing new evidence and presenting an extended model for policy purposes.

The assumption that training provision is efficiently provided by market forces is particularly restrictive. As shown in Chapter 1, a large strand of the literature has posited that training is likely to be under-provided. Chapter 3 proposes a more realistic setup, where firms fail to fully internalise training returns, thus they offer less training than optimal. I present empirical evidence that the proposed channel actually plays a role in the determination of the training participation rate at the aggregate level. In particular, I find a statistically significant relation between the separation rate of workers and their training probability, after controlling for endogeneity.

The new theoretical framework introduces two different representative households, a skilled and an unskilled one, as well as progressive, and distortionary, taxation on households' total income. These innovations create a more realistic setting for policy analysis purposes.

To summarise, the main contributions of my work are three: (i) I build a new stylised model which features a range of endogenous channels affecting training provision and wage and productivity outcomes and a more complex fiscal menu at disposition of the policy-maker; (ii) I find novel and compelling evidence that training provision is negatively affected by the separation rate of workers, the so called poaching externality; and (iii) I provide a first quantitative evaluation of the effects of fiscal policies with respect to training subsidies that the UK government could implement.

Also, in comparison to Chapter 2, introducing two separate households, skilled and unskilled, allows to employ additional inequality measures, such as total income inequality or capital income inequality, and it has the advantage of not relying on an implicit consumption-insurance assumption, as with a single representative household. Such innovations allow to perform the welfare evaluation of fiscal policies related to training.

Within the new framework, training subsidies can be Pareto improving, yet there is a level of subsidization above which the distortions caused by income taxation outweigh the benefits of the intervention. Measuring welfare for *ex-ante* different households, it can be observed that the optimal subsidy is different for unskilled compared to skilled workers. For example, the former (latter) desire higher (lower) unskilled training subsidies. Even in this case, the government faces a trade-off between maximizing the total welfare and reducing the inequality between unskilled and skilled workers.

Once I impose a criterion for optimality, e.g. the average worker's welfare, I can identify the optimal combination of skilled and unskilled training subsidies. Both households will benefit by training subsidies, however only for small subsidy rates. For higher rates, average welfare may still increasing, but the benefits are unevenly distributed between workers. As a consequence, in this case, there is at least one group who opposes the policy reform.

Chapter 3 is organised as follows: I first introduce additional empirical motivation that supports the assumptions used for the model; then, Section 3.4 presents the new framework; finally, Section 3.5 reports the results of the several counterfactual exercises run for policy purposes.

3.2 Empirical motivation

Building on the results discussed in the previous chapters, the main intent of this section is to provide evidence suggesting that the level of training provided by companies is suboptimal. In this regard, Brunello and Gambarotto (2004) provides compelling evidence that firms located in denser areas of the UK are less willingly to train workers. In their view, the high likelihood that employees get poached by competitors reduces the incentive to train their workforce as they 'free ride' on training activities of other firms located in the same district. Technical reports from institutions are particularly concerned about the negative effects of workers' excess mobility which reduces firms' incentives to invest on their skills and leads to inefficient training outcomes, see e.g. OECD (1995) and the Sveriges Riksdag (1999), and Wolter and Stiftung (2018).

Academics have also shared this view, e.g. Booth and Snower (1996) and Acemoglu and Pischke (1999). Theoretical work, e.g. Katz and Ziderman (1990) and Stevens (1994), has shown that, under no commitment, the provision of general training is inefficient due to the poaching externality. In the same lines of thought, Moen and Rosén (2004) argue that, albeit worse than under commitment, the market equilibrium with poaching can be considered as 'constrained efficient'; that is the social planner would implement the same allocation as private firms, if all other choices were taken as given.[1] Yet, this result relies on the assumption that workers with different training levels separate into different sub-markets which, in turn, requires complete information to sort the workers according to their productivity and learning ability. The authors' opinion is that their assumption is restrictive, but may be partly verifiable. Hence, they do not exclude that training subsidies may be welfare improving (as long as sorting is imperfect).

[1]Under their set of assumptions, it is still possible to achieve the first best. However, to achieve the latter, a more complex policy is required. Depending on the model calibration, it might be necessary a specific combination of payroll taxes and either training subsidies or taxes.

Ideally, companies can engineer optimal contracts to internalise training externalities, for example making workers pay if they get hired by competitors, however it is a situation that rarely occurs due to limits to private contracting. The literature has observed few cases of formal enforcement. For example, Liebeskind (1997) reports that contracts that protect firm's knowledge from appropriation by rivals are costly and hard to enforce. When stakes are higher, e.g. to protect patents and R&D, companies do recur to non-compete clauses in employment contracts (see Marx (2011)), but this practice is limited to small segments of the labour market. In the case of firm-provided training, which has a general human capital component (see the literature review in Chapter 1), it is natural to expect that almost all workers have the alternative to change employer retaining valuable know-how and knowledge previously gained.

Empirically, this inefficiency channel has been found to have an economically relevant effect on apprenticeship provision in Germany. Dustmann and Schoenberg (2012) argue that apprenticeships with no commitment are only 28% as intensive as activities where firms have committed to high training levels. The low quality of the offer leads to fewer workers looking for apprenticeships as these programmes are less beneficial they would otherwise be. To strengthen this evidence, the next section studies the relationship between workers' separation rates and their training probability.

3.2.1 Poaching as source of training under-provision

The section shows that a higher separation rate affects negatively the workers' training probability. Using QLFS data, I create a cell "ID" based on the interaction between 8 industrial sectors, 9 occupational classes and 6 UK subregions.[2] The rationale for this choice is that what matters for firms are the local labour market conditions for a specific combination of skills. For example, a company which hires and trains software developers in Cambridge has to retain workers with specific IT skills and competencies (see e.g. Gathmann and Schönberg (2010)).

I compute the average separation rate of each cell by counting how many workers employed in a quarter are "new hires" in the next quarter and dividing it by the total number of workers.[3] Next, I assign to every observation (i.e. worker) the separation rate of the cell she belongs to.

The average training participation rate is negatively correlated with the cell's

[2] The starting point is the dataset employed in Chapter 1. Stata code is available on request. The classification is selected as to balance a greater level of detail with a sufficiently high number of responses within each cell. A finer classification, e.g. between sub-sectors or sub-regions, would compromise robust inference.

[3] As before, only employees who are 25 to 65 years old are kept. Also, I compute the separation rate only if there are more than 49 workers (for robustness) and I drop all observations with 0% separation rate. Only a negligible share of sectors has a 0% separation rate, i.e. less than 1% of the total observations. Thus, keeping these extra observations does not affect the results. Finally, I use only observations from 2005 to 2016 for two reasons: (i) to reduce the computational burden; and (ii) to have a more homogeneous reference period (also, this sub-sample is less unbalanced).

separation rate, but a standard regression could suffer from endogeneity of training and job-seeking outcomes. Notice that the relationship could even be positive as a higher turnover implies that there is a larger group of new hires that has to be trained. To control for endogeneity, I run an IV-probit model using three lags of the job separation rate as instrument variables of the current separation rate.[4] In more detail, the model I estimate is:

$$y_{i,j,t} = 1\left[\alpha_1 Z_{j,t} + X^0_{i,j,t}\bar{\alpha} + \alpha_j + \gamma_t + \varepsilon_{i,t} \geq 0\right] \tag{3.1}$$

in which $y_{i,j,t}$ is a binary outcome variable that describes whether a worker i in cell j at time t has received training; $1[\cdot]$ is an indicator function that assumes a value of either 0 or 1, depending on whether the latent variable inside the brackets is non-negative or positive; $X^0_{i,t}$ is a $1 \times L$ vector of exogenous explanatory variables; and the endogeneity of the separation rate $Z_{j,t}$ is addressed by estimating the first stage regression:

$$Z_{j,t} = \beta_j + \beta_1 Z_{j,t-1} + \beta_2 Z_{j,t-2} + \beta_3 Z_{j,t-3} + X^0_{i,j,t}\bar{\beta} + \epsilon_t \tag{3.2}$$

[4]Using any set of lags, from one to three, provides the same qualitative results, but different point estimates. Also, it is possible that the estimation sample differs depending on the number of lags set for the model.

where $Z_{j,t}$ is the job separation rate observed for cell j in the period t; $X^0_{i,j,t}$ is the set of exogenous variables that appear also in the main model and $\bar{\beta}$ is the $L \times 1$ parameter vector. Additionally, $(\varepsilon_{i,t}, \epsilon_t)$ is assumed to be distributed as a multivariate normal, $\sim N(\mathbf{0}, \Sigma)$, where, for identification purposes, the element σ_{11} is normalised to 1. For further assumptions on the validity of an IV-probit model and its solution methods see Amemiya (1978) and Wooldridge (2010).

Table 3.1: Effect of the separation rate on training probability

	skilled workers	unskilled workers
average marginal effect	-0.642	-0.330
p-value	0.00	0.00
implied elasticity	-0.038	-0.037

From Table 3.1, it can be observed that a higher separation rate causes the probability of training of the average worker to fall. The effect is small but significant at the 1% level. Also, the effect is twice as large for skilled workers than for unskilled workers. This is in line with the strand of literature that reports greater benefits from training for skilled workers, who are, *ceteris paribus*, more likely to change employers, and end up having a greater bargaining power when negotiating their wages (see e.g. Gertler *et al.* (2016)). The evidence suggests that, even though skilled workers are more likely to be trained, they suffer from the poaching externality more than unskilled workers and train less than the optimal amount.

I performed a series of robustness checks to validate my results. Using only one lag of the separation rate as instrumental variable, does not affect qualitatively the results (the coefficient is negative and statistically significant). However, using three lags of the separation rate as IV allows me to test for the over-identifying restrictions, that is I can verify whether the data suggests that my instrumental variables are endogenous. The Hansen J statistics has a p-value of 0.35, which means the null hypothesis that the instruments are valid cannot be rejected. Likewise, the Kleibergen-Paap LM and F tests do not provide any evidence that the estimation suffers from for weak instruments or under-identification.[5]

For calibration purposes, the elasticity is a more relevant information. I use the average separation rate and training rates to compute the elasticity of training to an increase in the separation rate for skilled and unskilled workers. Those are -0.038 and -0.037, respectively. As reported in Table 3.1, the elasticity is similar between the two groups of workers despite the marginal effect is almost twice as large for skilled workers. This is due to the fact that skilled workers have, *ceteris paribus*, lower separation rates and higher training participation rates.[6]

The new compelling evidence suggests that training is under-provided and mo-

[5] These tests can be computed in Stata with the command "ivreg2" after estimating a linear model. At the moment, the implementation of the tests does not support non-linear regression models, so running a linear model for the robustness checks is the best that can be done. Also, it must be noted that linear models have been frequently used in the literature on job-related training. References for the techniques employed can be found in Baum *et al.* (2018) and accompanying material.

[6] As usual, the elasticity is defined as $\frac{\Delta t}{\Delta s}\frac{s}{t}$, where s is the separation rate, t is the training participation rate, and $\frac{\Delta t}{\Delta s}$ is the average marginal effect reported above.

tivates building a general equilibrium model to study the effects of policies that address this inefficiency. That is the objective of Section 3.5. Before that, I discuss some of the assumptions imposed in the chapter.

3.3 Theoretical motivation

The models of Chapter 2 and Chapter 3 are based on the assumption that households, or agents, differ from one another only in respect to ex-ante characteristics (i.e. educational attainment). At a first glance, a model featuring (ex-post) heterogeneous workers sounds more appealing as it better captures the process that leads to training outcomes, e.g. think about the state dependence of training outcomes discussed in Chapter 1. However, I argue that the latter approach has a series of drawbacks and disadvantages, which lead me to opt for the representative agent fiction.

A main reason for such a choice is that previous works have found large externalities of training across workers.[7] For example, Metcalfe and Sloane (2007) find that increasing the education level of all co-workers by about one year results in larger wage increases (~12%) than increasing the worker's education by one year (~7%). Other works suggesting large spillovers from education of co-workers are

[7] This represents an extension of the strand of the literature looking at human-capital externalities.

Battu *et al.* (2003), for UK, and Bratti and Leombruni (2014) for Italy.[8]

All these interactions would be lost in a wage bargaining model à la Mortensen and Pissarides (1999), or, if included, they would make the model intractable Conversely, by using a closed-form equation for the accumulation of skills, I obtain a more compact and more limpid model. Also, I bypass any issue with this kind of externality, as their effects are already accounted for by the representative agent fiction. While a fancier model would require information that is not currently available, the calibration of my model is straightforward since it has predictions that can be directly compared to the empirical estimates I perform or report.

Lastly, thinking in terms of aggregate quantities, the attention can be focused on the poaching externality as a channel of inefficiencies in the training sector. The complexity of an heterogeneous agent model would not add sufficient benefits to justify its use.

3.4 The model

To evaluate the quantitative implications of policies that raise firms' incentives to train differently educated workers, I construct a dynamic general equilibrium

[8] These works are an extension to the literature that finds human-capital spillovers on average productivity (and wages), at the level of the sector, region, or even a whole country. See, among the others, the work of Acemoglu and Angrist (2001) and that of Moretti (2004a). A summary of the early literature can be found in Moretti (2004).

model that coheres with the main stylised facts relating to the UK job-related training and wage inequality data. The key features of the model are: (i) *ex ante* skill heterogeneity between non-University educated (unskilled) and University (skilled) household, leading to wage inequality under capital-skill complementarity in production; (ii) a government collecting revenues through a progressive income tax; (iii) training and skill creation undertaken by firms separately for skilled and unskilled workers; and, most importantly, (iv) under-provision of training at the steady state as firms fail to fully internalise training returns.

I calibrate the model in such a way that its steady state matches key quantities observed for the UK economy. Also, I verify that the model's predictions are in line with the empirical estimates reported in Section 3.2.[9] This constraint allows to identify the size of the externality in the theoretical model. Then, I simulate deterministic transition paths leading to new steady-states characterised by larger training subsidies.

A first exercise follows closely the one presented in Chapter 2, by considering separately unskilled and skilled training subsidies reforms. The exercise allows me to evaluate the effect of increasing subsidies to training in terms of efficiency, measured as total welfare gains, and equity, measured as either wage or income

[9]The model's dynamics after a permanent change in the poaching externality is reported in Appendix F.

premium. To complete the picture, I present an optimal policy exercise based on few basic objective functions, e.g. welfare, inequality, or productivity.[10] All the results are reported in Section 3.5.

3.4.1 The households

I introduce *ex-ante* heterogeneity between two households whose size is constant and equal to n_u and $n_s = 1 - n_u$, as total population is normalised to 1.[11] Both households can accumulate physical capital stock, but only skilled workers can purchase equity. With respect to the former, the two households face different (but finite) capital holding costs, while only the skilled household faces equity holding costs. These costs are modelled following closely the literature, e.g. Persson and Tabellini (1992) and Benigno (2009). The government collects revenues through a progressive income tax on total income of each household.

[10] This exercise is performed to emphasise that, depending on the policy-maker's goal, the optimal policy may be very different.

[11] In the current framework, workers' educational attainment is assumed to be exogenous. This assumption greatly simplifies the model and the general equilibrium solution. Also, I found no evidence suggesting that workers decide whether to study or to work depending on their training opportunities. Although a higher training provision leads to steeper wage profiles, returns to education are large enough to provide the strongest incentive to pursue a University degree or higher education. Arguably, higher training subsidies should impact workers' education decision only marginally; thus, the assumption can be considered a good approximation.

3.4.1.1 Skilled workers household

This infinitely-lived representative household is comprised of skilled members only. Thus, the superscript s distinguishes this household from the unskilled one. The head of the household makes all choices on behalf of its members. She maximises the discounted lifetime utility of its members:

$$U^s = \sum_{t=0}^{\infty} \beta^t \frac{\left\{ \left(c_t^s\right)^{\psi_s} \left[\left(1 - l_t^s\right) \right]^{1-\psi_s} \right\}^{(1-\sigma)}}{1-\sigma}, \tag{3.3}$$

where $0 < \beta < 1$ is the time discount factor (identical for both households); c_t^s is consumption; l_t^s is skilled labour supplied to firms; $\sigma > 1$ is the coefficient of relative risk aversion; and the parameter $0 < \psi_s < 1$ represents the weight that the household attaches to consumption and leisure $\left(1 - l_t^s\right)$ in utility. The household's budget constraint is:

$$c_t^s + I_t^s + z_{t+1}^s P_t^z = \left(w_t^s l_t^s + r_t k_t^s + z_t^s \pi_t\right)\left(1 - \tau_t^{h,s}\right) - \psi_s^k \left(k_t^s\right)^2 + z_t^s P_t^z - \psi_s^z \left(z_t^s\right)^2 - T_t^s, \tag{3.4}$$

where k_t^s is physical capital; $0 < \delta_k < 1$ is the capital depreciation rate; w_t^s is the skilled wage rate; r_t is the net return to capital; z_t^s is the (predetermined) share of firms' equity owned by the skilled household. In each period, the household earns a fraction of the firms' dividends π_t proportional to the shares it owns, then, given the price P_t^z, it can decide how many shares to hold, i.e. z_{t+1}^s. Under this assumption, the discount factor for the profits of the representative firm can be easily derived. This assumption has been used extensively in the literature, e.g. Merz and Yashiv (2007), Lee (2008), and Cheng and Lai (2015). The skilled household faces finite holding costs for both capital, ψ_s^k, and firm equity holdings, ψ_s^z. The latter is calibrated to match the value of equity to total wealth of skilled workers. Investments, I_t^s, drive the accumulation of physical capital stock according to the law of motion:

$$k_{t+1}^s = (1-\delta_k)\, k_t^s + I_t^s \tag{3.5}$$

The Lagrangian for the household's maximization problem is:

$$\Lambda^s = \textstyle\sum_{t=0}^{\infty} \left\{ \beta^t \frac{\left[\left(c_t^s\right)^{\psi_s} \left(1-l_t^s\right)^{1-\psi_s} \right]^{(1-\sigma)}}{1-\sigma} - \beta^t \lambda_t^{b,s} [c_t^s + k_{t+1}^s - (1-\delta_k)\, k_t^s + z_{t+1}^s P_t^z - \right.$$
$$\left. - \left(w_t^s l_t^s + r_t k_t^s + z_t^s \pi_t\right) \left(1-\tau_t^{h,s}\right) - z_t^s P_t^z + \psi_s^z \left(z_t^s\right)^2 + \psi_s^k \left(k_t^s\right)^2 + T_t^s \right] \right\} \tag{3.6}$$

where $\lambda_t^{b,s} > 0$, is the Lagrange multiplier for the household budget. The household chooses $\left\{c_t^s, l_t^s, k_{t+1}^s, z_{t+1}^s\right\}_{t=0}^{\infty}$ taking the initial conditions, $\left\{k_0^s, z_0^s\right\}$, fiscal policy $\left\{T_t^s, \tau_t^{h,s}\right\}_0^{\infty}$, prices, $\left\{w_t^s, r_t, P_t^z\right\}_{t=0}^{\infty}$, and profits $\{\pi_t\}_{t=0}^{\infty}$ as given. The static first-order condition (FOC) for consumption:

$$\lambda_t^{b,s} = \psi_s \left(\frac{1 - l_t^s}{c_t^s} \right)^{1-\psi_s} \left[\frac{\left(c_t^s\right)^{\psi_s}}{\left(1 - l_t^s\right)^{1-\psi_s}} \right]^{\sigma} \tag{3.7}$$

states that the shadow price of the budget constraint (3.4) is equal to the marginal utility of consumption, $\frac{\partial U^s}{\partial c_t^s}$, at time t.

The *intratemporal* FOC for skilled labour supply:

$$\frac{1 - \psi_s}{\psi_s} \frac{\left(c_t^s\right)^{1+\psi_s}}{\left(1 - l_t^s\right)^{-\psi_s}} = w_t^s \left(1 - \tau_t^{h,s}\right) \tag{3.8}$$

implies that the marginal rates of substitution between leisure and consumption at time t, i.e. $\frac{\partial U^s}{\partial \left(1 - l_t^s\right)} / \frac{\partial U^s}{\partial c_t^s}$, is proportional to the after-tax skilled wage rate.The Euler equation for capital:

$$\frac{\lambda_t^{b,s}}{\lambda_{t+1}^{b,s}} = \beta \left[1 - \delta_k + r_{t+1} \left(1 - \tau_{t+1}^{h,s}\right) - 2\psi_s^k \left(k_{t+1}^s\right) \right], \tag{3.9}$$

says that the marginal rate of substitution between consumption at time t and $t+1$, $\frac{\lambda_t^{k,s}}{\lambda_{t+1}^{k,s}} \equiv \frac{\partial U^s}{\partial c_t^s} / \frac{\partial U^s}{\partial c_{t+1}^s}$, is equal to the gross return to capital, $1 + r_{t+1}$, net of capital depreciation and a function of capital holding costs, all multiplied by the discount rate β. Finally, the first order condition with respect to the equity holdings:

$$P_t^z = \beta \frac{\lambda_{t+1}^{b,s}}{\lambda_t^{b,s}} \left[\pi_{t+1} \left(1 - \tau_{t+1}^{h,s} \right) + P_{t+1}^z - 2\psi_s^z \left(z_{t+1}^s \right) \right] \tag{3.10}$$

links the asset price at time t to its future price and the dividends it will pay out.

3.4.1.2 Unskilled workers household

The second infinitely-lived household is comprised of unskilled members only, and the superscript u is to denote such attribute. As above, the head of the household makes all choices on behalf of its members. In particular, she maximises the discounted lifetime utility:

$$U^u = \sum_{t=0}^{\infty} \beta^t \frac{\left[\left(c_t^u \right)^{\psi_u} \left(1 - l_t^u \right)^{1-\psi_u} \right]^{1-\sigma}}{1-\sigma}, \tag{3.11}$$

where $0 < \beta < 1$ is the time discount factor; c_t^u is unskilled consumption; l_t^u is the unskilled labour supplied to firms; $\sigma > 1$ is the coefficient of relative risk aversion; and the parameters $0 < \psi_u < 1$ represents the weight that the household attaches to consumption and leisure $\left(1 - l_t^u \right)$ in utility, respectively.

The household's budget constraint is:

$$c_t^u + I_t^u = \left(w_t^u l_t^u + r_t k_t^u\right)\left(1 - \tau_t^{h,u}\right) - \psi_u^k \left(k_t^u\right)^2 - T_t^u, \tag{3.12}$$

where k_t^u is physical capital; $0 < \delta_k < 1$ is the capital depreciation rate; w_t^u is the wage rate; r_t is the gross return to capital; $\psi_u^k > 0$ measures the holding costs for capital; and T_t^u is lump-sum transfers from the government. I assume that, unlike skilled workers, unskilled workers cannot purchase the equity of the firm. Such assumption does not affect quantitatively the results, and it allows to simplify the discount factor used by the representative firm.[12]

This assumption is also well grounded in empirical research and used in theoretical work. According to the Wealth and Assets Survey (WAS), in the years 2012-2014, unskilled workers owned a small amount of equity, which represents only 4% of their total wealth.[13] Also, empirical evidence suggests that non-University educated workers have little financial literacy. For example, Lusardi and Mitchell (2014) reports that college-educated individuals scored more than twice than high-

[12]Imagine that, if equity were shared between households, the discount factor would be a time-varying endogenous function of equity share, income tax rate, and marginal utility of consumption of each household.

[13]This figure is computed from the raw database, selecting individuals who are between 25 and 65 years old and controlling for educational attainment. Thus, the classification is as consistent as possible with the classification used for the data from the QLFS. In case of individuals with positive net wealth, I calculate the ratio between the financial (equity) assets and the total wealth. Then, I average the ratio across all individuals.

school educated individuals in a financial literacy test of US households. In turn, financial literacy strongly correlates with participation in stock markets (see Yoong (2010) and Arrondel *et al.* (2012) among the others). Although participation may influence the level of financial literacy, the literature favours the other direction of causality. For example, according to Campbell (2006), poorer and less educated workers are bad investors and, as they are aware of their limitations, they avoid purchasing equity. Hence, it is not a surprise if theoretical models generally assume that unskilled workers do not have access to equity markets (e.g. Favilukis(2013)) or any savings (e.g. Gali *et al.* (2004)).

To conclude the presentation of the unskilled household's constraints, investments I_t^u drive the accumulation of physical capital stock according to the law of motion:

$$k_{t+1}^u = (1-\delta_k)\, k_t^u + I_t^u. \tag{3.13}$$

Accordingly, the Lagrangian for the household is given by:

$$\begin{aligned} \Lambda^u = \textstyle\sum_{t=0}^{\infty} \Bigg\{ \beta^t \frac{\left[\left(c_t^u\right)^{\psi_u}\left(1-l_t^u\right)^{1-\psi_u}\right]^{1-\sigma}}{1-\sigma} - \beta^t \lambda_t^{b,u} [c_t^u + k_{t+1}^u - (1-\delta_k)\, k_t^u - \\ - \left(w_t^u l_t^u + r_t k_t^u\right)\left(1-\tau_t^{h,u}\right) + \psi_u^k \left(k_t^u\right)^2 + T_t^u] \Bigg\} \end{aligned} \tag{3.14}$$

where $\lambda_t^{b,u} > 0$ is the Lagrange multiplier for the budget constraint. The household chooses $\left\{c_t^u, l_t^u, k_{t+1}^u\right\}_{t=0}^{\infty}$ taking the initial condition, $\left\{k_0^u\right\}$, fiscal policy, $\left\{T_t^u, \tau_t^{h,u}\right\}_0^{\infty}$, and prices, $\left\{w_t^u, r_t\right\}_{t=0}^{\infty}$, as given. The first-order condition (FOC) for consumption:

$$\lambda_t^{b,u} = \psi_u \left(\frac{1-l_t^u}{c_t^u}\right)^{1-\psi_u} \left[\frac{\left(c_t^u\right)^{\psi_u}}{\left(1-l_t^u\right)^{1-\psi_u}}\right]^{\sigma} \tag{3.15}$$

states that the shadow price of the budget constraint (3.12) is equal to the marginal utility of consumption, $\frac{\partial U^u}{\partial c_t^u}$, at time t.

The *intratemporal* FOC for unskilled labour supply:

$$\frac{1-\psi_u}{\psi_u} \frac{\left(c_t^u\right)^{1+\psi_u}}{\left(1-l_t^u\right)^{-\psi_u}} = w_t^u \left(1-\tau_t^{h,u}\right) \tag{3.16}$$

implies that the marginal rate of substitution between leisure and consumption at time t, i.e. $\frac{\partial U^u}{\partial\left(1-l_t^u\right)} / \frac{\partial U^u}{\partial c_t^u}$, is proportional to the unskilled wage rate, net of income tax. The Euler equation for capital:

$$\frac{\lambda_t^{b,u}}{\lambda_{t+1}^{b,u}} = \beta \left[1 - \delta_k + r_{t+1}\left(1-\tau_{t+1}^{h,u}\right) - 2\psi_u^k \left(k_{t+1}^u\right)\right], \tag{3.17}$$

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

3.4.2 Representative firm

For the production sector, I assume that there is an infinitely-lived representative firm, which is owned by skilled households and which 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:

$$\widetilde{y}_t^f = A\left\{\mu\left(q_t^u\right)^\alpha + (1-\mu)\left[\rho\left(k_t^f\right)^\nu + (1-\rho)\left(q_t^s\right)^\nu\right]^{\frac{\alpha}{\nu}}\right\}^{\frac{1}{\alpha}}, \tag{3.18}$$

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; q_t^i is the effective labour input used in production $(i = s, u)$; k_t^f is the demand for capital; and $\alpha, \nu < 1$ are the parameters defining the factor elasticities, i.e. $1/(1-\alpha)$ is the elasticity of substitution between capital and effective unskilled labour as well as between effective skilled and effective unskilled labour; whereas $1/(1-\nu)$ 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 occurs because an increase in capital stock and capital augmenting technology 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 are used for production, as some of the workers' time is used for training purposes. Denoting the share of workers' time in job-related training by t_t^i implies that the net time actually used for production is given by $l_t^{f,i}\left(1-t_t^i\right)$, whereas $l_t^{f,i}t_t^i$ is the 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), the stocks of skills, accumulated via job-related training, evolve according to the following laws of motion:

$$h_{t+1}^u = \left(1-\delta^u-\delta_\varepsilon^u\right)h_t^u + \delta_\varepsilon^u\bar{h}_t^u + H^u\left(l_t^{f,u}t_t^u h_t^u\right)^{\gamma^u}, \tag{3.19}$$

and

$$h_{t+1}^s = \left(1-\delta^s-\delta_\varepsilon^s\right)h_t^s + \delta_\varepsilon^s\bar{h}_t^s + H^s\left(l_t^{f,s}t_t^s h_t^s\right)^{\gamma^s}, \tag{3.20}$$

where $0<\delta^u,\delta^s<1$ is the depreciation rate for skills accumulated by unskilled and skilled workers, respectively, which can be attributed to retirement of workers and the technological change which depreciates outdated skills; $0<\delta_\varepsilon^u,\delta_\varepsilon^s<1$ measure the flows of human capital due to poaching and mobility of workers across companies. The firm fails to internalise that the loss of skills due to poaching is compensated by hiring from a pool of workers with average human capital, $\bar{h}_t^u$ and $\bar{h}_t^s$, equivalent to that of its own employees. For simplicity, I assume that, at each time t, the net flow of workers is constant. In other words, the firm hires as many workers as those poached away by competitors.

These assumptions help to include an externality widely recognised by the empirical literature without introducing unemployment in the model. If unemployment were endogenous, it would be difficult to gauge the direct effect of training subsidies on income and welfare. This alternative setup increases the model's degrees of freedom, and, as a consequence, it requires an even larger set of assumptions and targets. Conversely, my approach allows to introduce the poaching externality with a precise upper bound on the size of the inefficiency. In particular, the human capital poaching must be a fraction of the average share of workers who move from one employer to another within a quarter, i.e. the job separation rate. As shown in Section 3.2, higher separation rates reduce the training probability of workers. The model replicates the empirical evidence. If the representative firm observes a decrease in the expected value of its training investments due to poaching, it reacts by curtailing the provision of training.

Respecting the notation introduced in Chapter 2, $H^i\left(l_t^{f,i} t_t^i h_t^i\right)^{\gamma^i}$ are the 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, H^i is normalised to unity while γ^i is calibrated as to capture differences in learning ability associated with University education.

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

The firm incurs in both an opportunity cost, in terms of foregone output, and a monetary cost when training its employees. The benefit for the firm is that the labour productivity generated by job-related training increases the effective labour input. In particular, the latter, q_t^i, is a function of workers' time and of labour productivity:

$$q_t^s = \left[l_t^{f,s} \left(1 - t_t^s\right) \right]^{\omega} \left[h_t^s \right]^{1-\omega}, \tag{3.21}$$

and

$$q_t^u = \left[l_t^{f,u} \left(1 - t_t^u\right) \right]^{\omega} \left[h_t^u \right]^{1-\omega}, \tag{3.22}$$

where $0 < \omega < 1$ measures the elasticity of effective labour with respect to production time.[14] Note that the constant returns to scale restriction in (3.21)-(3.22) implies that the production function (3.18) is also constant returns to scale in its five inputs $\left[l_t^{f,i} \left(1 - t_t^i\right), h_t^i \right]_{i=s,u}$, and k_t^f.

This setup implies that the firm pays for the training and it partly owns the job-related skills associated with h_t^i, thus capturing rents that are generated by job-related training. I use the expression "it partly owns" in that workers leaving the company bring with them some of the skills they acquired through training. To best understand the setup, suppose that training is indeed general, but firms have a strong bargaining power against the workers in setting the wages, thus they capture most of the training rents. The firm having the upper hand may be due to information asymmetry, in the sense that the rival company is uncertain about the quality of a poached worker and of the training he received. Kahn (2013) found

[14]It could be the case that ω is different between the skilled and the unskilled labour input. However, there is no clear evidence in favour of either assumption. Thus, for simplicity the parameter is assumed to be the same for both.

empirical evidence confirming these asymmetries among US firms. This aspect is captured by the model through the assumption that the representative firm is hiring equally trained workers (as the ones who left) but it fails to internalise this when it decides how much training to provide.

To summarise, the setup is consistent with the observation that (i) firms pay for most of job-related training of their employees, (ii) firms' productivity 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, and (iii) firms provide insufficient training to their workforce because of the risk of losing the skill they invested in.

Thus, in this specification, and given that the production function in (3.18) features 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. Notwithstanding the poaching externality, the relationship between ω and the firm's profitability is the same as in Chapter 2. Therefore, lower values of the parameter imply a higher contribution of the firm-owned factor h_t^i to production, increasing the profitability of the training investments.

The firm's problem is formalised as follows. The representative firm aims to maximise the present discounted value of lifetime profits:

$$\Pi^f = \sum_{t=0}^{\infty} Q_t \pi_t^f, \tag{3.23}$$

where:

$$Q_t = \beta^t \frac{\lambda_{t+1}^{b,s}}{\lambda_0^{b,s}} \frac{1 - \tau_{t+1}^{h,s}}{1 - \tau_0^{h,s}}, \tag{3.24}$$

defines the discount factor (for the present value at $t = 0$), $\forall t \geq 1$. The discount factor entails that the firms maximises its stakeholders' wealth, which is given by the sum of current after-tax profits and the market price of the equity, the latter being a function of discounted future after-tax dividends. The reason why tax rates matter in this context is quite straightforward. The value of equity depends on how high is the income tax on profits. This implies that firms may have an incentive to shift temporally profits through human capital (dis-)investment. For example, if the government commits to decreasing future skilled tax rates, the company will retain profits by investing more in training and capitalise the higher profits when those are taxed at the lower rate. The representative firm's profits are defined as:

$$\pi_t^f = \widetilde{y}_t^f - r_t k_t^f - w_t^u l_t^{f,u} - w_t^s l_t^{f,s} - \phi^u t_t^u l_t^{f,u} \left(1 - \tau_\phi^u\right) - \phi^s t_t^s l_t^{f,s} \left(1 - \tau_\phi^s\right), \tag{3.25}$$

the revenue from selling the final good, minus the costs of capital, the costs of unskilled and skilled labour, as well the training monetary costs. The parameter $0 < \phi^i < 1$ refers to the fixed monetary cost per training hour; and τ_ϕ^i is a subsidy or tax on training activities. The *intertemporal* trade-off associated with training time is evident in equations (3.18)-(3.22) and (3.25). All else constant, an increase in training time raises new skills at time t and the stock of skills in $t + 1$. Hence, in $t + 1$, both effective labour and output increases. However, training incurs a resource outlay. In addition, at time t, by lowering the time available for work at time t, effective labour and output fall.

The setup creates different incentives for the firm to train its skilled and unskilled employees which can be matched to the UK data. Since the employees have different marginal products of effective labour, there is a 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. the evidence in Almeida and Faria (2014)).

On the other hand, if training unskilled workers implies a relatively higher monetary cost, $\phi^u > \phi^s$, then the firm has more incentives to train skilled than to train unskilled workers. Finally, as skilled workers are more likely to change job and be employed in a new company, i.e. $\delta_{\varepsilon}^{s} > \delta_{\varepsilon}^{u}$, the disincentives to invest in training due to the poaching externality are (slightly) stronger in the case of skilled workers, consistently with the results of Section 3.2.

In equilibrium, the relative training allocation between skilled and unskilled workers depends on the quantitative evaluation of this trade-off, which, as explained above, depends on a multiplicity of concurring factors.

Given that several factors affects the provision of training, I could have envisioned a range of policies alternative to subsidies to monetary training costs, such as increasing the productivity of workers with extra education or off-the-job training. However, as discussed earlier, firm-provided training is the most effective form of employees formation compared to training spells paid by the government or by the household. Also, it is not clear how the new acquired skills will affect the training provision of firms. They could be complementary or substitute to the accumulation of firm-provided skills. Finally, the literature has shown that training costs have a big impact on the decision of firms to provide training, e.g. Muehlemann *et al.* (2007) and Blatter *et al.* (2012).[15] As a consequence, it is most natural to consider training subsidies as a tool for policy-makers to influence

[15]Furthermore, Muehlemann *et al.* (2010) shows that Swiss companies that do not provide training are in average facing higher training costs than companies offering training to their employees.

the behaviour of firms.[16]

Taking the initial conditions, $\left\{k_0^f, h_0^s, h_0^u\right\}$, prices, $\left\{w_t^s, w_t^u, r_t\right\}_{t=0}^{\infty}$, the fiscal stance $\left\{\tau_t^s, \tau_t^u\right\}_{t=0}^{\infty}$, 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 (3.23), subject to (3.19) and (3.20). The Lagrangian for the firm is given by:

$$\begin{aligned}\Lambda^f = \textstyle\sum_{t=0}^{\infty}\{Q_t[y_t^f - r_t k_t^f - w_t^u l_t^{f,u} - w_t^s l_t^{f,s} - \\ -\phi^u t_t^u l_t^{f,u}\left(1-\tau_t^u\right) - \phi^s t_t^s l_t^{f,s}\left(1-\tau_t^s\right)] - \\ -Q_t\lambda_t^u[h_{t+1}^u - \left(1-\delta^u-\delta_\varepsilon^u\right)h_t^u - \delta_\varepsilon^u\bar{h}_t^u - H^u\left(l_t^{f,u}t_t^u h_t^u\right)^{\gamma^u}] - \\ -Q_t\lambda_t^s[h_{t+1}^s - \left(1-\delta^s-\delta_\varepsilon^s\right)h_t^s - \delta_\varepsilon^s\bar{h}_t^s - H^s\left(l_t^{f,s}t_t^s h_t^s\right)^{\gamma^s}]\},\end{aligned} \tag{3.26}$$

where λ_t^i, for $i = s, u$, are the shadow prices associated the skill accumulation constraints (3.19) and (3.20); and y_t^f is given by:

$$\begin{aligned}y_t^f = A_t\left\{\mu\left(\left[l_t^{f,u}\left(1-t_t^u\right)\right]^{\omega}\left[h_t^u\right]^{1-\omega}\right)^{\alpha} + \right.\\ \left. +\left(1-\mu\right)\left[\rho\left(k_t^f\right)^{\nu} + \left(1-\rho\right)\left(\left[l_t^{f,s}\left(1-t_t^s\right)\right]^{\omega}\left[h_t^s\right]^{1-\omega}\right)^{\nu}\right]^{\frac{\alpha}{\nu}}\right\}^{\frac{1}{\alpha}}.\end{aligned} \tag{3.27}$$

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

$$r_t = \frac{\partial y_t^f}{\partial k_t^f}, \tag{3.28}$$

[16]Remember that more coercive policies, such as the train-or-pay policy, have been dismissed due to the protest of private sector. So, I exclude this type of intervention as well.

$$w_t^u + \phi^u t_t^u \left(1 - \tau_t^u\right) = \frac{\partial y_t^f}{\partial l_t^{f,u}} + \lambda_t^u \frac{\partial h_{t+1}^u}{\partial l_t^{f,u}}, \tag{3.29}$$

and

$$w_t^s + \phi^s t_t^s \left(1 - \tau_t^s\right) = \frac{\partial y_t^f}{\partial l_t^{f,s}} + \lambda_t^s \frac{\partial h_{t+1}^s}{\partial l_t^{f,s}}, \tag{3.30}$$

equate their respective marginal costs to their marginal products.[17] 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:

$$\phi^u l_t^{f,u} \left(1 - \tau_t^u\right) - \frac{\partial y_t^f}{\partial t_t^u} = \lambda_t^u \frac{\partial h_{t+1}^u}{\partial t_t^u}, \tag{3.31}$$

[17] As in Chapter 2, some of the partial derivatives are not fully expressed to provide a more intuitive understanding of the first order conditions. The formulas for these derivatives are reported in Appendix D.

$$\phi^s l_t^{f,s}\left(1-\tau_t^s\right) - \frac{\partial y_t^f}{\partial t_t^s} = \lambda_t^s \frac{\partial h_{t+1}^s}{\partial t_t^s}, \tag{3.32}$$

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.

Finally, the Euler equations for unskilled and skilled skills acquisition:

$$\lambda_t^u = \frac{Q_{t+1}}{Q_t}\left(\frac{\partial y_{t+1}^f}{\partial h_{t+1}^u} + \lambda_{t+1}^u \frac{\partial h_{t+2}^u}{\partial h_{t+1}^u}\right), \tag{3.33}$$

$$\lambda_t^s = \frac{Q_{t+1}}{Q_t}\left\{\frac{\partial y_{t+1}^f}{\partial h_{t+1}^s} + \lambda_{t+1}^s \frac{\partial h_{t+2}^s}{\partial h_{t+1}^s}\right\}, \tag{3.34}$$

state that the shadow price of skill acquisition at time t, λ_t^i is equal to the discounted value of the net benefits of skill accumulation, i.e. $\frac{\partial y_{t+1}^f}{\partial h_{t+1}^i} + \lambda_{t+1}^u \frac{\partial h_{t+2}^i}{\partial h_{t+1}^i}$, 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 $t+1$, i.e. λ_{t+1}^i.

3.4.3 Government budget

To focus on policies to reduce training inequality, the government's balanced budget constraint is assumed to be:

$$\tau_t^u \left(\phi^u t_t^u l_t^{f,u}\right) + \tau_t^s \left(\phi^s t_t^s l_t^{f,s}\right) = n^u \left[T_t^u + \left(w_t^u l_t^u + r_t k_t^u\right) \tau_t^{h,u}\right] + \\ + n^s \left[T_t^s + \left(w_t^s l_t^s + r_t k_t^s + z_t^s \pi_t\right) \tau_t^{h,s}\right], \qquad (3.35)$$

where the sum of total-income taxes, net of lump-sum transfers, is equal to the total expenditure for training subsidies. I assume that any lump-sum transfer is equally divided between the two households according to their size, imposing that $n_s T_t^s = n_u T_t^u = T_t/2$. To ensure that the government budget is always balanced, the lump-sum transfers are the residual policy instrument, unless stated otherwise.

3.4.4 Market clearing conditions

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

$$k_t^f = k_t^u n^u + k_t^s n^s, \qquad (3.36)$$

$$l_t^{f,u} = n^u l_t^u, \tag{3.37}$$

$$l_t^{f,s} = n^s l_t^s, \tag{3.38}$$

$$z_t^f = n_s z_t^s = 1, \tag{3.39}$$

$$\pi_t^f = n_s z_t^s \pi_t = \pi_t, \tag{3.40}$$

and

$$\begin{aligned} y_t^f &= n_s c_t^s + n_u c_t^u + n_s \left[k_{t+1}^s - (1-\delta_k) k_t^s\right] + n_u \left[k_{t+1}^u - (1-\delta_k) k_t^u\right] + \\ &+ \phi_u t_t^u l_t^{f,u} + \phi_s t_t^s l_t^{f,s} + n_s \psi_s^z \left(z_t^s\right)^2 + n_s \psi_s^k \left(k_t^s\right)^2 + n_u \psi_u^k \left(k_t^u\right)^2 . \end{aligned} \tag{3.41}$$

3.4.5 The decentralised equilibrium

Given initial conditions, the decentralised equilibrium is defined to be an allocation $\left\{c_t^s, c_t^u, l_t^u, l_t^s, l_t^{f,u}, l_t^{f,s}, \pi_t^f, \pi_t, k_t^f, k_t^s, k_t^u, t_t^u, t_t^s, h_{t+1}^u, h_{t+1}^s\right\}_{t=0}^{\infty}$, prices $\left\{r_t, w_t^u, w_t^s, P_t^z\right\}_{t=0}^{\infty}$, shadow prices $\left\{\lambda_t^{b,s}, \lambda_t^{b,u}, \lambda_t^s, \lambda_t^u\right\}_{t=0}^{\infty}$, and policy instruments $\left\{T_t, \tau_t^s, \tau_t^u, \tau_t^{h,s}, \tau_t^{h,u}\right\}_{t=0}^{\infty}$ 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. At the aggregate level, the representative firm skill-stock variables, h_t^u and h_t^s, are equal to the economy-wide variables $\bar{h}_t^u$ and $\bar{h}_t^s$.

Using Walras' law I discard the skilled household's budget constraint, thus the decentralised equilibrium consists of the following 27 equations: (i) the two households' FOCs, i.e. the 3-equations (3.15)-(3.17) and the 4-equations (3.7)-(3.10); (ii) the firm's 2-skill accumulation equations (3.19)-(3.20); (iii) the firm's 7-FOCs, i.e. equations (3.28)-(3.34); (iv) the government's budget constraint, represented by equation (3.35); (v) the 6-market clearing conditions, i.e. equations (3.36)-(3.41); and, finally, the unskilled household budget constraint (3.12). To reduce the size of the model, I drop the equity stock, z_t^s, which is constant by assumption, and I use a single variable, T_t, to represent lump-sum transfers.

3.4.6 Calibration

I set the parameters appearing in the DE equations with the overall aim that the model generates a steady-state solution which implies quantities similar to the data for the UK. The parameters are summarised in Table 3.2. These are divided in three groups: one for the households, the second for the firm, and the last for physical and human capital.[18]

The productivity parameters which work as scaling factors $\{A, H^u, H^s\}$ are all normalised to unity and omitted from the table. Following many dynamic general equilibrium studies, I set the coefficient of relative risk aversion $\sigma = 2$.[19] The

[18]Note that the first four parameters of the household, the first two parameters of the firm, and the first three parameters of the capital are set. All the others parameters are calibrated, as explained below.

[19]As mentioned in Chapter 2, Browning *et al.* (1999), Ionescu (2009) and Bakış *et al.* (2015)

common discount factor β is set so that the interest rate on capital is equal to about 2.9% at the steady state, which is equivalent to the average risk free interest rate for UK over the last fifteen years before the great recession.[20]. The unskilled and skilled population shares are the average shares from the last two decades of data.

The coefficient for the capital holding cost for the skilled household, ψ_s^k, is normalised to $1.3e^{-04}$ so that it has negligible effect on the resources allocation.[21] By setting $\psi_u^k = 4.5e^{-04}$, I pin down the ratio of the stock of physical capital of skilled workers over that of unskilled workers, which is 2.2. Also, I set equity holding costs, ψ_s^z, equal to $2.7e^{-04}$ to target the ratio of physical capital to total wealth for skilled households, that, at the time of the fourth Wealth and Assets Survey (WAS), was about 92%. As a consequence, holding equity is more expensive than holding physical capital.

The poaching coefficients for the skill capital are the most critical parameters. To identify an upper bound, I estimate the job-to-job transition rates of skilled and unskilled workers, respectively, from the longitudinal QLFS dataset. For simplicity, I assume that a worker has been poached if she is employed for two

are some of the works which use a value of 2 for the households' risk aversion coefficient.

[20]The data source is reported in Appendix A

[21]If ψ_s^k is any smaller than the specified value, MATLAB may fail to converge to the solution for the steady state due to approximation. Larger values could be used, but that implies bringing aggregate values, e.g. $\frac{\bar{k}}{\bar{y}}$, away from their empirical counterparts.

Table 3.2: Calibration

symbol	value	definitions
		Household
β	0.9985	quarterly time discount factor
δ^k	0.025	quarterly capital depreciation rate
σ	2	coefficient of relative risk aversion
n^u	0.660	share of unskilled to total household members
ψ_u	0.385	leisure weight in utility (unskilled)
ψ_s	0.418	leisure weight in utility (skilled)
ψ_s^k	1.3e-04	capital holding costs (skilled)
ψ_s^z	2.7e-04	equity holding costs (skilled)
ψ_u^k	4.5e-04	capital holding costs (unskilled)
		Firm
ν	-0.495	effective skilled labour to capital substitution parameter
α	0.401	effective unskilled labour substitution parameter
ω	0.938	elasticity of effective labour with respect to time
μ	0.564	share of composite input to output
ρ	0.869	share of capital to the composite input
		Capitals
δ_k	0.025	depreciation rate for physical capital
δ^u	0.025	depreciation rate for accumulated skills (unskilled)
δ^s	0.025	depreciation rate for accumulated skills (skilled)
δ_ε^s	5.39e-04	firm poaching (skilled)
δ_ε^u	4.91e-04	firm poaching (unskilled)
ϕ^u	4.469	fixed cost per training hour (unskilled)
ϕ^s	3.953	fixed cost per training hour (skilled)
γ^u	0.501	returns to scale for creating new skills (unskilled)
γ^s	0.531	returns to scale for creating new skills (skilled)

consecutive quarters but her (self-reported) tenure in the subsequent quarter is less than three months, in other words she is a new hire in the subsequent quarter.[22]

In the UK from 2005 to 2016 the average quarterly separation rate has been

[22]Potentially, this leads to overestimating job-to-job transitions, as a laid-off worker may find job within the interview period and be confused for someone who was hired by competitors. However, it can be noted that more than 50% of movers report that they resigned from their previous job, while only about 15% report that they had a temporary job or they were made redundant.

2% and 1.6% for skilled and unskilled workers, respectively. Setting δ_ε^i to match workers' separation rate entails the additional assumption that all workers possess the same average skill-stock, but it is well known that trained workers are less likely to leave the company and that training tends to reduce the probability of changing employer. For that reason, a realistic target for these parameters should be a fraction of the observed job-to-job transition rate.

The parameters can be set in such a way that the model simulated moments match the empirical data on the poaching externality. The larger δ_ε^i is, the stronger is the elasticity of training to a shock to δ_ε^i. Thus, I set $\delta_\varepsilon^s = 5.4e^{-4}$ and $\delta_\varepsilon^u = 4.9e^{-4}$, so that the model replicates the elasticities estimated in Section 3.2 with a negligible margin of error.[23]

Table 3.3 reports the set of parameters related to the fiscal policy tools. The total income tax rate for skilled and unskilled households is obtained by targeting two observables. The analysis of the British Household Panel Survey (BHPS) data suggests that the ratio between the average skilled labour tax rate and the average unskilled labour tax rate is about 1.17 (thus, on average, skilled workers face a 17% higher tax rate). Furthermore, Piketty and Saez (2006) report that the total effective income tax rate for UK in the reference year 2004 (which is in the midst of the sample period) is 23.7%. Given the evidence, I set the tax rate

[23] These are -0.038 and -0.037 for skilled and unskilled workers, respectively.

26.2% and 22.4% for skilled and unskilled household income, respectively.

As reported in Chapter 2, The UK training subsidies amount to 2.4% of the monetary training costs. This sum is much less than the revenues from the income tax. As a consequence, the residual is rebated to households in form of lump-sum transfers.[24]

Table 3.3: Calibration of fiscal policy

		policy
τ^u	0.024	public subsidy for training activities (unskilled)
τ^s	0.024	public subsidy for training activities (skilled)
$\tau^{h,u}$	0.224	total income tax (unskilled)
$\tau^{h,s}$	0.262	total income tax (skilled)
$\bar{T}$	-0.131	lump-sum transfers

Since the depreciation of job-related skills is hard to measure, I assume $\delta^s = \delta^u = 0.25$. 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%

[24]Unspent revenues could be considered as government spending with limited consequences on the analysis that follows.

and 3.26%. Thus, the value of 2.5% is in-between these estimates. The physical capital depreciation is set to 2.5 (10% yearly), a rate that is frequently employed in the literature (see e.g. Basu and Thoenissen (2007), Leeper *et al.* (2010), and Andreasen (2012)) and matches the one set for the skill capital depreciation.

Next, I discuss the group of parameters relating to training and production $\{\nu, \alpha, \mu, \rho, \phi^u, \phi^s, \gamma^u, \gamma^s, \omega\}$, starting with ω, which is linked to firms' returns on job-related training, i.e. the firms' rent. As already discussed, estimates on the profitability of job-related training do not exist for the UK. Blundell *et al.* (1999) measure a 10% private return to participating to job-related training in the UK 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. Thus, ω is chosen by relating firm profitability to a valuation of the return to 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. $\frac{\pi_t}{\phi^u t_t^u l_t^{f,u}(1-\tau^u)+\phi^s t_t^s l_t^{f,s}(1-\tau^s)+w_t^u t_t^u l_t^{f,u}+w_t^s t_t^s l_t^{f,s}}$. Almeida and Carneiro (2009) estimate the return to be between 8.6 and 13.8 percentage points for training firms in Portugal. Setting $\omega = 0.938$ ensures that, in conjunction with the remaining parameters, the rate of return of training investments is about 9 percentage points.

Table 3.4: Steady-state

variable	definition	model	data
$\frac{(1-\tau^{h,s})w^s}{(1-\tau^{h,u})w^u}$	skill premium	1.578	1.589
t^s	skilled training to total time share	0.022	0.023
t^u	unskilled training to total time share	0.013	0.013
t^s/t^u	training differential	1.726	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	0.016	0.017
l^s	skilled labour to total hours	0.315	0.310
l^u	unskilled labour to total hours	0.292	0.290
k/y	capital-to-output	8.20	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	0.027	0.026
$\frac{\tau^s\phi_s t^s l^s n^s+\tau^u\phi_u t^u l^u n^u}{y}$	public spending on training costs-to-output	0.0006	0.0006
rk/y	capital income-to-output (model gross tax)	0.322	0.299
$\frac{n^s l^s w^s+n^u l^u w^u}{y}$	labour income-to-output (model gross tax)	0.647	0.671

The remaining parameters in the production function are chosen so that the model's steady-state solution is consistent with factor income shares and inequality indices. In particular, I 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. Thus, for this set of parameters, the calibration approach follows closely the one used in Chapter 2. The steady-state solution implied by the parameters in Table 3.2 is summarised in Table 3.4. As can be seen, the model's predictions for the long-run quantities are close to the targets, except for the capital to output ratio. However, the model generated quantity is consistent with the ratio obtained using a more restrictive definition of productive capital.[25]

3.5 Evaluation of fiscal policies

I use the framework developed above to evaluate a range of fiscal policies with respect to training activities. The main policy intervention is a fiscal subsidy to either skilled or unskilled training. To compensate the extra costs, I assume that the government can finance its spending in different ways: (i) taxing both skilled and unskilled workers (in which case the progressivity of the tax system is preserved), or (ii) taxing only skilled workers. In Appendix G, I produce a counterfactual experiment where the same fiscal subsidy is financed through lump-sum transfers from both households to net out the effects of distortionary taxation on income. Lastly, to complete the picture, I consider the optimal training subsidies in the case they are financed by taxes on both households (first case).

In the next sections, I start presenting the impulse response functions obtained from the model's simulation to provide an intuitive interpretation of he fiscal policy effects. Then, I report a more detailed analysis of each policy intervention providing quantitative results.

[25]It is also possible to get closer to the 10:1 capital-output ratio by either increasing the steady-state interest rate or decreasing the physical capital depreciation. This has not been done to allow a closer comparison between the model in Chapter 3 and the one in Chapter 2.

3.5.1 Impulse Response Functions

To perform a preliminary inspection of the economy's reaction to training subsidies, I assume that the economy is at steady state and I let the policy reform occur after the first period. Then, I simulate the path of endogenous variables from the initial steady state to the new equilibrium. The results are reported in the following subsections. I consider the effects of unskilled and skilled training subsidies, respectively. In each case, I show an arbitrarily large intervention, that covers one fourth of monetary training costs. The intervention is financed according to either one of the two scenarios mentioned above. In addition, I consider a 1% increase in progressivity of UK tax system as a measure to reduce inequality alternative to the training subsidy policies.

3.5.1.1 Permanent increase in unskilled training subsidy

Figure 3.1 shows the dynamics of key endogenous variables after a permanent increase in unskilled training subsidies from 0.024 to 0.25. The blue line represents the impulse response functions (IRFs) when the subsidy is financed through increasing both skilled and unskilled income tax rates equally, whereas the orange line represents the dynamics of the model when the subsidy is financed by skilled workers only.[26] The skilled and unskilled wages reported in the figure are before tax so that I can compare the two scenarios. The wage inequality gives an idea of the overall redistributive effect, since it is measured as the ratio between skilled and unskilled after-tax wages.

It can be observed that, if government subsidises unskilled training (raising taxes on both households), unskilled workers are offered higher wages as the demand of unskilled labour input and unskilled training time increases (see sub-plots (2,3) and (4,5), respectively).[27] This reduces the wage premium and, even more so, labour income inequality through its indirect effect on labour supply (see sub-plot (3,3), (3,4), and (2,3), respectively).

The policy thus helps reduce the market failure caused by the training externality as it increases the provision of training for unskilled workers (see sub-plot (4,5)). The effects of the policy on skilled training is less clear, as it depends on the financing. In both cases, the policy has a very limited effect on skilled training, as it deviates from steady state only marginally (see sub-plot (4,4)).

In the first scenario (both taxed), the skilled household is initially subjected to a loss of utility due to the taxation, however in the long-run the loss is smaller (see

[26]In Appendix G, Figure G.1 compares the dynamics of the ”both taxed” case with the case where the policy is financed through lump-sum taxes.

[27]Note that, if taxes are raised only on the skilled household, we observe a different short term dynamics but over the long run the effects are qualitatively the same.

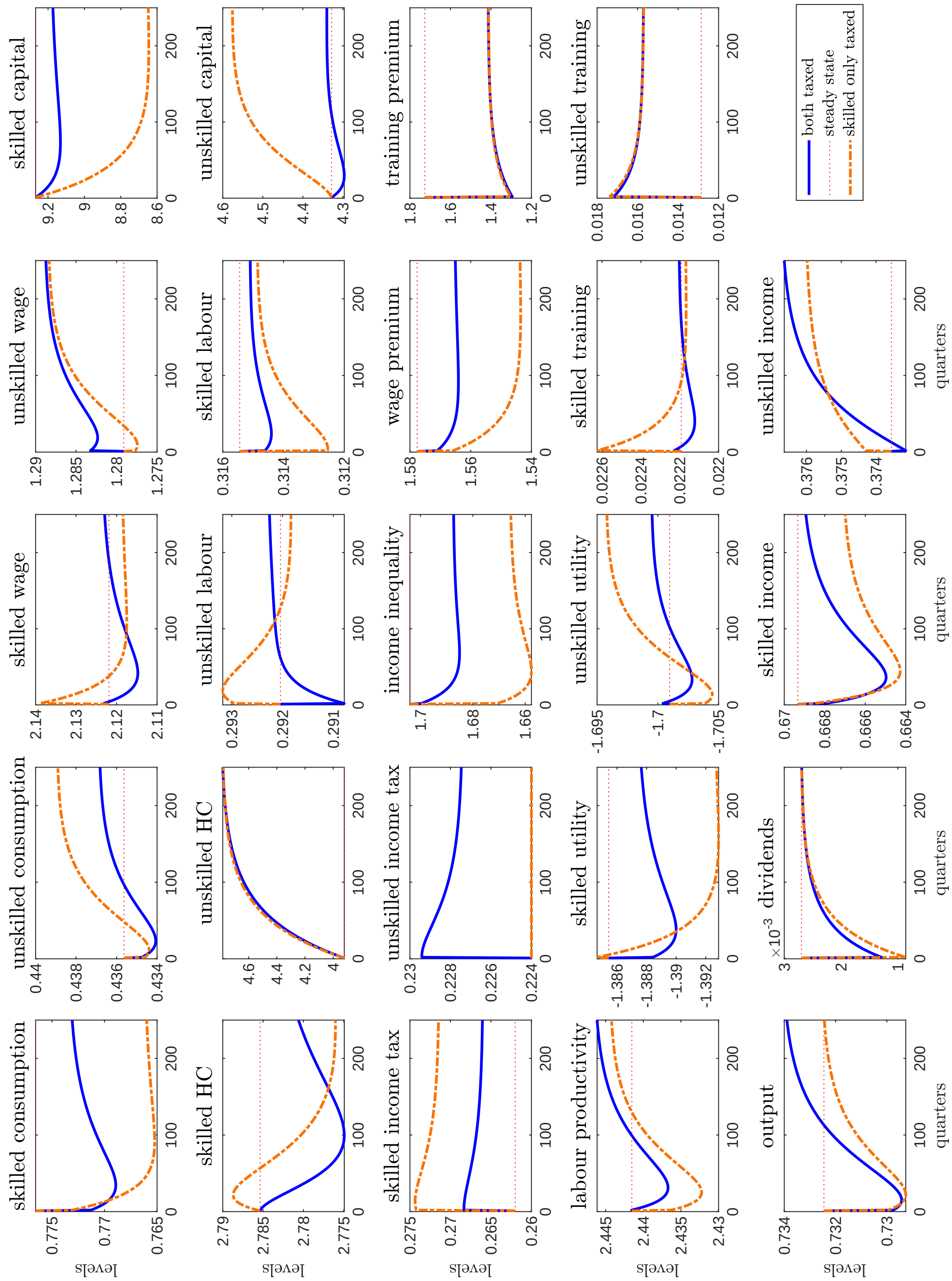


Figure 3.1: Permanent increase of unskilled training subsidy to 0.25

sub-plot(4,2), blue line). Also, this loss is largely compensated by the increase in unskilled workers' utility.

When the subsidy is financed entirely by the skilled household (second scenario), the decrease of utility for this group is much larger and persistent. In this case, the skilled household is worse off both in the long and in the short run (see sub-plot (4,2), orange line).

As the subsidy helps the unskilled group, at impact, their utility increases above the original steady state, it decreases below steady state in the short run, when the costs of higher taxation are felt most, then it increases substantially in the long run, as the benefits of higher human capital stock are accrued (see sub-plot (4,3)). Thus, the benefits of the policy, in terms of utility, materialise only in the long run, but the reduction of wage and income inequality is immediate (see sub-plots (3,3) and (3,4)). This is true under both scenarios.

In the long run, in both cases, labour productivity is positively affected by the policy due to the higher stock of physical and skill capital (see sub-plot (4,1) compared to sub-plots (1,5), (2,5), (2,1), and (2,2)).[28] The dynamics shows that initially productivity falls due to: (i) more training time for unskilled workers to accumulate skills; (ii) higher monetary investments on training; and, consequently,

[28]With respect to the latter, the average worker's skills are increased since the decrease of skilled human capital is more than compensated by the increase in unskilled human capital.

(iii) higher tax rates causing a temporary contraction of consumption and output.

The pattern for unskilled training is almost identical across the two scenarios, whereas skilled training follows different paths which eventually converge towards similar steady-state values (see sub-plot (4,5)).[29]

With respect to the redistribution of consumption between households, the policy has a much stronger effect when only skilled workers are paying for the extra subsidy (see sub-plots (1,1) and (1,2), orange lines). The high tax rate leads skilled workers to reduce the provision of labour (see sub-plot (2,4), orange line). In the second scenario, the substitution effect should be stronger due to the tax hike, yet the reduction of labour supply is quantitatively similar, suggesting that, in this case, a stronger income effect induces the skilled household to supply more labour.

3.5.1.2 Permanent increase in skilled training subsidy

Figure 3.2 shows the effects of a subsidy to skilled training that covers one quarter of monetary training costs. The blue line represents the impulse response functions (IRFs) when the subsidy is financed through increasing both skilled and unskilled income tax rates equally, whereas the orange line represents the dynamics of the

[29]Remember that each scenario differs by the source of financing: the subsidy costs can be paid by both skilled and unskilled households, only the skilled household, or through lump-sum transfers – the last case being reported in Appendix G.

model when the subsidy is financed entirely by skilled workers.[30]

As can be noted, the economy reacts as it did for the unskilled subsidies. Yet, this time, inequality measures move in the opposite direction, i.e. upwards in favour to skilled workers (see sub-plots (2,5), (2,4), and (2,3)). With respect to the short term dynamics, the main difference is that, after an increase in unskilled training subsidies, the unskilled labour supply contracts under the first scenario and expands under skilled-only financing, whereas, after the increase in skilled training subsidies, skilled labour supply always fall below steady state (see sub-plot (2,4)).[31]

This breaks the qualitative symmetry of the IRFs between unskilled and skilled training subsidies. Also, in the case of shared financing, it helps explain why, after the skilled subsidy policy, the wage premium does not increase monotonically towards the new steady state but it jumps up at impact and it deflates in the medium run before converging to its new long-run equilibrium (see sub-plot (3,4), blue line).

Another difference is that the aggregate effects produced by the skilled subsidies

[30]Figure G.2 in Appendix G, shows the dynamics of the endogenous variables in the case the policy is financed through lump-sum taxes, contrasting it with the first scenario.

[31]In the long term, it is the skilled labour supply that is always above the old steady state after the increase in skilled training subsidy. Conversely, the unskilled labour supply is below the old steady state if the unskilled training subsidy is financed by skilled workers and above in the other case.

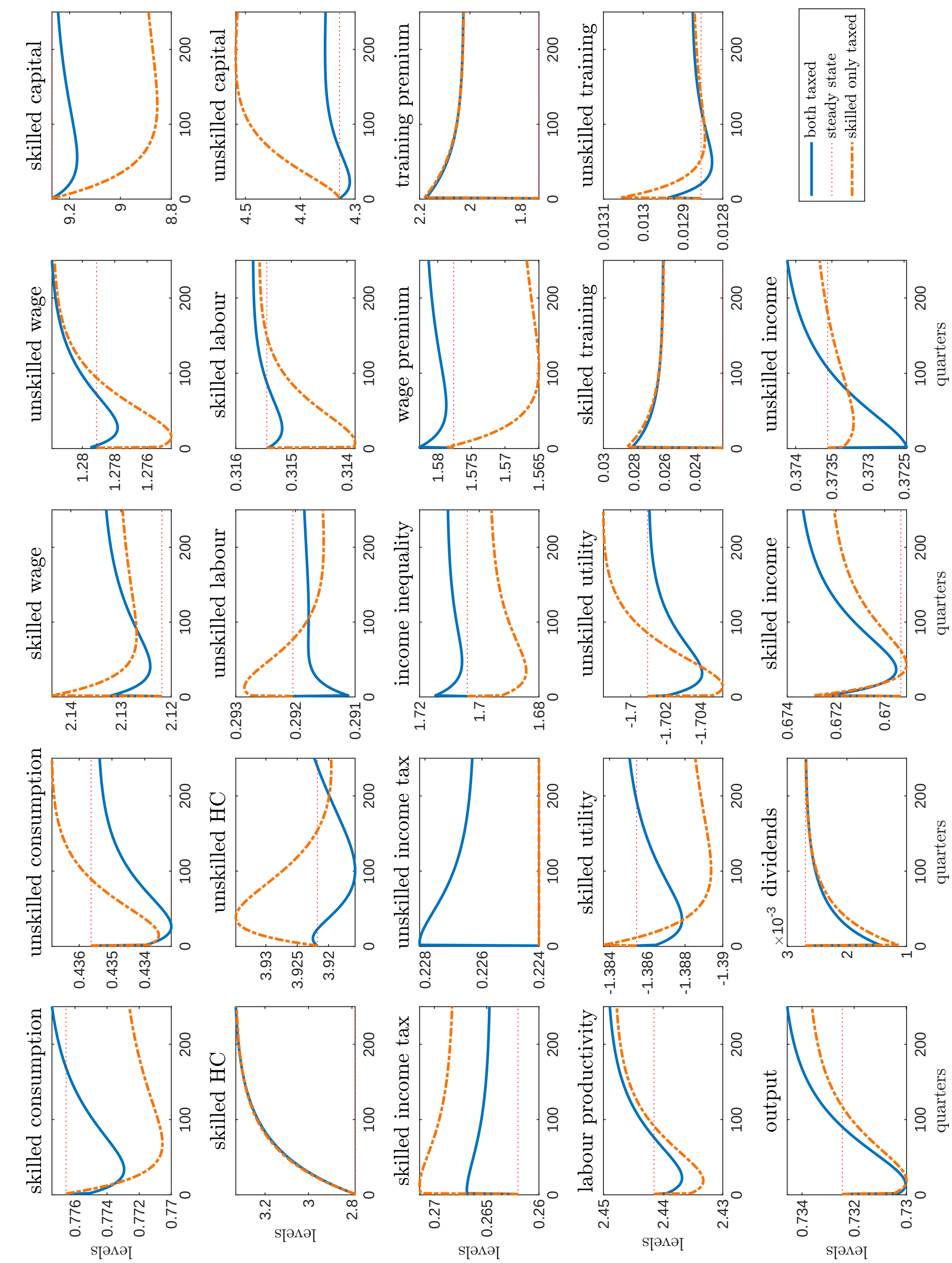


Figure 3.2: Permanent increase of skilled training subsidy to 0.25

are marginally larger than those caused by the unskilled subsidies, especially in terms of labour productivity and capital stock accumulation (see sub-plot (4,1) and sub-plots (1,5) and (2,5), respectively). This can be explained by a series of factors: (i) the physical capital complementarity; (ii) the stronger poaching externalities that characterise the skilled labour input; and (iii) the distortions due to income tax. These differences are quantified and commented in Section 3.5.2.

3.5.1.3 Permanent increase in tax progressivity

I simulate an increase in the progressivity of the tax system to compare this alternate policy with the training subsidy policies. To do so, I assume that the additional revenues by taxing skilled workers must be equal to the revenues lost by reducing the tax rate for unskilled workers. Under this condition, the progressivity of the income tax can be amended without any direct change to lump-sum transfers or training subsidy rates.

Figure 3.3 presents the IRFs to an increase of 5% in the tax progressivity, defined as ratio between skilled and unskilled tax rate $\tau^{h,s}/\tau^{h,u}$. The increase is similar in size to the change in tax rates that is necessary to compensate the training subsidy expenditures in the sections above.

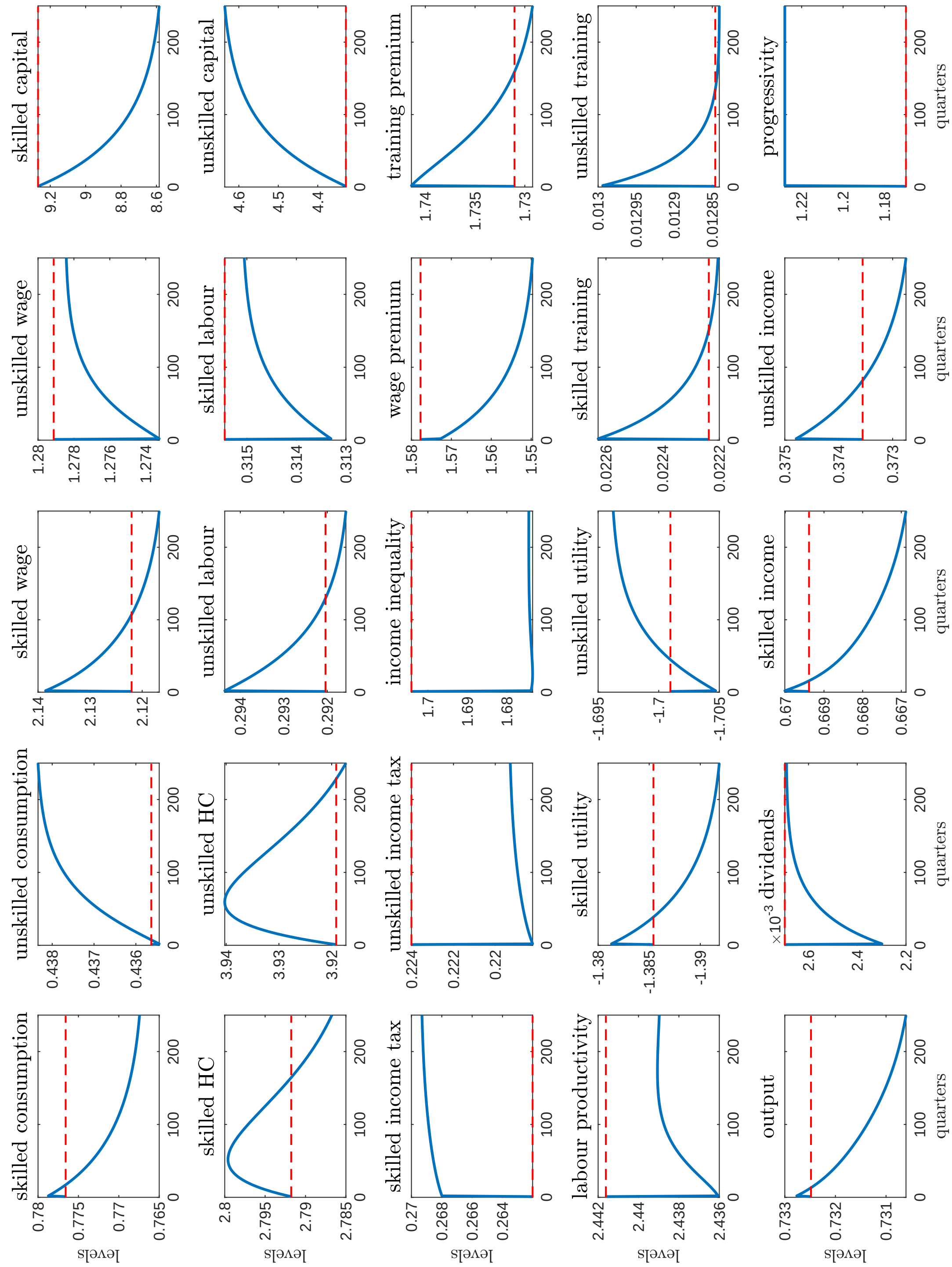


Figure 3.3: Permanent increase of tax progressivity by 5%

From Figure 3.3, it can be observed that, despite the government raises the same revenues, the increase in progressivity reduces output in both the short and the long run. The fall in labour productivity (sub-plot (4,1)) causes both unskilled and skilled wages to decrease below the original steady state, even though the new steady state labour supply is lower.[32]

Higher taxation of skilled workers' income entails a fall in the physical capital stock as well as in the skill capital stocks (in the long run only). Higher (lower) tax rates on skilled (unskilled) income entails a downward (upward) shift of the labour supply, which means higher (lower) costs that the firms has to pay for that specific production factor.

The higher (lower) cost to use skilled (unskilled) labour for production entails that there is a relocation towards the use of more unskilled labour input and unskilled human capital stock. As a consequence, the long-run decline of the latter is smaller than the decline observed for the skilled human capital stock. This relocation translates into the observed fall of labour productivity since unskilled workers are, by definition, less productive than skilled ones.

In terms of equality, the policy has positive effects as it produces a sizeable

[32]Remember that sub-plots (1,3) and (1,4) are before-tax wage rates. Hence, the peak of skilled wage only proves that the tax burden is shared between employers and employees. In particular, the former need to increase the bidding wage to incentive skilled workers to provide labour (sub-plot(2,4)).

reduction in both wage and labour income inequality. The after-tax incomes of skilled and unskilled households are equalised after this intervention, as shown by the inequality measures reported. Evaluated at the steady state, the after tax reduction in inequality is 5.8%, with respect to wage, and 3.8% with respect to labour income.

Conversely, the policy does not produce significant changes in the training inequality. In the short run, the policy incentives higher skilled and unskilled training as firms can use it as a buffer (human capital is the only production factor not taxed). Since skilled labour is more expensive than unskilled one, skilled human capital is most valuable. Thus, initially, the ratio of skilled training to unskilled training increases (sub-plot (3,5)). Later, at the steady-state, the training premium is similar to its initial value.

To conclude, it appears that there is a relevant efficiency cost (from resource mis-allocation) when the government exploits the progressivity of income tax rates to reduce between-groups income inequality. Thus, the government may prefer policies based on training subsidies to avoid this trade-off. After these preliminary considerations, I look more in detail at the effects of the fiscal reforms under scrutiny. The quantitative analysis of these policies is conduced in the next section.

3.5.2 Quantitative evaluation of the policies

This section analyses in more detail the policies considered above (skilled or unskilled training subsidy, either financed by both households or only the skilled one).[33] In both cases, revenues are increased so that the government budget is balanced after the increase in training subsidies.

To evaluate and compare the fiscal interventions, I report the percentage change from the initial steady state (common to all the experiments) to the new steady state to which the economy converges after the fiscal reform is implemented. This provides a first rough measure of the overall effect of each policy on the economy. Moreover, to take into account the dynamics, i.e. the transition from one steady state to the other, I compute fiscal multipliers following the literature on fiscal policy evaluation (see e.g. Leeper *et al.* (2010)). Accordingly, the lifetime multiplier of e.g. the variable x is defined as:

$$\mathrm{LM}_x = \frac{\sum_{t=0}^{S} Q_t^i \left(x_t - x\right)}{\sum_{t=0}^{S} Q_t^i \left(T_t - T\right)} \tag{3.42}$$

where Q_t^i, for $i = u, s$, is the household discount factor as implicitly defined by the

[33] In Appendix G, I consider the case where the policies are financed by lump-sum transfers. This allows to more closely compare the results here reported with those from Chapter 2.

respective physical capital Euler equation, i.e. equation (3.17) for unskilled and (3.9) for skilled workers. When the income measure of interest is an aggregate, e.g. aggregate output y_t, the average of the two discount factors, weighted by the population share, is used.[34]

[34] Note that using a constant discount factor, i.e. the β, provides virtually the same results as the ones reported here.

3.5.2.1 Subsidies to unskilled training

Steady state comparative statics

I start with the case of training subsidies in favour of unskilled workers. Table 3.5 reports the percent changes from the old to the new steady state of key variables such as training, inequality measures, and utility among the others. These quantities are shown for different levels of the subsidies to the monetary training costs for each scenario. With respect to utility, since its steady state value is negative, I compute the percent change as the difference between new and old steady-state divided by the absolute value of the latter. Following this definition, a positive change implies welfare gain (as the utility gets closer to zero) whereas a negative change implies welfare losses (as utility gets further away from zero).

Several observations are in order. In first instance, the effect of the policy on training time is strong, subsidising 15% of monetary training costs implies an 11% increase in the share of unskilled labour input dedicated to training activities (see cell(1,3) in Table 3.5).

This is accompanied by a large decrease in training inequality which is also substantial (see cell (3,3) in Table 3.5) and does not change significantly if I change the source of financing. These results suggest that, at least for low level of subsidies, the effect on training is similar between the two scenarios.[35]

Table 3.5: Steady-state effects of increasing unskilled training subsidies

	both tax rates			skilled-only		
$\tau^u =$	0.05	0.10	0.15	0.05	0.10	0.15
$\%\Delta t^u$	2.14	6.51	11.27	2.14	6.51	11.25
$\%\Delta t^s$	0.01	0.03	0.04	0.00	-0.01	-0.03
$\%\Delta \frac{t^s}{t^u}$	-2.08	-6.09	-10.09	-2.09	-6.12	-10.14
$\%\Delta w^u \left(1-\tau^{h,u}\right)$	0.04	0.12	0.20	0.08	0.24	0.40
$\%\Delta w^s \left(1-\tau^{h,s}\right)$	-0.03	-0.11	-0.20	-0.11	-0.36	-0.65
$\%\Delta \frac{w^s(1-\tau^{h,s})}{w^u(1-\tau^{h,u})}$	-0.08	-0.23	-0.39	-0.19	-0.59	-1.04
$\%\Delta w^u l^u \left(1-\tau^{h,u}\right)$	0.05	0.16	0.25	0.08	0.23	0.38
$\%\Delta w^s l^s \left(1-\tau^{h,s}\right)$	-0.04	-0.13	-0.25	-0.13	-0.40	-0.73
$\%\Delta \frac{(1-\tau^{h,s})l^s w^s}{(1-\tau^{h,u})l^u w^u}$	-0.10	-0.29	-0.50	-0.20	-0.63	-1.10
$\%\Delta \frac{y}{(l^s n^s + l^u n^u)}$	0.01	0.02	0.03	-0.11	-0.35	-0.63
$\%\Delta U^u/\lvert U^u\rvert$	0.01	0.03	0.05	0.03	0.09	0.16
$\%\Delta U^s/\lvert U^s\rvert$	-0.01	-0.03	-0.06	-0.04	-0.13	-0.24
$\%\Delta(t^s l^s \phi^s \tau^s + t^u l^u \phi^u \tau^u)$	0.51	1.54	2.67	0.61	1.89	3.30

In terms of after-tax wages and incomes, Table 3.5 shows that unskilled households benefit from the policy both when the tax burden is on skilled workers and when this is shared – for reasonably low subsidy rates. Conversely, skilled households are increasingly damaged by the policy as the subsidy rate gets larger.

Chapter 2 shows the presence of significant spillovers from unskilled workers to skilled workers when looking at labour income outcomes. In this case, I do not

[35]Note that if the subsidy to unskilled training were very large, then the type of taxation will have a large impact on the effects of the policy on unskilled training. In particular, distortions on skilled labour market due to a very high skilled-only tax hike could reduce the incentives to train unskilled workers. Such results are not presented here to focus on more ”reasonable” subsidy rates.

observe any manifest spillover effect.[36]

In terms of financial cost, it can be observed that the policies have a relatively modest impact on public finances. A permanent increase in training subsidies leads to an increase in public spending for training between 0.5 and 3.3% depending on the size of the intervention and the source of financing (see the last row of Table 3.5). This corresponds to a GDP share between 0.01% and 0.07%.

Taxing skilled workers only allows the government to further reduce wage inequality, but this comes at the cost of lower efficiency. In fact, the change in labour productivity (third row from the bottom, in Table 3.5) is shown to become negative. After commenting on labour productivity, I look at welfare changes. Unskilled utility improves in both scenarios, and, as it can be expected, the most positive outcome is when the financial burden is on skilled workers only. Skilled utility instead decreases under all interventions and financing options. Skilled utility loss is larger the larger is the subsidy size, and largest under skilled-only tax hike.

[36] To be more correct, spillovers would be seen under very small subsidy rate increases (e.g. 0.035 – a case that is not reported). For larger rates, like the one reported in Table 3.5, the negative effects of distortionary taxation kicks in, affecting the skilled household's market outcomes.

So far, I considered the change from the old steady state to the new one after the policy reform, thus neglecting the role played by the transition phase towards the new equilibrium. To better evaluate the impact of the unskilled training subsidies, Table 3.6 reports the present value of income multipliers as well as the present value of the change in utility $\sum_{t=0}^{\infty} Q_t^i \left(U_t^i - \bar{U}^i\right)$.[37]

Table 3.6: Lifetime multipliers after increasing unskilled training subsidies

	both tax rates			skilled-only		
$\tau_u =$	0.05	0.10	0.15	0.05	0.10	0.15
$w^u l^u \left(1 - \tau^{h,u}\right)$	0.75	0.69	0.63	1.35	1.30	1.24
$w^s l^s \left(1 - \tau^{h,s}\right)$	-1.79	-1.87	-1.96	-4.47	-4.63	-4.81
$w^u l^u (1 - \tau^{h,u}) + w^s l^s (1 - \tau^{h,s})$	-0.93	-1.00	-1.08	-2.49	-2.62	-2.76
$(1 + r\left(1 - \tau^{h,s}\right) - \delta_k) k^s$	-51	-53	-54	-322	-330	-339
$(1 + r\left(1 - \tau^{h,u}\right) - \delta_k) k^u$	4.16	3.92	3.67	127	129	132
π	-0.66	-0.66	-0.67	-5.85	-6.00	-6.16
y	-0.13	-0.23	-0.34	-6.19	-6.47	-6.78
$U^u - \bar{U}^u$	0.09	0.26	0.41	0.26	0.78	1.33
$U^s - \bar{U}^s$	-0.12	-0.39	-0.71	-0.35	-1.10	-1.98

As Table 3.6 shows, the results are more nuanced when looking at the effects of policies in terms of present value (taking into account the transition phase towards the new steady state). From this perspective, it is also possible to evaluate wealth redistribution. In particular, if the tax rate is increased only for skilled workers, the model predicts a large shift of capital income from skilled to unskilled workers

[37] This is equivalent to taking only the numerator instead the full lifetime multiplier as defined in equation (3.42). I show this value because it can be more easily interpreted.

as the reduction in labour inequality leads to a reduction in wealth inequality.[38]

An important quantity is the aggregate output multiplier. According to the simulation, spending 1£ on unskilled training decrease the present value of aggregate output by 0.23£ when increasing training subsidies up to 10% of monetary costs (see cell (7,2) in Table 3.6). Thus, when I take into account the distortion caused by higher income tax rates, the policy does not have an efficiency *rationale* but it can be justified in terms of utility gains and because it alleviates inequality.[39]

Welfare evaluation

The welfare evaluation suggest that unskilled subsidies are beneficial to unskilled households, but detrimental to skilled households. For low level of subsidies, I expect that the average welfare is increased by the reform because skilled households are only about 34% of the total population. Yet, high level of subsidies have increasingly larger skilled utility losses and those are unlikely to be compensated by the gains of the unskilled households. It goes without saying that households have contrasting views about the size of the optimal subsidy: unskilled workers benefit more and desire higher subsidies than skilled workers. The latter are always

[38] I do not discuss further this aspect as the focus of the present work is on wage and labour income.

[39] As already mentioned, the case of non-distortionary taxation is considered in Appendix G, and it represents a main assumption in Chapter 2, where I focus on the "pure" effects of training subsidy reforms.

penalised for any of the fiscal reform here contemplated.

To identify the optimal level of unskilled subsidies, I simulate the model for a range of unskilled training subsidies, τ^u, and I plot the change in present value of utility for skilled, unskilled, and average worker in Figure 3.4. They are reported with blue dashed, red continuous, and yellow dotted lines, respectively. The figure is composed of three sub-plots. The first one considers the case when the policy is financed through lump-sum taxes. This case, unsurprisingly, is characterised by larger utility gains. Sub-plot (b) shows the utility change when the policy is financed through higher taxes for both skilled and unskilled workers. Lastly, sub-plot (c) presents the case where the fiscal burden only on skilled workers.

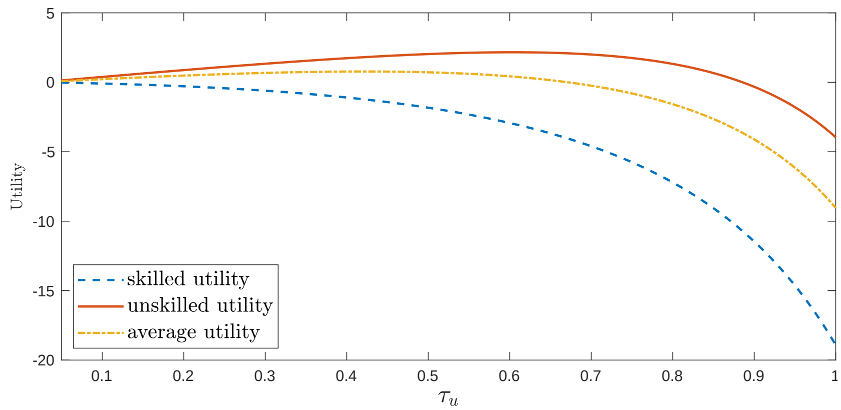
Sub-plot (b) of Figure 3.4 suggests that the optimal subsidy for unskilled workers is $\tau^u = 0.37$, whereas skilled workers prefer the *status quo* if both households bear the costs. The policy that maximises average utility under this scenario is $\tau^u = 0.11$, but this intervention will be opposed by skilled households.

If I assume that the costs are paid by the skilled household, there is no optimal subsidy level for this group. Unskilled workers benefit more and inequality is reduced. However, for skilled workers, the costs always outweigh the benefits of the intervention. Thus, unskilled workers desire training subsidies as high as possible under this scenario.

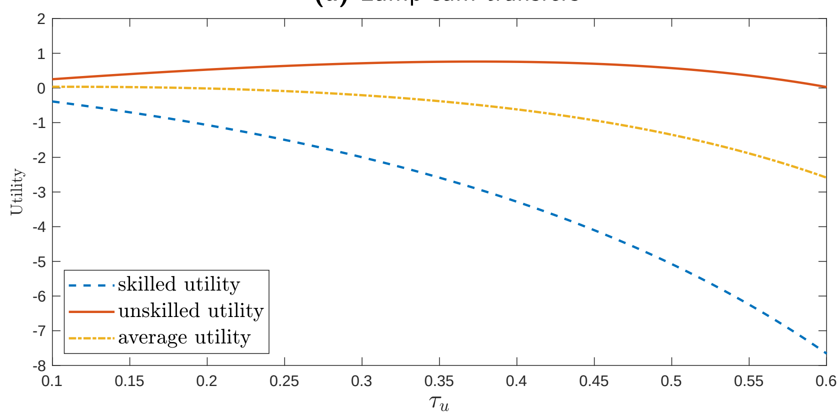
According to the model, if the government implements a subsidy to unskilled training equal to 15% of its monetary costs, financed through higher taxes on both households, the overall reduction in after-tax wage inequality is about 0.4% whereas, if the government finances an unskilled training subsidy of 15% by taxing only skilled workers, the reduction in after-tax wage inequality is about 1%. Thus, inequality reduction is more than two times larger, but distortions have very negative effects on average welfare and aggregate productivity.

In conclusion, financing unskilled training finds the opposition of the skilled group. However, I have also verified that if the fiscal burden were on unskilled workers only, they would not be not willing to pay for their subsidies.[40] In the case of shared burden, the social conflict lies in the level of subsidisation. Unskilled workers desire higher subsidies than skilled workers (who prefer the *status quo*), and only a subsidy rate in-between these two values will not meet the opposition of skilled workers.

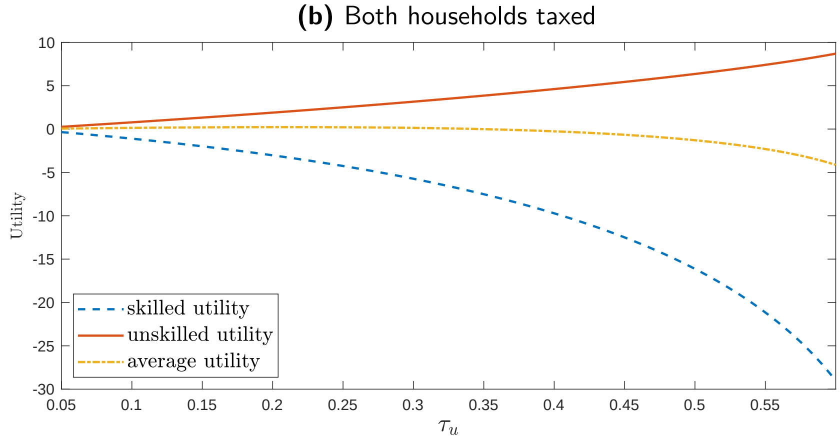
[40] THe welfare evaluation for this additional financing option is reported further below, in Table 3.9.



(a) Lump-sum transfers



(b) Both households taxed



(c) Skilled only taxed

Figure 3.4: Change in present value of utility by unskilled subsidy

3.5.2.2 Subsidies to skilled training

Steady state comparative statics

I now move on to consider the effects of training subsidies in favour of skilled workers. Table 3.7 reports the percent change from the old to the new steady state of key variables such as training, inequality measures, and utility. These values are shown for different levels of subsidies of the monetary training costs under both the financing options. As in the previous section, the percent change in utility is computed taking the absolute value of the steady-state utility (as denominator) to have a more intuitive figure.

The effect of subsidies to skilled training is quantitatively similar to that of unskilled training subsidies. In this case, training inequality rises and it exacerbates inequality of outcomes among workers. Yet, the effects are dampened by the strong spillovers in favour of unskilled workers.[41] In fact, unskilled wages are higher than the old steady state for τ^s values up to at least 0.10.

The benefits of skilled training subsidy are represented by higher productivity (see row 10 in Table 3.8) and higher after-tax wage rates for both households (see rows 4 and 5 in the same table). These benefits do not materialise if the policy is financed by increasing only the skilled tax rate (last three columns of Table 3.8).

[41] Interestingly, there are stronger spillovers from skilled policies to unskilled workers than from unskilled policies to skilled workers.

In this case, though, the spillover effects are magnified.

With respect to the costs, this policy is more expensive than the policy based on unskilled subsidies as evident by comparing the last row of Table 3.8 with the last row of table 3.6. This is true for both scenarios and for any subsidy size. This may seem unexpected since the monetary costs of unskilled training are higher than those of skilled training ($\phi_u > \phi_s$). However, it must be noted that in the case of skilled subsidies, extra resources are spent not only for skilled training but also for unskilled training (whose steady state is higher than before the reform).

Table 3.7: Steady-state effects of increasing skilled training subsidies

	both tax rates			skilled-only		
$\tau^s =$	0.05	0.10	0.15	0.05	0.10	0.15
$\%\Delta t^u$	0.02	0.06	0.10	0.02	0.05	0.09
$\%\Delta t^s$	1.75	5.29	9.07	1.74	5.26	9.03
$\%\Delta \frac{t^s}{t^u}$	1.73	5.22	8.97	1.72	5.20	8.93
$\%\Delta w^u \left(1-\tau^{h,u}\right)$	0.00	0.00	-0.02	0.02	0.07	0.12
$\%\Delta w^s \left(1-\tau^{h,s}\right)$	0.03	0.09	0.14	-0.02	-0.08	-0.16
$\%\Delta \frac{w^s(1-\tau^{h,s})}{w^u(1-\tau^{h,u})}$	0.03	0.09	0.15	-0.05	-0.15	-0.27
$\%\Delta w^u l^u \left(1-\tau^{h,u}\right)$	0.00	-0.01	-0.03	0.01	0.04	0.05
$\%\Delta w^s l^s \left(1-\tau^{h,s}\right)$	0.04	0.12	0.18	-0.02	-0.06	-0.13
$\%\Delta \frac{(1-\tau^{h,s})l^s w^s}{(1-\tau^{h,u})l^u w^u}$	0.04	0.13	0.22	-0.03	-0.10	-0.19
$\%\Delta \frac{y}{(l^s n^s + l^u n^u)}$	0.03	0.10	0.15	-0.05	-0.15	-0.28
$\%\Delta U^u / \lvert U^u \rvert$	0.00	0.00	0.00	0.01	0.04	0.07
$\%\Delta U^s / \lvert U^s \rvert$	0.01	0.02	0.04	-0.01	-0.04	-0.08
$\%\Delta(t^s l^s \phi^s \tau^s + t^u l^u \phi^u \tau^u)$	1.31	3.95	6.78	1.38	4.19	7.21

In terms of after-tax wage, Table 3.7 shows that both households benefit from the policy even when the fiscal burden is shared, up to the 10% subsidy rate (see rows 4 and 5). This is a stronger result compared to the case of unskilled training subsidies. Even in this case though, the effect is stronger for moderate increases of τ^s and it weakens when the subsidy gets larger due to the adverse effect of distortionary taxation. The spillovers from skilled workers to unskilled workers are stronger than those from the opposite direction,[42] but the policy-maker faces the unpleasant drawback of exacerbating wage and labour inequality.

The skilled training subsidies favour the accumulation of skilled human capital which is marginally more productive than unskilled one. This has direct and indirect benefits for both households. As a consequence, most of the values in Table 3.7 are positive, at least when the burden is shared. With respect to the skilled-only tax burden, the benefits of the fiscal reform are outweighed by the negative effects of higher tax rate on skilled workers income. The pattern is similar to what observed for unskilled training subsidies for the same scenario. In this case, the negative effects are so strong that even a 5% subsidy rate depresses labour productivity.

With respect to welfare, unskilled workers have utility gains in both scenarios. Sharing the burden seems to lead to lower unskilled utility benefits. The greater the increase in skilled training subsidies, the greater are the benefits accrued by

[42]This can be concluded by comparing size of the unskilled utility losses after the change in skilled training subsidies versus the size of skilled utility losses after the change in unskilled training subsidies (in the shared tax burden case).

the skilled household as measured by steady-state changes in lifetime utility when the financing is shared.

Transition comparative dynamics

Once controlling for the transition path, results are more nuanced. Table 3.8 reports the present value of income multipliers and the present value of the change in utility $\sum_{t=0}^{\infty} Q_t^i \left(U_t^i - \bar{U}^i\right)$ as computed in the previous section for the τ^u policy.

Table 3.8: Lifetime multipliers after increasing skilled training subsidies

	both tax rates			skilled-only		
$\tau_s =$	0.05	0.10	0.15	0.05	0.10	0.15
$w^u l^u \left(1-\tau^{h,u}\right)$	-0.08	-0.09	-0.10	0.05	0.04	0.04
$w^s l^s \left(1-\tau^{h,s}\right)$	0.25	0.22	0.20	-0.34	-0.38	-0.43
$w^u l^u(1-\tau^{h,u}) + w^s l^s(1-\tau^{h,s})$	0.14	0.12	0.10	-0.21	-0.24	-0.27
$(1+r\left(1-\tau^{h,s}\right)-\delta_k)k^s$	-2.77	-3.35	-3.93	-62	-64	-66
$(1+r\left(1-\tau^{h,u}\right)-\delta_k)k^u$	4.30	4.16	4.02	31	32	32
π	-0.12	-0.13	-0.13	-1.27	-1.30	-1.33
y	0.41	0.37	0.34	-0.93	-0.99	-1.07
$U^u - \bar{U}^u$	-0.01	-0.06	-0.12	0.11	0.32	0.53
$U^s - \bar{U}^s$	0.03	0.08	0.11	-0.13	-0.42	-0.78

The multipliers of this exercise look similar to the ones in Table 3.8, however there are important differences.

First, multipliers for the skilled training subsidy are smaller in absolute values than those associated with unskilled training subsidies. Hence, the group favoured by the policy gain less in terms of e.g. labour income, but at the same time the other group loses less. Moreover, the average labour income multiplier increases, whereas it decreases with unskilled training subsidies.[43]

Capital income multipliers are different as well. In Table 3.8, columns 1 to 3, skilled capital multipliers are negative whereas unskilled capital multipliers are positive, despite the policy works in favour of skilled workers. This is due to the fact that this policy favours physical capital accumulation and drives capital income up through its complementarity with skilled labour input (thus the monetary spillover in favour to unskilled workers).

Table 3.8 also shows that skilled training policies have positive output multipliers when the tax burden is shared. When only skilled households bear the cost, the policies have less negative output multiplier than unskilled training subsidies. Thus, the income tax related distortions are much larger if the policy is financed by increasing only the skilled income tax rate as if it exacerbates the efficiency costs linked to promoting inequality. The second scenario can thus be considered as the combination of skilled training subsidies and an increase in tax progressivity (studied alone in Section 3.5.1.3).[44] The loss of efficiency (less output and lower productivity) can be seen as the negative consequence of favouring unskilled

[43]This result holds also under skilled-only taxation, despite aggregate multipliers being negative. It's still true that unskilled training subsidies have stronger negative effects under skilled-only taxation compared to the skilled training subsidies (compare e.g. row 7 of Table 3.8 and 3.6).

[44]Skilled workers receive higher subsidies to train while facing higher income tax rates.

labour input over skilled one.

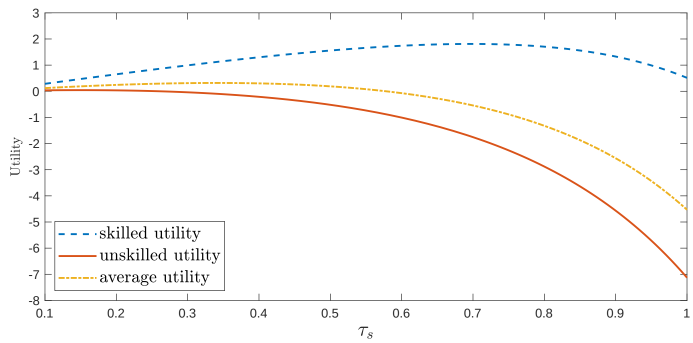
Welfare evaluation

I now discuss about welfare. To identify the optimal level of subsidies, taking $\tau^s = 0.024$ as starting point, I simulate the model economy for different levels of (higher) skilled training subsidies and I plot in Figure 3.5 the change present value of utility for skilled and unskilled workers, and for the average worker. Those are represented by the blue, red, and yellow lines respectively.

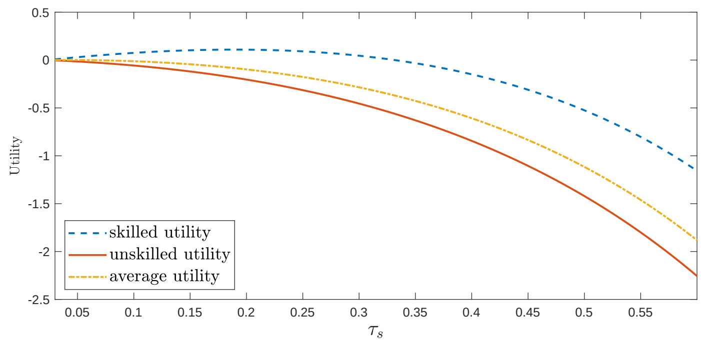
From Figure 3.5, sub-plot (b), it can be inferred that the optimal subsidy for skilled workers is $\tau^s = 0.19$, whereas unskilled workers prefer the *status quo* if both households bear the financial costs of the intervention. The policy that maximises average utility under this scenario is $\tau^s = 0.04$, and this project would be backed by the skilled household only, as unskilled households pay part of the costs but the policy's spillovers are not large enough to compensate them.[45]

If the costs are paid by the skilled household, i.e. Figure 3.5, sub-plot (c), the optimal subsidy levels are different. The optimal subsidy is $\tau^s = 0.60$ from the perspective of unskilled workers, and it is $\tau^s = 0.03$ from the perspective of skilled workers. Not surprisingly, in this case skilled workers desire a much smaller

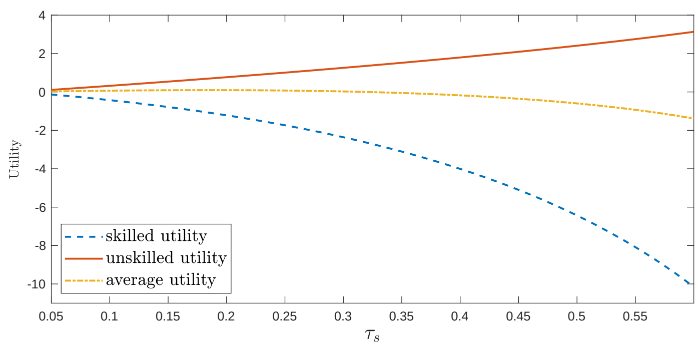
[45] Simulating the model with larger values of the externality, i.e. δ^s_ε, suggests that both households would prefer the reform to the *status quo*, confirming the importance of a robust calibration for this parameter.



(a) Lump-sum transfers



(b) Both households taxed



(c) Skilled only taxed

Figure 3.5: Change in present value of utility by skilled subsidy

Table 3.9: Desired tax rate by worker type

	τ_u policy			τ_s policy		
	unskilled	skilled	average	unskilled	skilled	average
Lump-sum	0.61	-	0.42	0.15	0.70	0.34
Both taxed	0.37	-	0.11	-	0.19	0.04
Skilled only	0.63	-	0.21	0.60	0.03	0.19
Unskilled only	-	0.81	-	-	1.00	-

subsidy, since they have to fully pay for it through higher total income taxes. The unskilled workers' support for large subsidies to skilled training is due to the spillover effects that higher skilled workers productivity entails and the indirect benefit of higher taxes on skilled income.

To summarise the results, Table 3.9 reports the optimal rates for both the unskilled training subsidies (first three columns) and the skilled training subsidies (last three columns). Each column indicates which worker's utility is maximised under a given policy. Hence, *ceteris paribus*, each policy rate is "optimal" in that it maximises the welfare gains of either: (i) the unskilled worker, (ii) the skilled worker, or (iii) the average worker. The rows indicate which source of revenues is used to finance the subsidy. I report the lump-sum case as a reference, and the three cases where skilled, unskilled, or both households pay higher taxes to finance the new training policy. For example, in the cell (2,1), $\tau^u = 0.37$ is the unskilled subsidy, financed by both households, that maximises the welfare of unskilled workers. A missing value means that the worker is worse off for any

Table 3.10: Change in inequality under the desired tax rate

	τ_u policy			τ_s policy		
	unskilled	skilled	average	unskilled	skilled	average
Lump sum	-1.70	-	-0.97	0.24	1.90	0.69
Both taxed	-1.14	-	-0.22	-	0.25	0.02
Skilled only	-4.55	-	-0.68	0.39	0.01	0.17
Unskilled only	-	-3.39	-	-	3.20	-

increment of the subsidy rate.[46]

The general conclusions that can be drawn from Table 3.9 are the following. First, the subsidy rates, optimal for the average worker, are relatively low (meaning lower than 22%). The beneficiary group desires much higher subsidy rates (up to 60%) but the other group often desire no subsidy at all. Lastly, due to population size, the unskilled welfare has a large impact on the average worker's.

Table 3.10 shows the effects of every policy indicated in Table 3.9 on the steady-state wage inequality (in percentage points). For example, looking at the cell (3,1) of each table, it can be concluded that the policy $\tau_u = 0.63$ financed by a higher tax rate on skilled income lead to a reduction of 4.55% in the before-tax wage inequality, i.e. w_s/w_u.

In conclusion, skilled training subsidies improve the welfare of both unskilled and skilled households, but they exacerbate inequality and may not be seen favourably by a policy-maker that is concerned about the latter. On the other side,

[46]This analysis assumes that the other subsidy rate is unchanged. I consider what happens when both rates are reformed in the next section.

unskilled subsidies have, in average, positive effects but they are more polarising as they are liked by either skilled or unskilled workers, but not by both. Finally, only unskilled subsidies can mitigate wage inequality and the achieved reduction is relatively small.[47]

In the next section, I consider what happens when both τ_s and τ_u are optimised over a specific target.

3.5.3 Optimal policy

In this context, the policy evaluation can be performed in different ways, depending on the preferences and the goals of the policy-maker. I consider three main policy targets: (i) the average utility of workers; (ii) the steady state level of output per worker, as a proxy for labour productivity; and (iii) the present value of future aggregate output, as a proxy for the overall size of the economy. Since unskilled workers represent the larger population share, the average present value of utility favours more unskilled than skilled workers, yet the weights could be changed in consideration of equity issues.

I focus on the case where the fiscal subsidy is financed by both households, omitting the analysis of optimal policy under other scenarios as it would be either

[47] I have already discussed why wage (and income) inequality cannot be addressed through unskilled training subsidies in Section 2.6 of Chapter 2.

trivial or uninteresting.[48]

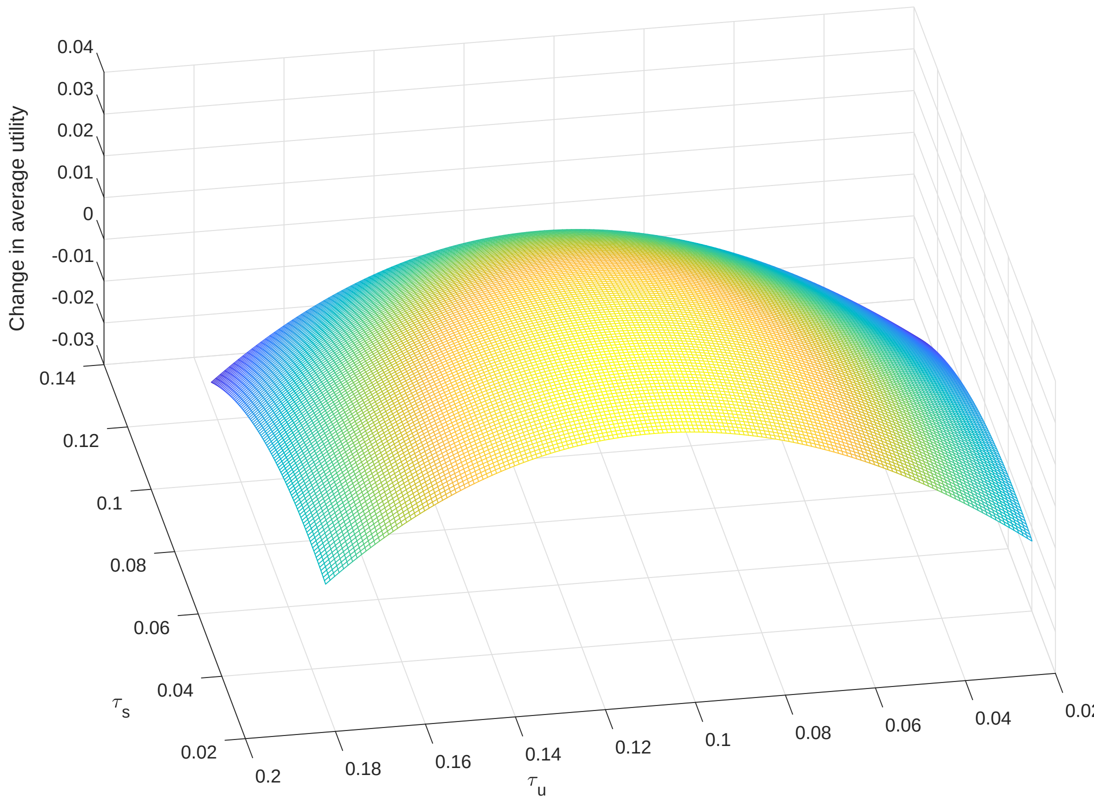


Figure 3.6: Effects of training subsidies on aggregate utility

Since *a priori*, it is not possible to prove uniqueness of the optimal policy, especially if the latter is defined according to the criterion (ii), I consider a closed-subset of fiscal policies. In particular, I numerically evaluate the effects of combination of policies for $\tau_u \in [0.0241, 0.18]$ and $\tau_s \in [0.0241, 0.12]$. Given the results

[48]In the previous section, I concluded that fiscal reforms that are subsidised by one single household are never Pareto-improving, thus the payers do not approve the change in policy.

of the previous section, this set of combinations includes any reasonable level of tax subsidies.

If the policy-maker wishes to maximise the average present value of utility, the optimal policy would be to set $\tau_u = 0.108$ and $\tau_s = 0.037$. This corresponds to an increase in the average present value of utility by 0.038%. According to the results, skilled training subsidies should be incremented marginally, whereas the government should be much more active in favour of unskilled training subsidies. Figure 3.6 shows the welfare gain associated to different levels of skilled and unskilled training subsidies. There is a clear decrease in welfare gain that is related to how far subsidy rates are from their optimal level. So, it is unlikely that combinations of subsidy rates outside the current range may produce higher welfare.

Compared to the partial analysis conducted before, here the mutual benefits imply that the increase in welfare is larger than the sum of welfare gains due to single policies (i.e. subsidising skilled training keeping constant unskilled training subsidies and *vice versa*). In terms of optimal policy rates, these are the same as the ones reported in Table 3.9 for the case of a single fiscal subsidy tool. Thus, it appears that, at least for the welfare-based measure, the optimal training subsidy for one group is independent from the level of subsidies set for the other group.

With respect to the increase in labour productivity, I consider changes from the

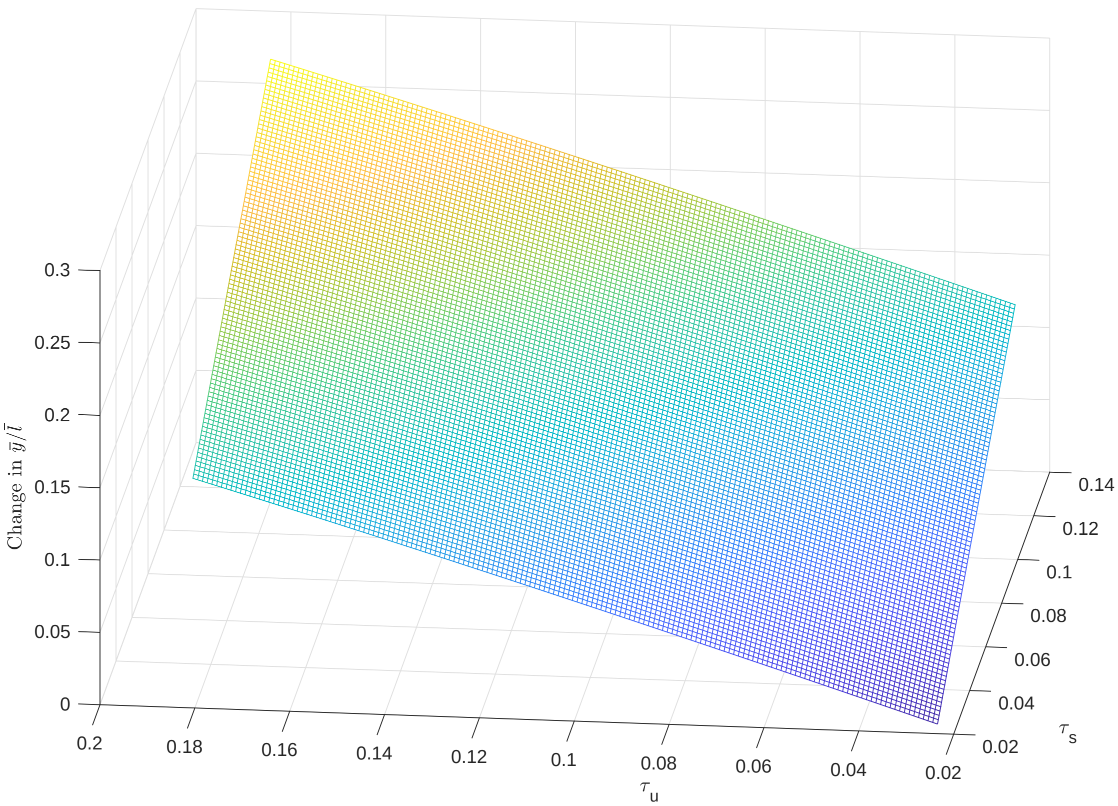


Figure 3.7: Effects of training subsidies on productivity

original steady state to the new one where the policy has been implemented to observe the long run gains from subsidising training activities. Figure 3.7 shows that there is no maximum increase in steady-state labour productivity $\bar{y}/\bar{l}$, at least for the set of subsidy rates taken into account. According to this metric, large training subsidies are desirable and can support productivity and wages.[49]

Despite the welfare evaluation favours unskilled training over skilled training

[49]Incidentally, the low performance of UK economy in terms of labour productivity has been reported and largely discussed in the literature (see e.g. Barnett *et al.* (2014) and Blundell *et al.* (2014)).

subsidies, here the conclusion is opposite. In fact, extrapolating iso-productivity lines from Figure 3.7, it can be noted that an increase in skilled training subsidies has a larger impact on labour productivity compared to an equally large increase in unskilled training subsidies.[50]

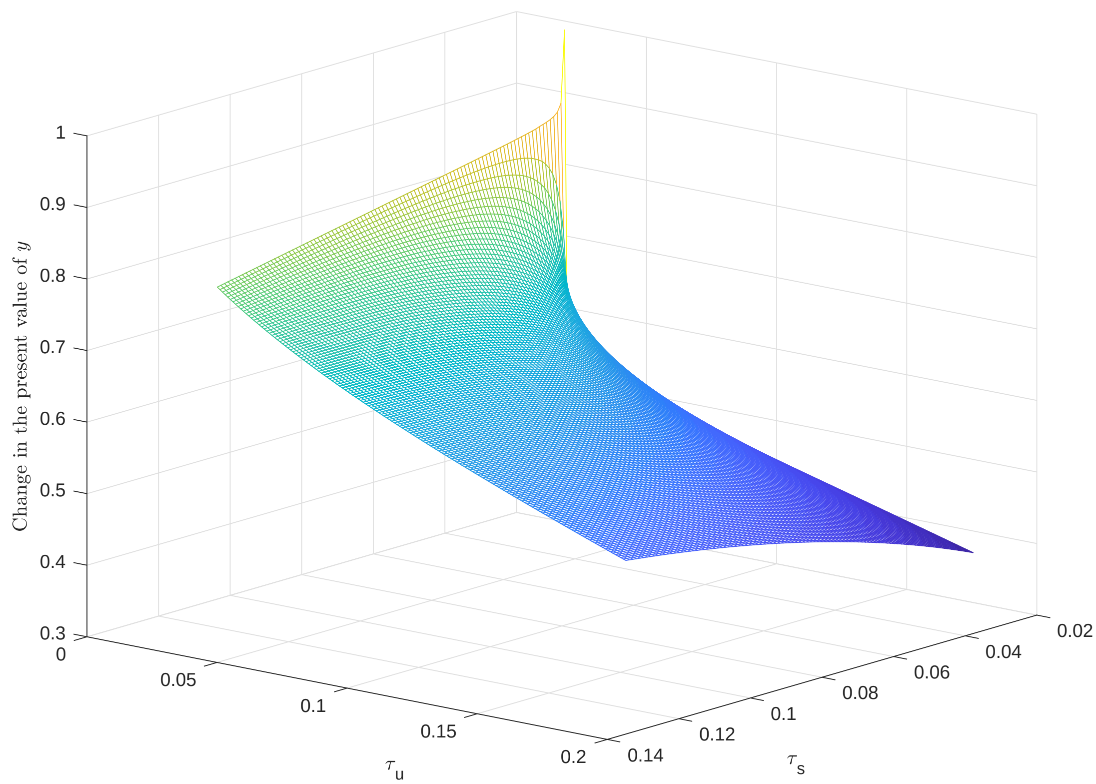


Figure 3.8: Effects of training subsidies on aggregate output

Lastly, maximising the present value of aggregate output with respect to the

[50]By this time, such a result should not be surprising. The crowding-out of the (more productive) skilled input caused by unskilled training subsidies has been already discussed above.

training subsidy rates leads to totally different conclusions. The effect of each policy combination is reported in Figure 3.8. It can be noted that the higher percentage increase in discounted output is associated with the smallest increase in training subsidies rates, i.e. $\tau_u = 0.0241$ and $\tau_s = 0.0241$. Also, the reduction in the benefits of the policy reform is stronger for unskilled training subsidies than for skilled ones (see the different slope of the two edges in the graph). The unskilled subsidies have a weaker positive effect on aggregate output than skilled subsidies. This is to be expected, given that in the production function the skilled input is substitute to unskilled input. The subsidy hinders productivity as it incentivises the unskilled input at the expenses of the other inputs.

3.6 Conclusions

This chapter, building on the experience and knowledge gained from Chapter 2, performs a thorough policy evaluation of the effects of fiscal subsidies on firm-provided training activities for the UK economy.

I find evidence suggesting that firms provide less training opportunities than optimal because competitors can free-ride on the skills of workers they invested in (poaching externalities). In particular, I estimate that a 1% increase in the job-to-job separation rate leads to a fall in the training participation rate by -0.038%

and -0.037% for skilled and unskilled workers, respectively. The fact that these values are small can explain why the literature has not agreed on the existence of training under-provision.

The evidence suggests that some (moderate) government intervention may be desirable. According to the data of the Continuing Vocational Training Survey, the UK government supports training activities paying 2.4% of their monetary costs with resources from general taxation. To find out which level of training subsidisation is desirable, I simulate a model that captures the salient features of UK economy, investigating the effects of policy reforms.

First, I consider the effects of a partial intervention, which includes either unskilled or skilled training subsidies. This allows to evaluate the usefulness of unskilled training subsidies to reduce wage inequality and to evaluate the impact of each policy separately. The results are consistent with the conclusions of Chapter 2. The lever offered by training subsidies to reduce wage and income inequality is limited but not negligible. A trade-off arises, between the pursuit of lower inequality between skilled and unskilled workers and the larger efficiency gains produced by subsidising skilled training.

All in all, I find that an increase of the subsidies to job-related training is beneficial to the UK workforce. To maximise the present value of the average worker's welfare, the policy-maker should subsidise 10.8% and 3.7% of the monetary training costs of unskilled and skilled workers, respectively. This policy increases welfare by about 3.4% and reduces wage inequality by 0.24%. Despite the gains for the average worker, such a policy damages skilled workers who disapprove this level of subsidisation.[51] Thus, the valuation of these policies cannot be separated by a judgement of social justice (Rawls (1971)).[52]

The recent trends observed in the UK labour market are likely to affect this evaluation. In particular, the increased supply of skilled workers is expected to increase the desirable amount of training subsidies in their favour and to reduce their opposition to subsidies in favour of unskilled workers. In any case, these findings consolidate the role of training subsidies as part of a larger set of government policies concerning education and labour market participation.

[51] A Pareto-improving policy, that makes both groups better off, would be subsidies corresponding to 3.7% and 2.7% of the monetary training costs for skilled and unskilled workers, respectively.

[52] This is true even when the policy-maker subsidises both unskilled and skilled training at the same rate.

Conclusions

The accumulation of human capital is a key research subject in economics. Recently, post-education activities providing skills and competences have received greater and greater attention, both by the academia and by institutions. This Ph.D. dissertation focuses on firm-provided training and the pertaining fiscal policies.

In the first chapter, I bring forth evidence on the returns to training in terms of higher wages and profits, for workers and firms, respectively. Taking as reference the UK economy, I study the aggregate figures of training participation and the determinants of training at the individual level.

Since it emerges that job-related training in the UK does not have a cyclical component, I consider more in detail its trend in the last twenty years. In this regard, the training participation rate has decreased since the peak observed in the early 2000s. This fall has been only tentatively explained by the literature.

My data analysis excludes some of the proposed explanations (e.g. educational changes) and emphasises the importance of decomposing training participation into sub-groups. In particular, the previous literature has neglected important differences between University educated and non-University educated workers, and between workers close to the retirement age and all the other workers. Future work is needed to fully comprehend the driver of these trends.

Lastly, I observe a large and steady difference between training rates of University educated and non-University educated workers in favour to the former. Exploiting this workers classification, I contribute to the literature showing a positive relationship between training inequality and wage inequality. This last result brings up the research question for the second chapter.

The second chapter investigates whether subsides to job-related training could improve earnings for the lower skilled workers and reduce wage inequality. To this end, I use a general equilibrium (GE) model that incorporates skilled and unskilled labour, capital-skill complementarity in production and an endogenous training allocation.

The quantitative policy analysis suggests 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. Training subsidies to skilled workers, while increasing skilled and unskilled earnings, raise the former by more, worsening wage inequality. Therefore, there is a trade-off associated with subsidies to skilled training.

Training subsidies to unskilled workers improve earnings for both skilled and unskilled workers without a negative impact on inequality. However, the positive spillover effects to skilled workers imply that the effects of training subsidies on inequality are small. Hence, training subsidies are not a very effective in reducing inequality, but this is not a negative result. The effectiveness of the policy to propagate benefits throughout the labour force increases its social value.

To focus on the redistributive aspect of the fiscal policy, the second chapter assumes that the market provision of training is efficient. Under such assumption, I can also test the model's consistency with UK empirical evidence, and I can observe more easily the effect of training subsidies. Yet, to fully understand the consequences of these policies, it is important to take into account externalities and distortionary taxation.

The third chapter addresses these issues. I hypothesise that firms provide less training opportunities than optimal because competitors can free-ride on the skills of workers they invested in, the so-called poaching externality. My analysis suggests that a 1% increase in the job-to-job separation rate leads to a fall in the training participation rate by -0.038% and -0.037% for skilled and unskilled workers, respectively.

I amend the GE model to accommodate this evidence, and to allow for welfare evaluation. Performing similar policy exercises, preliminary results confirm what emerges from Chapter 2. The lever offered by training subsidies to reduce wage and income inequality is limited but not negligible. A trade-off arises, between the pursuit of lower inequality between skilled and unskilled workers and the larger efficiency gains produced by subsidising skilled training.

The analysis concludes that a moderate increase of the subsidies to job-related training is beneficial to the UK workforce. To maximise the average welfare, the policy-maker should subsidise 10.8% and 3.7% of the monetary training costs of unskilled and skilled workers, respectively. This policy increases welfare by about 3.4% and reduces wage inequality by 0.24%. Despite the average gains, such a policy damages skilled workers who disapprove this level of subsidisation.

The Ph.D. thesis provides a comprehensive picture of firm-provided training and it brings about new empirical evidences and theoretical insights. The work validates the importance of government intervention in the training sector and it aims to rekindle a fertile discussion around the issues of equity, productivity, and skill accumulation in modern economies.

A

Data description

This Appendix reports the list of data used to perform the analyses and to calibrate the model of Chapter 2 and 3. Additional information is available on request. Unless stated otherwise, data are pooled to compute aggregate quantities. The samples include employed workers who are 25 to 65 years old. This restriction is commonly used in the literature (e.g. Booth (1993) and Hara (2014)) and is intended to exclude part-timers and apprentices.

Table A.1: Data list

variable	source	frequency	description
average wage, skilled workers	QLFS	1995.1-2015.4	NSA, nominal
average wage, unskilled workers	QLFS	1995.1-2015.4	NSA, nominal
training participation rate, total	QLFS	1994.3-2015.4	NSA, share of workers who trained in last 13 weeks to all workers
training participation rate, skilled workers	QLFS	1994.3-2015.4	NSA, share of workers who trained in last 13 weeks to all workers
training participation rate, unskilled workers	QLFS	1994.3-2015.4	NSA, share of workers who trained in last 13 weeks to all workers
average weekly hours, skilled workers	QLFS	1994.1-2015.4	NSA
average weekly hours, unskilled workers	QLFS	1994.1-2015.4	NSA
average 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, average by SIC sector in 2005 & 2010
per firm training costs	CVTS 3 & 4	2005 & 2010	nominal, average by SIC sector in 2005 & 2010
average firm size (number of employees)	CVTS 3 & 4	2005 & 2010	average number of employees by SIC sector in 2005 & 2010
average 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

B

Derivatives firm's FOCs

The derivatives used in Chapter 2 to simplify the representative firm's FOCs are defined as follows:

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

$$\frac{\partial y_t^f}{\partial l_t^{f,u}} = \mu\omega A\{\mu([l_t^{f,u}\left(1-t_t^u\right)]^{\omega}\left[h_t^u\right]^{1-\omega})^{\alpha} + (1-\mu)\times$$
$$\times[\rho(k_t^f)^{\nu} + (1-\rho)\left([l_t^{f,s}\left(1-t_t^s\right)]^{\omega}\left[h_t^s\right]^{1-\omega}\right)^{\nu}]^{\frac{\alpha}{\nu}}\}^{\frac{1}{\alpha}-1}\times \tag{B.2}$$
$$\times([l_t^{f,u}\left(1-t_t^u\right)]^{\omega}\left[h_t^u\right]^{1-\omega})^{\alpha}(l_t^{f,u})^{-1},$$

$$\frac{\partial h_{t+1}^u}{\partial l_t^{f,u}} = \gamma^u H^u\left(t_t^u h_t^u\right)^{\gamma^u}(l_t^{f,u})^{\gamma^u-1}, \tag{B.3}$$

$$\frac{\partial y_t^f}{\partial l_t^{f,s}} = \omega\left(1-\mu\right)\left(1-\rho\right)A\{\mu([l_t^{f,u}\left(1-t_t^u\right)]^{\omega}\left[h_t^u\right]^{1-\omega})^{\alpha} + (1-\mu)\times$$
$$\times[\rho(k_t^f)^{\nu} + (1-\rho)\left([l_t^{f,s}\left(1-t_t^s\right)]^{\omega}\left[h_t^s\right]^{1-\omega}\right)^{\nu}]^{\frac{\alpha}{\nu}}\}^{\frac{1}{\alpha}-1}[\rho(k_t^f)^{\nu} + (1-\rho)\times \tag{B.4}$$
$$\times([l_t^{f,s}\left(1-t_t^s\right)]^{\omega}\left[h_t^s\right]^{1-\omega})^{\nu}]^{\frac{\alpha}{\nu}-1}([l_t^{f,s}\left(1-t_t^s\right)]^{\omega}\left[h_t^s\right]^{1-\omega})^{\nu}(l_t^{f,s})^{-1},$$

$$\frac{\partial h_{t+1}^s}{\partial l_t^{f,s}} = \gamma^s H^s\left(t_t^s h_t^s\right)^{\gamma^s}(l_t^{f,s})^{\gamma^s-1}, \tag{B.5}$$

$$\frac{\partial y_t^f}{\partial t_t^u} = -\omega\mu A\{\mu([l_t^{f,u}\left(1-t_t^u\right)]^{\omega}\left[h_t^u\right]^{1-\omega})^{\alpha} + (1-\mu)\times$$
$$\times[\rho\left(k_t\right)^{\nu} + (1-\rho)\left([l_t^{f,s}\left(1-t_t^s\right)]^{\omega}\left[h_t^s\right]^{1-\omega}\right)^{\nu}]^{\frac{\alpha}{\nu}}\}^{\frac{1}{\alpha}-1}\times \tag{B.6}$$
$$\times\frac{([l_t^{f,u}\left(1-t_t^u\right)]^{\omega}\left[h_t^u\right]^{1-\omega})^{\alpha}}{1-t_t^u},$$

$$\frac{\partial h_{t+1}^u}{\partial t_t^u} = \gamma^u H^u\left(L_t^u h_t^u\right)^{\gamma^u}\left(t_t^u\right)^{\gamma^u-1}, \tag{B.7}$$

$$
\begin{aligned}
\frac{\partial y_t^f}{\partial t_t^s} = &-(1-\rho)(1-\mu)\,\omega A\{\mu([l_t^{f,u}\left(1-t_t^u\right)]^{\omega}\left[h_t^u\right]^{1-\omega})^{\alpha}+(1-\mu)\times \\
&\times[\rho\,(k_t)^{\nu}+(1-\rho)\,([l_t^{f,s}\left(1-t_t^s\right)]^{\omega}\left[h_t^s\right]^{1-\omega})^{\nu}]^{\frac{\alpha}{\nu}}\}^{\frac{1}{\alpha}-1}[\rho\,(k_t)^{\nu}+(1-\rho)\times \\
&\times([l_t^{f,s}\left(1-t_t^s\right)]^{\omega}\left[h_t^s\right]^{1-\omega})^{\nu}]^{\frac{\alpha}{\nu}-1}([l_t^{f,s}\left(1-t_t^s\right)]^{\omega}\left[h_t^s\right]^{1-\omega})^{\nu}(l_t^{f,s})^{-1},
\end{aligned}
\tag{B.8}
$$

$$
\frac{\partial h_{t+1}^s}{\partial t_t^s} = \gamma^s H^s(l_t^{f,s}h_t^s)^{\gamma^s}\left(t_t^s\right)^{\gamma^s-1}, \tag{B.9}
$$

$$
\begin{aligned}
\frac{\partial y_{t+1}^f}{\partial h_{t+1}^u} = &A\mu\,(1-\omega)\,\{\mu([l_{t+1}^{f,u}\left(1-t_{t+1}^u\right)]^{\omega}\left[h_{t+1}^u\right]^{1-\omega})^{\alpha}+(1-\mu)\times \\
&\times[\rho(k_{t+1}^f)^{\nu}+(1-\rho)\,([l_{t+1}^{f,s}\left(1-t_{t+1}^s\right)]^{\omega}\left[h_{t+1}^s\right]^{1-\omega})^{\nu}]^{\frac{\alpha}{\nu}}\}^{\frac{1}{\alpha}}\times \\
&\times([l_{t+1}^{f,u}\left(1-t_{t+1}^u\right)]^{\omega}\left[h_{t+1}^u\right]^{1-\omega})^{\alpha}(h_{t+1}^u)^{-1},
\end{aligned}
\tag{B.10}
$$

$$
\frac{\partial h_{t+2}^u}{\partial h_{t+1}^u} = 1-\delta^u+\gamma^u H^u\left(t_{t+1}^u l_{t+1}^{f,u}\right)^{\gamma^u}\left(h_{t+1}^u\right)^{\gamma^u-1}, \tag{B.11}
$$

$$
\begin{aligned}
\frac{\partial y_{t+1}^f}{\partial h_{t+1}^s} = &A\,(1-\mu)\,(1-\rho)^{\nu}\,(1-\omega)\,\{\mu([l_{t+1}^{f,u}\left(1-t_{t+1}^u\right)]^{\omega}\left[h_{t+1}^u\right]^{1-\omega})^{\alpha}+ \\
&+(1-\mu)\,[\rho(k_{t+1}^f)^{\nu}+(1-\rho)\,([l_{t+1}^{f,s}\left(1-t_{t+1}^s\right)]^{\omega}\left[h_{t+1}^s\right]^{1-\omega})^{\nu}]^{\frac{\alpha}{\nu}}\}^{\frac{1}{\alpha}}\times \\
&\times[\rho\left(k_{t+1}^f\right)^{\nu}+(1-\rho)\,([l_{t+1}^{f,s}\left(1-t_{t+1}^s\right)]^{\omega}\left[h_{t+1}^s\right]^{1-\omega})^{\nu}]^{\frac{\alpha}{\nu}-1}\times \\
&\times([l_{t+1}^{f,s}\left(1-t_{t+1}^s\right)]^{\omega}\left[h_{t+1}^s\right]^{1-\omega})\left[h_{t+1}^s\right]^{-1},
\end{aligned}
\tag{B.12}
$$

$$\frac{\partial h^s_{t+2}}{\partial h^s_{t+1}} = 1 - \delta^s + \gamma^s H^s \left(t^s_{t+1} l^{f,s}_{t+1}\right)^{\gamma^s} \left(h^s_{t+1}\right)^{\gamma^s - 1}. \tag{B.13}$$

C

The social planner solution

This appendix shows that the solution to the social planner problem is equivalent to the solution to the decentralised economy. This proves that any government intervention does not Pareto improve the market allocation described in Section 2.3.

The planner maximizes the welfare of households, by choosing consumption, investments and all inputs of production. Also, he decides how to divide working time between production and skill-capital accumulation. The social planner guarantees the same level of consumption and welfare to all its members, irrespective of individual labour market status, as the representative household does in the decentralized economy. He decides how much agents should work, and their savings. He maximizes the lifetime utility function:

$$E_0 \sum_{t=0}^{\infty} \beta^t \frac{\left[c_t^{\psi_1} \left(1 - n^u l_t^u\right)^{\psi_2} \left(1 - n^s l_t^s\right)^{\psi_3}\right]^{(1-\sigma)}}{1-\sigma} \tag{C.1}$$

where instantaneous utility is a CRRA function. Keeping the same notation as in the main text, $c_t \equiv (c_t^s)^{n^s}(c_t^u)^{n^u}$ is a weighted average of skilled and unskilled members' consumption; $l_t^i \in (0,1)$, for $i = u, s$, is the household's typical member's labour supply; n^i, for $i \in \{u, s\}$, is the household's share of the relevant type of member, i.e. $n_t^s = (1 - n^u)$; $\sigma \in [0,1)$ measures the intertemporal elasticity of substitution; and $\beta \in (0,1)$ is the discount factor. The first constraint for the planner is the technology that translates factors into the undifferentiated final good. This follows the CES production function:

$$
\begin{aligned}
y_t = A_t \Big\{ & \mu \left(\left[n^u l_t^u \left(1 - t_t^u\right) \right]^{\omega} \left[h_t^u \right]^{1-\omega} \right)^{\alpha} \\
& + (1-\mu) \left[\rho \left(k_t\right)^{\nu} + (1-\rho) \left(\left[n^s l_t^s \left(1 - t_t^s\right) \right]^{\omega} \left[h_t^s \right]^{1-\omega} \right)^{\nu} \right]^{\frac{\alpha}{\nu}} \Big\}^{\frac{1}{\alpha}}
\end{aligned}
\tag{C.2}
$$

where $1/(1-\nu)$ and $1/(1-\alpha)$ are the elasticity of substitution between capital and skilled labour input and between the latter and unskilled labour input, respectively. Provided that $\mu, \rho \in (0,1)$, the input shares of capital, skilled labour and unskilled labour are $(1-\mu)\rho$, $(1-\mu)(1-\rho)$ and μ. Labour inputs are a Cobb-Douglas combination of skill stock and labour units: h_t^i represents the stock of skills accumulated through on-the-job training, and $n^i l_t^i \left(1 - t_t^i\right)$ represents the workers' time that is devoted to production. In this regard, ω and $1-\omega$ are respectively the weight of labour units and of skills in compounding the effective labour input. Finally, A_t represents the exogenous dynamic process of total factor productivity (TFP).

The planner is also constrained by the aggregate condition that consumption plus investment must be equal to total output, net of monetary training costs:

$$
c_t + \frac{k_{t+1}}{P_t} - (1-\delta_k)\frac{k_t}{P_t} = y_t - \phi_u t_t^u n^u l_t^u - \phi_s t_t^s n^s l_t^s \tag{C.3}
$$

where P_t is the inverse of the relative price of capital, following an AR(1) process. Finally, the law of accumulation for skilled and unskilled capital, which is respectively:

$$h^u_{t+1} = (1 - \delta_u)\, h^u_t + H^u \left(n^u l^u_t t^u_t h^u_t\right)^{\gamma^u}, \tag{C.4}$$

and

$$h^s_{t+1} = (1 - \delta_s)\, h^s_t + H^s \left(n^s l^s_t t^s_t h^s_t\right)^{\gamma^s}. \tag{C.5}$$

To sum up, the planner maximizes (C.1) subject to (C.2)-(C.5) with respect to consumption, skilled and unskilled labour, next period physical capital, skilled and unskilled training, and skilled and unskilled human capital. The first order condition with respect to consumption is:

$$\frac{\psi_1}{c_t} \left[c_t^{\psi_1} \left(1 - L^s_t\right)^{\psi_3} \left(1 - L^u_t\right)^{\psi_2}\right]^{1-\sigma} = \lambda^h_t. \tag{C.6}$$

where L^i_t has been used to replace $n^i l^i_t$ for $i = s, u$. The FOCs with respect to skilled and unskilled labour are:

$$\begin{aligned} &\tfrac{\psi_3}{1-L^s_t} \left[c_t^{\psi_1} \left(1 - L^s_t\right)^{\psi_3} \left(1 - L^u_t\right)^{\psi_2}\right]^{1-\sigma} = \\ &\quad H^s \gamma^s \lambda^s_t \frac{\left(h^s_t t^s_t\right)^{\gamma^s}}{\left(L^s_t\right)^{1-\gamma^s}} + \lambda^h_t \left[\frac{(1-\mu)(1-\rho)\omega A^\alpha_t y^{1-\alpha}_t \left(\left[h^s_t\right]^{1-\omega} \left(L^s_t\left[1-t^s_t\right]\right)^{\omega}\right)^{\nu}}{L^s_t\left(\rho k^\nu_t + \left(\left[h^s_t\right]^{1-\omega}\left(L^s_t\left(1-t^s_t\right)\right)^{\omega}\right)^{\nu}(1-\rho)\right)^{1-\frac{\alpha}{\nu}}} - \phi^s t^s_t\right] \end{aligned} \tag{C.7}$$

and

$$\frac{\psi_2}{1-L_t^u}\left(c_t^{\psi_1}\left(1-L_t^s\right)^{\psi_3}\left(1-L_t^u\right)^{\psi_2}\right)^{1-\sigma} = \\ H^u\gamma^u\lambda_t^u\frac{\left(h_t^u t_t^u\right)^{\gamma^u}}{\left(L_t^u\right)^{1-\gamma^u}}+\lambda_t^h\left[\frac{\mu\omega}{L_t^u}A_t^\alpha y_t^{1-\alpha}\left(\left[h_t^u\right]^{1-\omega}\left[L_t^u\left(1-t_t^u\right)\right]^{\omega}\right)^{\alpha}-\phi^u t_t^u\right]. \tag{C.8}$$

Next, maximizing for the stock of next period's physical capital entails the Euler equation:

$$\frac{\lambda_t^h}{P_t\beta}=E_t\lambda_{t+1}^h\left\{\frac{A_{t+1}^\alpha\rho k_{t+1}^{\nu-1}y_{t+1}^{1-\alpha}\left(1-\mu\right)}{\left[\rho k_{t+1}^\nu+\left(\left[h_{t+1}^s\right]^{1-\omega}\left[L_{t+1}^s\left(1-t_{t+1}^s\right)\right]^{\omega}\right)^{\nu}\left(1-\rho\right)\right]^{1-\frac{\alpha}{\nu}}}+\frac{1}{P_{t+1}}\left(1-\delta_k\right)\right\}, \tag{C.9}$$

whilst the optimal training choice requires that:

$$H_s\gamma_s\frac{\lambda_t^s}{\lambda_t^h}\frac{\left(L_t^s h_t^s\right)^{\gamma_s}}{\left(t_t^s\right)^{1-\gamma_s}}=\phi^s L_t^s+\lambda_t^h\frac{\left(1-\mu\right)\left(1-\rho\right)\omega A_t^\alpha y_t^{1-\alpha}\left(\left[h_t^s\right]^{1-\omega}\left[L_t^s\left(1-t_t^s\right)\right]^{\omega}\right)^{\nu}}{\left(1-t_t^s\right)\left[\rho k_t^\nu+\left(\left[h_t^s\right]^{1-\omega}\left[L_t^s\left(1-t_t^s\right)\right]^{\omega}\right)^{\nu}\left(1-\rho\right)\right]^{1-\frac{\alpha}{\nu}}} \tag{C.10}$$

for skilled training, and

$$H^u\gamma^u\frac{\lambda_t^u}{\lambda_t^h}\frac{\left(L_t^u h_t^u\right)^{\gamma^u}}{\left(t_t^u\right)^{1-\gamma^u}}=\phi^u L_t^u+\frac{\mu\omega A_t^\alpha y_t^{1-\alpha}}{1-t_t^u}\left(\left[h_t^u\right]^{1-\omega}\left[L_t^u\left(1-t_t^u\right)\right]^{\omega}\right)^{\alpha} \tag{C.11}$$

for unskilled training. The two FOCs with respect to the next period skill-capital stock are:

$$\begin{aligned}\lambda_t^s =&\beta E_t \lambda_{t+1}^h \frac{(1-\mu)(1-\omega)(1-\rho) A_{t+1}^{\alpha} y_{t+1}^{1-\alpha}\left(\left[h_{t+1}^s\right]^{1-\omega}\left[L_{t+1}^s\left(1-t_{t+1}^s\right)\right]^{\omega}\right)^{\nu}}{h_{t+1}^s\left[\rho k_{t+1}^{\nu}+\left(\left[h_{t+1}^s\right]^{1-\omega}\left[L_{t+1}^s\left(1-t_{t+1}^s\right)\right]^{\omega}\right)^{\nu}(1-\rho)\right]^{1-\frac{\alpha}{\nu}}}\\ &+\beta E_t \lambda_{t+1}^s\left(H^s \gamma^s \frac{\left(L_{t+1}^s t_{t+1}^s\right)^{\gamma^s}}{\left(h_{t+1}^s\right)^{1-\gamma^s}}-\delta^s+1\right),\end{aligned} \tag{C.12}$$

and

$$\begin{aligned}\lambda_t^u =& \beta E_t \lambda_{t+1}^h\left[\frac{(1-\omega)\mu}{h_{t+1}^u} A_{t+1}^{\alpha} y_{t+1}^{1-\alpha}\left(\left[h_{t+1}^u\right]^{1-\omega}\left[L_{t+1}^u\left(1-t_{t+1}^u\right)\right]^{\omega}\right)^{\alpha}\right]\\ &+\beta E_t \lambda_{t+1}^u\left[H^u \gamma^u \frac{\left(L_{t+1}^u t_{t+1}^u\right)^{\gamma^u}}{\left(h_{t+1}^u\right)^{1-\gamma^u}}-\delta^u+1\right].\end{aligned} \tag{C.13}$$

Initial and transversality conditions

The first order conditions are necessary condition for a solution of the planner's problem. To ensure the existence of a solution, I assume that:

$$k_0 = \bar{k} > 0, \tag{C.14}$$

$$h_0^s = \bar{h}^s > 0, \tag{C.15}$$

and

$$h_0^u = \bar{h}^u > 0. \tag{C.16}$$

Further, since agents are infinitely-living, to ensure a finite and unique solution to the model exists (see e.g. Kamihigashi (2001)), I impose the following transversality conditions:

$$\lim_{t\to\infty} E_t \beta^t \lambda_t^h k_{t+1} = 0, \tag{C.17}$$

$$\lim_{t\to\infty} E_t \beta^t \lambda_t^h h_{t+1}^s = 0, \tag{C.18}$$

and

$$\lim_{t\to\infty} E_t \beta^t \lambda_t^h h_{t+1}^u = 0. \tag{C.19}$$

Social Planner solution

Given initial conditions (C.14)-(C.16) and the path of exogenous innovations $\{A_t, P_t\}_{t=0}^{\infty}$, the social planner solution is defined to be an allocation $\{c_t, l_t^s, l_t^u, t_t^s, t_t^u, k_{t+1}, h_{t+1}^s, h_{t+1}^u, \lambda_t^h, \lambda_t^s, \lambda_t^u, y_t\}_{t=0}^{\infty}$ such that (i) the planner's budget is binding (C.3); (ii) transversality conditions hold; (iii) all FOCs (C.6)-(C.12) hold; (iv) skill-capital stocks evolve according to equations (C.5) and (C.4).

Equivalence of social planner and decentralised economy

The equivalence between the solution of the planner and the decentralised economy can be appreciated by observing that, once prices and wages are substituted out, the first order conditions (2.6)-(2.9) and (2.20)-(2.26) are equivalent to the planner's, so that the solution must also be the coincident.

As additional check, I solve numerically for the steady state the social planner economy using the same calibration reported in Chapter 2 and I find that the allocation, e.g. the level of consumption, physical capital and etcetera, is the same as for the steady state of the competitive economy.

In conclusion, the decentralised equilibrium is a Pareto optimal allocation that respects the first Welfare Theorem. Broadly speaking, this implies that the model economy is characterised by: (i) absence of externalities, (ii) completeness of markets, and (iii) absence of distorting taxes, such as income and sale taxes (see e.g. Hammond (1998)).

D

Derivatives firm's FOCs with poaching

This Appendix reports the derivatives used in Chapter 3 to simplify the representative firm's FOCs. As it can be noted, they are the same as those for Chapter 2 reported in Appendix B except the intertemporal derivatives of skilled and unskilled human capital. This result follows intuitively from the fact that the production function is the same and that poaching only affects the depreciation of skills (and not the marginal product of working or training time). Nonetheless, I report all the derivatives utilised in Chapter 3. Those are defined as follows:

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

$$\begin{aligned}
\frac{\partial y_t^f}{\partial l_t^{f,u}} &= \mu\omega A\{\mu([l_t^{f,u}\left(1-t_t^u\right)]^{\omega}\left[h_t^u\right]^{1-\omega})^{\alpha} + (1-\mu)\times \\
&\times[\rho(k_t^f)^{\nu} + (1-\rho)\,([l_t^{f,s}\left(1-t_t^s\right)]^{\omega}\left[h_t^s\right]^{1-\omega})^{\nu}]^{\frac{\alpha}{\nu}}\}^{\frac{1}{\alpha}-1}\times \\
&\times([l_t^{f,u}\left(1-t_t^u\right)]^{\omega}\left[h_t^u\right]^{1-\omega})^{\alpha}(l_t^{f,u})^{-1},
\end{aligned} \tag{D.2}$$

$$\frac{\partial h_{t+1}^u}{\partial l_t^{f,u}} = \gamma^u H^u\left(t_t^u h_t^u\right)^{\gamma^u}(l_t^{f,u})^{\gamma^u-1}, \tag{D.3}$$

$$\begin{aligned}
\frac{\partial y_t^f}{\partial l_t^{f,s}} &= \omega\,(1-\mu)\,(1-\rho)\,A\{\mu([l_t^{f,u}\left(1-t_t^u\right)]^{\omega}\left[h_t^u\right]^{1-\omega})^{\alpha} + (1-\mu)\times \\
&\times[\rho(k_t^f)^{\nu} + (1-\rho)\,([l_t^{f,s}\left(1-t_t^s\right)]^{\omega}\left[h_t^s\right]^{1-\omega})^{\nu}]^{\frac{\alpha}{\nu}}\}^{\frac{1}{\alpha}-1}[\rho(k_t^f)^{\nu} + (1-\rho)\times \\
&\times([l_t^{f,s}\left(1-t_t^s\right)]^{\omega}\left[h_t^s\right]^{1-\omega})^{\nu}]^{\frac{\alpha}{\nu}-1}([l_t^{f,s}\left(1-t_t^s\right)]^{\omega}\left[h_t^s\right]^{1-\omega})^{\nu}(l_t^{f,s})^{-1},
\end{aligned} \tag{D.4}$$

$$\frac{\partial h_{t+1}^s}{\partial l_t^{f,s}} = \gamma^s H^s\left(t_t^s h_t^s\right)^{\gamma^s}(l_t^{f,s})^{\gamma^s-1}, \tag{D.5}$$

$$\frac{\partial y_t^f}{\partial t_t^u} = -\omega\mu A\{\mu([l_t^{f,u}\left(1-t_t^u\right)]^\omega \left[h_t^u\right]^{1-\omega})^\alpha + (1-\mu)\times$$
$$\times[\rho\left(k_t\right)^\nu + (1-\rho)\left([l_t^{f,s}\left(1-t_t^s\right)]^\omega \left[h_t^s\right]^{1-\omega}\right)^\nu]^{\frac{\alpha}{\nu}}\}^{\frac{1}{\alpha}-1}\times \tag{D.6}$$
$$\times\frac{([l_t^{f,u}(1-t_t^u)]^\omega\left[h_t^u\right]^{1-\omega})^\alpha}{1-t_t^u},$$

$$\frac{\partial h_{t+1}^u}{\partial t_t^u} = \gamma^u H^u \left(L_t^u h_t^u\right)^{\gamma^u}\left(t_t^u\right)^{\gamma^u-1}, \tag{D.7}$$

$$\frac{\partial y_t^f}{\partial t_t^s} = -\left(1-\rho\right)\left(1-\mu\right)\omega A\{\mu([l_t^{f,u}\left(1-t_t^u\right)]^\omega \left[h_t^u\right]^{1-\omega})^\alpha + (1-\mu)\times$$
$$\times[\rho\left(k_t\right)^\nu + (1-\rho)\left([l_t^{f,s}\left(1-t_t^s\right)]^\omega \left[h_t^s\right]^{1-\omega}\right)^\nu]^{\frac{\alpha}{\nu}}\}^{\frac{1}{\alpha}-1}[\rho\left(k_t\right)^\nu + (1-\rho)\times \tag{D.8}$$
$$\times([l_t^{f,s}\left(1-t_t^s\right)]^\omega \left[h_t^s\right]^{1-\omega})^\nu]^{\frac{\alpha}{\nu}-1}([l_t^{f,s}\left(1-t_t^s\right)]^\omega \left[h_t^s\right]^{1-\omega})^\nu(l_t^{f,s})^{-1},$$

$$\frac{\partial h_{t+1}^s}{\partial t_t^s} = \gamma^s H^s (l_t^{f,s} h_t^s)^{\gamma^s}\left(t_t^s\right)^{\gamma^s-1}, \tag{D.9}$$

$$\frac{\partial y_{t+1}^f}{\partial h_{t+1}^u} = A\mu\left(1-\omega\right)\{\mu([l_{t+1}^{f,u}\left(1-t_{t+1}^u\right)]^\omega \left[h_{t+1}^u\right]^{1-\omega})^\alpha + (1-\mu)\times$$
$$\times[\rho(k_{t+1}^f)^\nu + (1-\rho)\left([l_{t+1}^{f,s}\left(1-t_{t+1}^s\right)]^\omega \left[h_{t+1}^s\right]^{1-\omega}\right)^\nu]^{\frac{\alpha}{\nu}}\}^{\frac{1}{\alpha}}\times \tag{D.10}$$
$$\times([l_{t+1}^{f,u}\left(1-t_{t+1}^u\right)]^\omega \left[h_{t+1}^u\right]^{1-\omega})^\alpha(h_{t+1}^u)^{-1},$$

$$\frac{\partial h_{t+2}^u}{\partial h_{t+1}^u} = 1-\delta^u-\delta_\varepsilon^u+\gamma^u H^u\left(t_{t+1}^u l_{t+1}^{f,u}\right)^{\gamma^u}\left(h_{t+1}^u\right)^{\gamma^u-1}, \tag{D.11}$$

$$
\begin{aligned}
\frac{\partial y^{f}_{t+1}}{\partial h^{s}_{t+1}} &= A\left(1-\mu\right)\left(1-\rho\right)^{\nu}\left(1-\omega\right)\{\mu([l^{f,u}_{t+1}\left(1-t^{u}_{t+1}\right)]^{\omega}\left[h^{u}_{t+1}\right]^{1-\omega})^{\alpha}+\\
&+\left(1-\mu\right)[\rho(k^{f}_{t+1})^{\nu}+\left(1-\rho\right)([l^{f,s}_{t+1}\left(1-t^{s}_{t+1}\right)]^{\omega}\left[h^{s}_{t+1}\right]^{1-\omega})^{\nu}]^{\frac{\alpha}{\nu}}\}^{\frac{1}{\alpha}}\times\\
&\times[\rho\left(k^{f}_{t+1}\right)^{\nu}+\left(1-\rho\right)([l^{f,s}_{t+1}\left(1-t^{s}_{t+1}\right)]^{\omega}\left[h^{s}_{t+1}\right]^{1-\omega})^{\nu}]^{\frac{\alpha}{\nu}-1}\times\\
&\times([l^{f,s}_{t+1}\left(1-t^{s}_{t+1}\right)]^{\omega}\left[h^{s}_{t+1}\right]^{1-\omega})\left[h^{s}_{t+1}\right]^{-1},
\end{aligned}
\tag{D.12}
$$

$$
\frac{\partial h^{s}_{t+2}}{\partial h^{s}_{t+1}} = 1-\delta^{s}-\delta^{s}_{\varepsilon}+\gamma^{s}H^{s}\left(t^{s}_{t+1}l^{f,s}_{t+1}\right)^{\gamma^{s}}\left(h^{s}_{t+1}\right)^{\gamma^{s}-1}. \tag{D.13}
$$

E

Total factor productivity in Chapter 3

As discussed in Section 2.4, after a total factor productivity (TFP) shock, the model of chapter 2 generates dynamics that are in line with the real business cycle (RBC) literature. This appendix assesses the role of training during business cycles fluctuations by comparing a simulation to a temporary TFP shock with the model of Chapter 3 with the simulation run on the same model but where training does not exists.

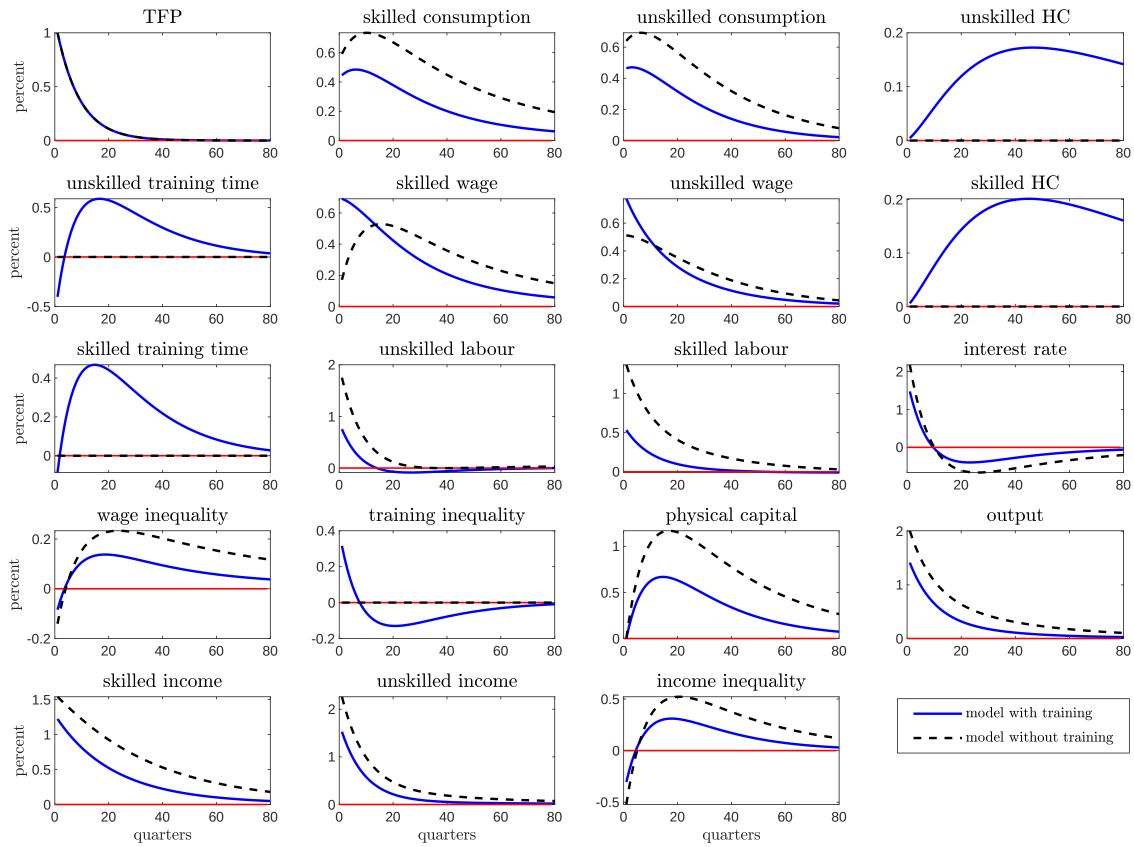


Figure E.1: Comparison of a 1% temporary increase in total factor productivity

To exclude the training channel, it is sufficient to assume that: the human capital stocks, h_t^s and h_t^u are fixed and equal to their respective steady state;[1] training time is zero, i.e. $t_t^s, t_t^u = 0$; and the monetary resources spent for training

[1]This way the stocks could be interpreted as a fixed production input, such as land, owned by firms.

subsidies are (unproductive) public spending.

Figure E.1 reports the impulse response functions of both models. Qualitatively, the dynamics after a temporary TFP shock is the same whether training is part of the model or not. The only major difference is that, with endogenous training, wages jumps up at the arrival of the shock. Conversely, without training, wages rise less and more slowly (sub-plots (2,2) and (2,3)).

This difference can be attributed to the effects of higher TFP on the demand of training time. As companies seek to accumulate human capital skills, they need to occupy workers both in training and production activities. Thus, firms offer higher wages to bid workers' labour input.[2]

Lastly, it can be noted that endogenous training reduces the wage inequality that is generated, in the medium run, by the temporary positive TFP innovation (sub-plot (4,1)). Possibly, this is related to the lower stock of physical capital accumulated in the presence of the training sector.

In conclusion, these results suggest that the inclusion of the channel for human capital accumulation through training does not alter the basic properties of the model.

[2]This explanations seems at odds with the higher labour supply I observe in the model without training (sub-plots (3,2) and (3,3)). However, it is important to remember that also the marginal utility of consumption matters for the labour market equilibrium.

F

Simulation of the poaching externality shock

This appendix reports the simulation exercise which I use to compute the elasticity of training with respect to the job-to-job separation rate. The experiment shows the effects of an increase in the poaching externality δ_{ε}^{s} and δ_{ε}^{u} on the endogenous variables of the model employed in Chapter 3.

Experimenting with the model proves that the elasticity of training to an exogenous shock to these parameters and their values are proportionally related. As explained in Section 3.4, this paves the way to the calibration of the poaching externality.

Figure F.1 reports the dynamics of endogenous variables to a 1% permanent increase in both skilled and unskilled firm poaching externality. As expected, the reaction of the economy to this shock is very small. Most variables deviate from their steady state by less than one basis point.

The shock has the greatest impact on the skill capital stocks which is depleted because firms are less willingly to train workers. In fact, both skilled and unskilled training decreases (sub-plots (4,5) and (5,1)). The increase in the separation rate has overall negative consequences. In particular, total output declines and households have lower labour income. The skill premium increases (sub-plot (3,5)), as skilled wage initially rises, but this change is more than compensated by the changes in labour supply, as indicated by the steady fall of income inequality (sub-plot (3,4)).

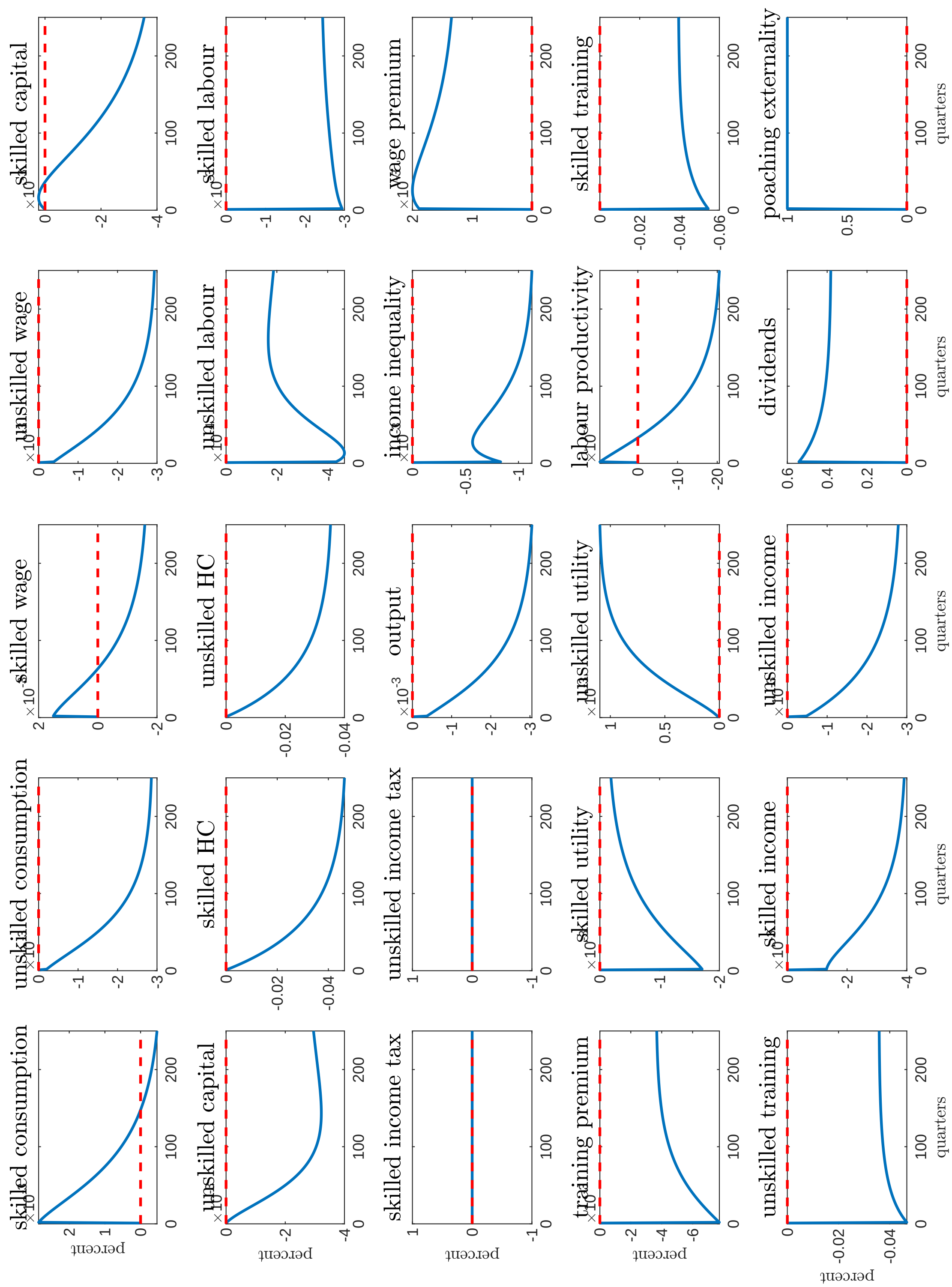


Figure F.1: Permanent increase in δ_ε^s and δ_ε^u by 1%

G

Further fiscal policies for Chapter 3

As explained in the main text, increasing the training subsidy rate requires additional fiscal resources to be collected by the government. In order to ensure a balanced budget, in Chapter 3 I have assumed that either both households get taxed at a higher rate or only the skilled tax rate is increased. Here I report the dynamics and the quantitative impact of skilled and unskilled training subsidies under the assumption that the government can collect lump-sum transfers to cover for the extra expenditures (the rest is still collected through income taxation).

Unskilled subsidies

Figure G.1 allows a comparison of the unskilled training when those are financed through distortionary taxed on both households and when those are financed through lump-sum transfers. Clearly, non-distortionary taxes entail a much more beneficial effect of training subsidies than otherwise.

When considering the steady state effects of the policies, under lump-sum transfers, it not surprising that those are generally larger than the effects when distortionary taxation is taken into account. Most remarkably, the path of physical capital is much higher whenever taxes are non-distortionary, and both skilled and unskilled consumption is much higher than it would be under distortionary taxation.

Table G.1 compares the effects of unskilled training subsidies under lump-sum and both household taxed assumptions in terms of percent changes from the original steady state (i.e. comparative statics).

It can be observed that financing the intervention through distortionary taxation causes a stronger reduction of inequality compared to lump-sum transfers, even if the burden is equally shared across households. After accounting for the

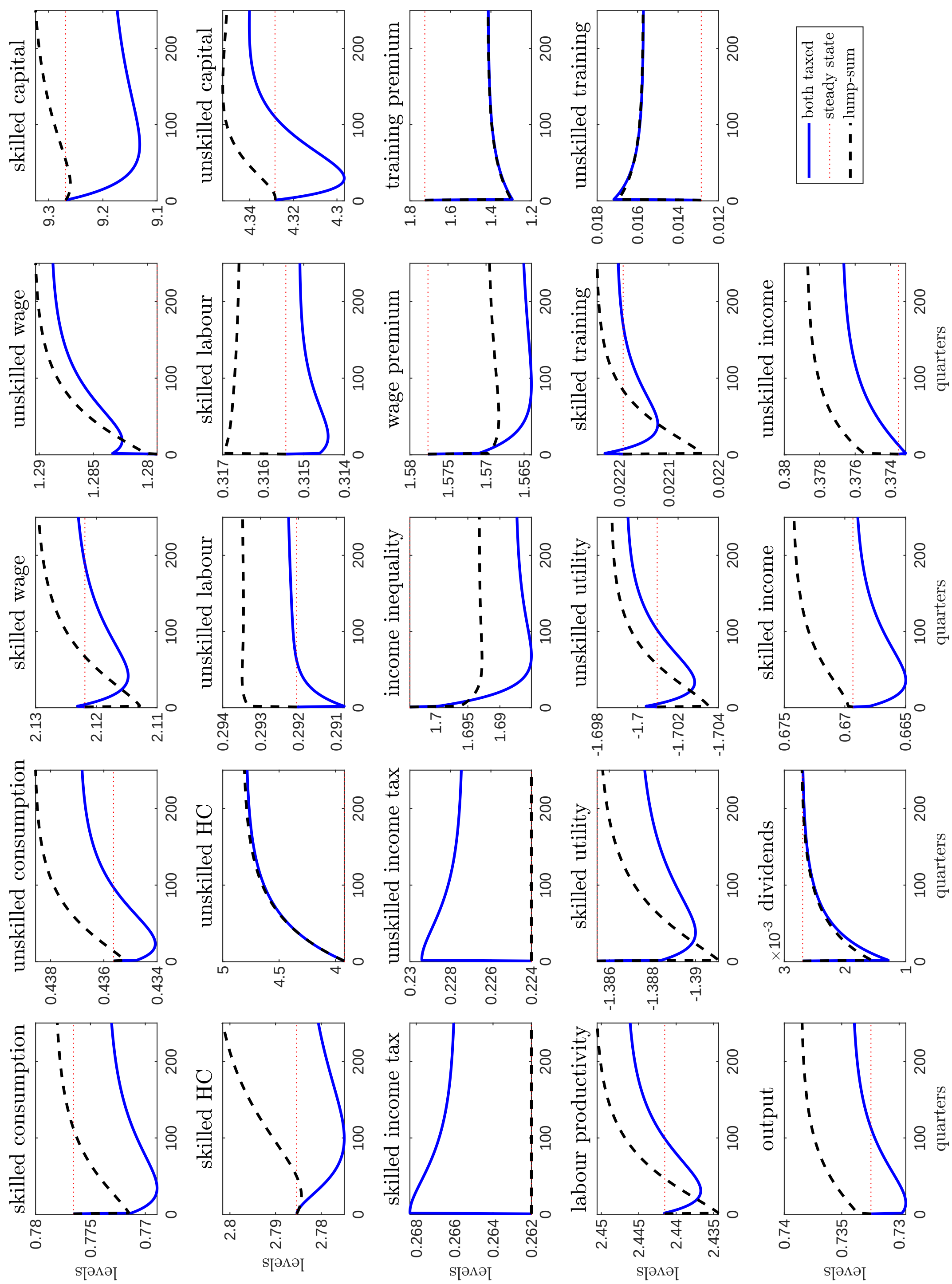


Figure G.1: Permanent increase of unskilled training subsidy to 0.25, lump-sum financing

Table G.1: Steady-state effects of increasing unskilled training subsidies under lump-sum financing

	lump-sum			both tax rates		
$\tau^u =$	0.10	0.20	0.30	0.10	0.20	0.30
$\%\Delta t^u$	6.5	17	29	6.5	17	28
$\%\Delta t^s$	0.1	0.2	0.3	0.0	0.1	0.1
$\%\Delta \frac{t^s}{t^u}$	-6.1	-14	-22	-6.1	-14	-22
$\%\Delta w^u \left(1-\tau^{h,u}\right)$	0.3	0.7	1.1	0.1	0.3	0.4
$\%\Delta w^s \left(1-\tau^{h,s}\right)$	0.1	0.3	0.5	-0.1	-0.3	-0.6
$\%\Delta \frac{w^s(1-\tau^{h,s})}{w^u(1-\tau^{h,u})}$	-0.2	-0.4	-0.6	-0.2	-0.6	-1.0
$\%\Delta w^u l^u \left(1-\tau^{h,u}\right)$	0.4	1.0	1.8	0.2	0.3	0.5
$\%\Delta w^s l^s \left(1-\tau^{h,s}\right)$	0.2	0.6	0.9	-0.1	-0.4	-0.7
$\%\Delta \frac{(1-\tau^{h,s})l^s w^s}{(1-\tau^{h,u})l^u w^u}$	-0.2	-0.5	-0.8	-0.3	-0.7	-1.2
$\%\Delta \frac{y}{(l^s n^s + l^u n^u)}$	0.1	0.3	0.5	0.0	0.0	0.0
$\%\Delta U^u/\lvert U^u \rvert$	0.0	0.1	0.2	0.03	0.07	0.10
$\%\Delta U^s/\lvert U^s \rvert$	0.0	0.0	-0.0	-0.03	-0.09	-0.18
$\%\Delta(t^s l^s \phi^s \tau^s + t^u l^u \phi^u \tau^u)$	1.5	3.8	6.5	1.5	3.9	6.8

distortionary taxation and the poaching externality, the effects of training subsidies are larger in term of inequality but smaller in terms of aggregate outcomes.

Comparing the lump-sum transfer case to the other scenario, it can be observed that distortionary taxation produces non-linear utility gains. The gains, in terms of unskilled utility, diminish as the subsidy rate increases. Valued at the steady state, unskilled utility improves by 0.03% when subsidies are 10%, by 0.07% when subsidies are 20%, but only by 0.1% when subsidies are 30%.

To complete the picture, Table G.2 reports the multipliers for unskilled subsidies under the two financing assumptions.

All fiscal multipliers are larger under the lump-sum transfers as the behaviour

Table G.2: Lifetime multipliers after increasing unskilled training subsidies under lump-sum financing

	lump-sum			both tax rates		
τ_u =	0.10	0.20	0.30	0.10	0.20	0.30
$w^u l^u \left(1-\tau^{h,u}\right)$	2.45	2.37	2.29	0.69	0.57	0.44
$w^s l^s \left(1-\tau^{h,s}\right)$	2.13	2.06	1.99	-1.87	-2.04	-2.23
$w^u l^u (1-\tau^{h,u}) + w^s l^s (1-\tau^{h,s})$	2.24	2.16	2.09	-1.00	-1.16	-1.32
$(1+r\left(1-\tau^{h,s}\right)-\delta_k)k^s$	33	32	31	-53	-56	-59
$(1+r\left(1-\tau^{h,u}\right)-\delta_k)k^u$	12	12	12	3.92	3.42	2.90
π	-0.13	-0.13	-0.12	-0.66	-0.67	-0.68
y	3.64	3.53	3.41	-0.23	-0.45	-0.68
$U^u - \bar{U}^u$	0.39	0.90	1.37	0.26	0.54	0.73
$U^s - \bar{U}^s$	-0.09	-0.29	-0.61	-0.39	-1.08	-2.01

of the households is not influenced by the latter.

Skilled subsidies

Figure G.1 allows a comparison of the skilled training when those are financed through distortionary taxed on both households and when those are financed through lump-sum transfers. Again, I observe that, most of the time, the IRFs under lump-sum transfers are above the IRFs with distortionary taxation for both income flow or stock variables.

The same analysis is re-proposed for the case of skilled training subsidies. Thus, Table G.3 reports the percent change from the original steady state of key variables, and Table G.4 contains the present value multipliers.

The welfare evaluation provides the same insights. Lump-sum transfers would

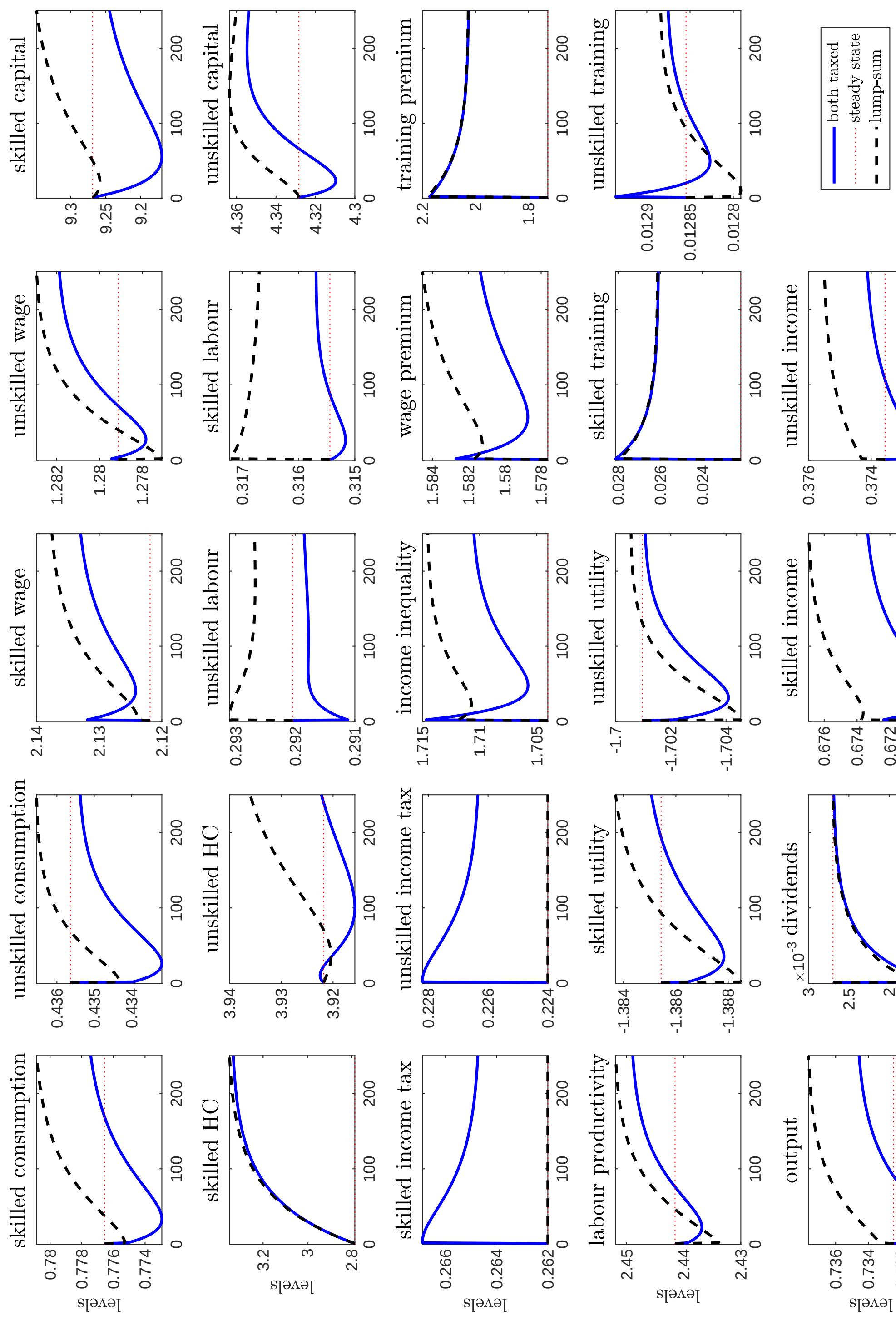


Figure G.2: Permanent increase of skilled training subsidy to 0.25, lump-sum financing imply large utility gains. When accounting for distortionary taxation, although benefits are still significant, the size of the subsidy plays a key role.

Table G.3: Steady-state effects of increasing skilled training subsidies under lump-sum financing

	lump-sum			both tax rates		
$\tau^s =$	0.10	0.20	0.30	0.10	0.20	0.30
$\%\Delta t^u$	0.08	0.18	0.29	0.1	0.1	0.2
$\%\Delta t^s$	5.3	13.2	22.4	5.3	13	22
$\%\Delta \frac{t^s}{t^u}$	5.2	13.0	22.0	5.2	13	22
$\%\Delta w^u \left(1-\tau^{h,u}\right)$	0.10	0.24	0.38	0.0	0.0	-0.1
$\%\Delta w^s \left(1-\tau^{h,s}\right)$	0.24	0.59	0.97	0.1	0.2	0.2
$\%\Delta \frac{w^s(1-\tau^{h,s})}{w^u(1-\tau^{h,u})}$	0.14	0.35	0.58	0.1	0.2	0.3
$\%\Delta w^u l^u \left(1-\tau^{h,u}\right)$	0.17	0.41	0.67	0.0	-0.1	-0.2
$\%\Delta w^s l^s \left(1-\tau^{h,s}\right)$	0.36	0.88	1.46	0.1	0.2	0.3
$\%\Delta \frac{(1-\tau^{h,s}) l^s w^s}{(1-\tau^{h,u}) l^u w^u}$	0.19	0.47	0.78	0.1	0.3	0.5
$\%\Delta \frac{y}{(l^s n^s + l^u n^u)}$	0.16	0.38	0.60	0.1	0.2	0.3
$\%\Delta U^u/\lvert U^u\rvert$	0.01	0.02	0.02	0.00	0.00	0.01
$\%\Delta U^s/\lvert U^s\rvert$	0.05	0.11	0.18	0.02	0.05	0.06
$\%\Delta(t^s l^s \phi^s \tau^s + t^u l^u \phi^u \tau^u)$	3.9	9.7	16.4	3.9	9.8	16.6

Table G.4 contains the fiscal multipliers for the skilled subsidy. As expected, the table shows that financing the subsidies with lump-sum transfers would be more effective than financing them through distortionary taxation.

With respect to the fiscal multiplier of output, unskilled subsidies have larger multipliers than skilled subsidies when these interventions are financed through lump-sum transfers. Conversely, skilled subsidies have larger output multipliers than unskilled subsidies when financed by distortionary taxes. For example, spending 1£ on skilled training can increase total output by 0.37£ when increasing

Table G.4: Lifetime multipliers after increasing skilled training subsidies under lump-sum financing

	lump-sum			both tax rates		
$\tau_s =$	0.10	0.20	0.30	0.10	0.20	0.30
$w^u l^u \left(1-\tau^{h,u}\right)$	0.30	0.29	0.28	-0.09	-0.11	-0.13
$w^s l^s \left(1-\tau^{h,s}\right)$	1.12	1.09	1.07	0.22	0.17	0.11
$w^u l^u (1-\tau^{h,u}) + w^s l^s (1-\tau^{h,s})$	0.84	0.82	0.80	0.12	0.07	0.03
$(1+r\left(1-\tau^{h,s}\right)-\delta_k)k^s$	16	15	14	-3.3	-4.5	-5.8
$(1+r\left(1-\tau^{h,u}\right)-\delta_k)k^u$	6.1	5.8	5.6	4.2	3.9	3.6
π	-0.01	-0.01	-0.01	-0.13	-0.13	-0.13
y	1.24	1.19	1.15	0.37	0.30	0.23
$U^u - \bar{U}^u$	0.04	0.04	-0.04	-0.06	-0.20	-0.46
$U^s - \bar{U}^s$	0.29	0.66	1.02	0.08	0.11	0.05

training subsidies to 10% of monetary costs whereas unskilled training depresses output (negative multiplier).

In conclusion, it appears that the combination of training subsidies and taxes affect the allocation of labour input with overall effects that depend the interplay of the different policy tools employed. I leave to the reader any further consideration.

References

[1] Acemoglu, D. and J. Angrist (2001). How large are human-capital externalities? evidence from compulsory-schooling laws. In B. S. Bernanke and K. Rogoff (Eds.), *NBER Macroeconomics Annual 2000*, Volume 15, pp. 9–74. MIT Press, Cambridge.

[2] Acemoglu, D. and D. Autor (2011). Skills, tasks and technologies: implications for employment and earnings. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 4B, Chapter 12, pp. 1043–1171. Elsevier Science B.V., Amsterdam.

[3] Acemoglu, D. and J.-S. Pischke (1999). Beyond becker: Training in imperfect labour markets. *The Economic Journal 109*(453), 112–142.

[4] Acemoglu, D. and J.-S. Pischke (2003). Minimum wages and on-the-job training. *Research in Labor Economics 22*, 153–202.

[5] Acemoglu, D. and J.-S. Pischke (2007). The structure of wages and investment in general training. *Journal of Political Economy 107*(3), 539–572.

[6] Adjemian, S. and M. Juillard (2011). Accuracy of the extended path simulation method in a new Keynesian model with zero lower bound on the nominal interest rate. Cepremap working papers, CEntre Pour la Recherche Economique et ses Applications.

[7] Akerlof, G. A. and J. L. Yellen (1990). The fair wage-effort hypothesis and unemployment. *The Quarterly Journal of Economics 105*(2), 255–283.

[8] Albert, C., C. García-Serrano, and V. Hernanz (2010). On-the-job training in Europe: determinants and wage returns. *International Labour Review 149*(3), 315–341.

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

[10] Almeida, R. K. and R. Aterido (2015). Investing in formal on-the-job training: are SMEs lagging much behind? *IZA Journal of Labor and Development 4*(1), 1–8.

[11] Almeida, R. K. 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), 1–19.

[12] Amemiya, T. (1978). The estimation of a simultaneous equation generalized probit model. *Econometrica 46*(5), 1193–1205.

[13] Andreasen, M. M. (2012). An estimated DSGE model: explaining variation in nominal term premia, real term premia, and inflation risk premia. *European Economic Review 56*(8), 1656–1674.

[14] Angelopoulos, K., A. Benecchi, and J. Malley (2017). Can subsidising job-related training reduce inequality? CESifo Working Paper Series 6605, Adam Smith Business School, University of Glasgow.

[15] Angelopoulos, K., S. Lazarakis, and J. Malley (2017a). Asymmetries in earnings, employment and wage risk in Great Britain. CESifo Working Paper Series 6400, Adam Smith Business School, University of Glasgow.

[16] Angelopoulos, K., S. Lazarakis, and J. Malley (2017b). Wealth inequality and externalities from ex ante skill heterogeneity. CESifo Working Paper Series 6572, Adam Smith Business School, University of Glasgow.

[17] Arrondel, L., M. Debbich, and F. Savignac (2013). Financial literacy and financial planning in France. *Numeracy 6*(2), 1–17.

[18] Arulampalam, W., A. L. Booth, and M. L. Bryan (2007). Is there a glass ceiling over Europe? exploring the gender pay gap across the wage distribution. *ILR Review 60*(2), 163–186.

[19] Autor, D. H. (2014). Skills, education, and the rise of earnings inequality among the other 99 percent. *Science 344*(6186), 843–851.

[20] Bakış, O., B. Kaymak, and M. Poschke (2015). Transitional dynamics and the optimal progressivity of income redistribution. *Review of Economic Dynamics 18*(3), 679 – 693.

[21] Barnett, A., S. Batten, A. Chiu, J. Franklin, and M. Sebastiá-Barriel (2014). The UK productivity puzzle. BOE Quarterly Bulletin 1, Bank of England.

[22] Barron, J. M., D. A. Black, and M. A. Loewenstein (1989). Job matching and on-the-job training. *Journal of Labor Economics 7*(1), 1–19.

[23] Bartel, A. P. (1994). Productivity gains from the implementation of employee training programs. *Industrial Relations 33*(4), 411–425.

[24] Bassanini, A., A. Booth, G. Brunello, M. D. Paola, and E. Leuven (2005). Workplace training in Europe. IZA Discussion Papers 1640, IZA Institute of Labor Economics.

[25] Basu, P. and C. Thoenissen (2007, January). Investment frictions and the relative price of investment goods in an open economy model. CDMA Working Paper Series 2007-04, Centre for Dynamic Macroeconomic Analysis.

[26] Battu, H., C. R. Belfieldà, and P. J. Sloane (2003). Human capital spillovers within the workplace: evidence for Great Britain. *Oxford Bulletin of Economics and Statistics 65*(5), 575–593.

[27] Baum, C., M. Schaffer, and S. Stillman (2018). Ivreg2: Stata module for extended instrumental variables/2SLS and GMM estimation. *Stata Journal 5*(4), 1–09.

[28] Beck, N., R. Kabst, and P. Walgenbach (2009). The cultural dependence of vocational training. *Journal of International Business Studies 40*(8), 1374–1395.

[29] Becker, G. S. (1962). Investment in human capital: a theoretical analysis. *Journal of Political Economy 70*(5), 9–49.

[30] Belfield, C., C. Crawford, and L. Sibieta (2017). Long-run comparisons of spending per pupil across different stages of education. IFS technical report R126, Institute for Fiscal Studies, London.

[31] Ben-Porath, Y. (1967). The production of human capital and the life cycle of earnings. *Journal of Political Economy 75*(4), 352–365.

[32] Benigno, P. (2009). Price stability with imperfect financial integration. *Journal of Money, Credit and Banking 41*(1), 121–149.

[33] Bilanakos, C., C. P. Green, J. S. Heywood, and N. Theodoropoulos (2017). Do dominant firms provide more training? *Journal of Economics and Management Strategy 26*(1), 67–95.

[34] Black, D., B. Noel, and Z. Wang (1999). On-the-job training, establishment size, and firm size: evidence for economies of scale in the production of human capital. *Southern Economic Journal 66*(1), 82–100.

[35] Black, S. and L. M. Lynch (2001). How to compete: the impact of workplace practices and information technology on productivity. *Review of Economics and Statistics 83*(3), 434–445.

[36] Black, S. E. . and L. M. Lynch (1996). Human-capital investments and productivity. *The American Economic Review 86*(2), 263–267.

[37] Blatter, M., S. Muehlemann, S. Schenker, and S. C. Wolter (2012). Hiring costs of skilled workers and the supply of firm-provided training. IZA Discussion Papers 6344, IZA Institute of Labor Economics.

[38] Blau, F. D. and L. M. Kahn (2017). The gender wage gap: extent, trends, and explanations. *Journal of Economic Literature 55*(3), 789–865.

[39] Blau, P. (1964). Justice in social exchange. *Sociological Inquiry 34*(2), 193–207.

[40] Blundell, R., C. Crawford, and W. Jin (2014). What can wages and employment tell us about the UK's productivity puzzle? *Economic Journal 124*(576), 377–407.

[41] Blundell, R., L. Dearden, and C. Meghir (1996). The determinants and effects of work-related training in Britain. IFS Working Paper, Institute for Fiscal Studies, London.

[42] Blundell, R., L. Dearden, C. Meghir, 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), 1–23.

[43] Blundell, R. and B. Etheridge (2010). Consumption, income and earnings inequality in Britain. *Review of Economic Dynamics 13*(1), 76–102.

[44] Booth, A. and D. Snower (1996). *Acquiring Skills: market failures, their symptoms and policy responses.* Cambridge University Press, Cambridge.

[45] Booth, A. L. (1993). Private sector training and graduate earnings. *Review of Economics and Statistics 75*(1), 164–170.

[46] Booth, A. L. and M. L. Bryan (2002). Who pays for general training? new evidence for British men and women. IZA Discussion Papers 486, IZA Institute of Labor Economics.

[47] Booth, A. L. and M. Francesconi (2003). Unions, work-related training, and wages: evidence for British men. *Industrial and Labor Relations Review 57*(1), 68–91.

[48] Bratti, M. and R. Leombruni (2014). Local human capital externalities and wages at the firm level: evidence from Italian manufacturing. *Economics of Education Review 41*, 161 – 175.

[49] 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.

[50] Brown, J. N. (1989). Why do wages increase with tenure? on-the-job training and life-cycle wage growth observed within firms. *American Economic Review 79*(5), 971–91.

[51] Browning, M., L. P. Hansen, and J. J. Heckman (1999). Micro data and general equilibrium models. In J. Taylor and M. Woodford (Eds.), *Handbook of Macroeconomics*, Volume 1, Chapter 8, pp. 543–633. Elsevier, Amsterdam.

[52] Brunello, G. (2009). The effect of economic downturns on apprenticeships and initial workplace training: a review of the evidence. *Empirical Research in Vocational Education and Training 1*(2), 145–171.

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

[54] Brunello, G. and F. Gambarotto (2004). Agglomeration effects on employer-provided training: evidence from the UK. CESifo Working Papers 1055, Center for Economic Studies.

[55] Brunello, G., P. Garibaldi, and E. Wasmer (2007). *Education and Training in Europe*. Oxford University Press, Oxford.

[56] Budría, S. (2012). The shadow value of employer-provided training. *Journal of Economic Psychology 33*(3), 494–514.

[57] Budría, S. and P. T. Pereira (2007). The wage effects of training in Portugal: differences across skill groups, genders, sectors and training types. *Applied Economics 39*(6), 787–807.

[58] Busemeyer, M. R. (2014). *Skills and Inequality: partisan Politics and the Political Economy of Education Reforms in Western Welfare States*. Cambridge University Press, Cambridge.

[59] Campbell, J. Y. (2006). Household finance. *The Journal of Finance 61*(4), 1553–1604.

[60] Caselli, F., I. Coleman, and W. John (2006). The world technology frontier. *American Economic Review 96*(3), 499–522.

[61] Chen, K.-J. and C.-C. Lai (2015). On-the-job learning and news-driven business cycles. *Journal of Money, Credit and Banking 47*(2-3), 261–294.

[62] Colombo, E. and L. Stanca (2008). The impact of training on productivity : evidence from a large panel of firms the impact of training on productivity. Journal Article 134, University of Milan Bicocca.

[63] Connolly, H. and P. Gottschalk (2008). Wage cuts as investment in future wage growth. *Labour 22*(1), 1–22.

[64] Cotton, J. (1988). On the decomposition of wage differentials. *The Review of Economics and Statistics 70*(2), 236–243.

[65] Cunha, F. and J. Heckman (2007). The technology of skill formation. *American Economic Review 97*, 31–47.

[66] De Grip, A. and J. Sauermann (2011). The effects of training on own and co-worker productivity: evidence from a field experiment. *The Economic Journal 122*(560), 376–399.

[67] 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.

[68] De la Fuente, Á. and A. Ciccone (2003). Human capital in a global and knowledge-based economy. Report 562, Office for Official Publications of the European Communities.

[69] Dearden, L., H. Reed, 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*(4), 397–421.

[70] DeJong, D. N. and B. F. Ingram (2001). The cyclical behavior of skill acquisition. *Review of Economic Dynamics 4*(3), 536–561.

[71] Destré, G., L. Lévy-Garboua, and M. Sollogoub (2008). Learning from experience or learning from others?. inferring informal training from a human capital earnings function with matched employer-employee data. *Journal of Socio-Economics 37*(3), 919–938.

[72] Dostie, B. (2013). Estimating the returns to firm-sponsored on-the-job and classroom training. *Journal of Human Capital 7*(2), 161–189.

[73] Dostie, B. (2015). Do train-or-pay schemes really increase training levels? *Industrial Relations: A Journal of Economy and Society 54*(2), 240–255.

[74] Dustmann, C. and U. Schönberg (2012). What makes firm-based vocational training schemes successful? the role of commitment. *American Economic Journal: Applied Economics 4*(2), 36–61.

[75] Einarsson, T. and M. H. Marquis (1998). An RBC model with growth: the role of human capital. *Journal of Economics and Business 50*(5), 431–444.

[76] Eisenberger, R., R. Huntington, S. Hutchison, and D. Sowa (1986). Perceived organizational support. *Journal of Applied Psychology 71*(3), 500–507.

[77] Favilukis, J. (2013). Inequality, stock market participation, and the equity premium. *Journal of Financial Economics 107*(3), 740–759.

[78] Felstead, A., F. Green, and N. Jewson (2013). An analysis of the impact of the 2008-09 recession on the provision of training in the UK. *Work, Employment and Society 26*(6), 968–986.

[79] Felstead, A., F. Green, N. Jewson, and P. Casey (2013). Training in recession : the impact of the 2008-2009 recession on training at work. UKCES Report 72, UK Commission for employment and skills.

[80] Frazis, H. and M. A. Loewenstein (2005). Re-examining the returns to training. *Journal of Human Resources 40*(2), 453–476.

[81] Frazis, H. and M. A. Loewenstein (2006). Wage compression and the division of returns to productivity growth: evidence from EOPP. BLS Working Paper, Bureau of Labour Statistics.

[82] Galí, J., J. D. López-Salido, and J. Vallés (2004). Rule-of-thumb consumers and the design of interest rate rules. *Journal of Money, Credit and Banking 36*(4), 739–763.

[83] Galindo-Rueda, F. and J. Haskel (2005). Skills, workforce characteristics and firm-level productivity: evidence from the matched ABI/Employer Skills Survey. IZA Discussion Papers 1542, IZA Institute of Labor Economics.

[84] Garloff, A. and A. Kuckulenz (2006). Training, mobility, and wages: specific versus general human capital. *Jahrbücher für Nationalökonomie und Statistik 226*(1), 55–81.

[85] Gathmann, C. and U. Schönberg (2010). How general is human capital? a task-based approach. *Journal of Labor Economics 28*(1), 1–49.

[86] Gertler, M., C. Huckfeldt, and A. Trigari (2016). Unemployment fluctuations, match quality, and the wage cyclicality of new hires. NBER Working Paper Series 22341, National Bureau of Economic Research.

[87] Goldin, C. and L. F. Katz (2007). The race between education and technology: the evolution of us educational wage differentials, 1890 to 2005. NBER Working Papers 12984, National Bureau of Economic Research.

[88] Goldin, C. and L. F. Katz (2008). Transitions: career and family life cycles of the educational elite. *American Economic Review 98*(2), 363–369.

[89] Green, F., A. Felstead, D. Gallie, H. Inanc, , and N. Jewson (2013). What has been happening to the training of British workers? LLAKES Research Paper 43, Centre for Learning and Life Chances in Knowledge Economies and Societies.

[90] Green, F., A. Felstead, D. Gallie, H. Inanc, and N. Jewson (2015). The declining volume of workers' training in Britain. *British Journal of Industrial Relations 54*, 422–448.

[91] 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*(2), 223–263.

[92] Grund, C. and J. Martin (2012). Determinants of further training: evidence for Germany. *International Journal of Human Resource Management 23*(17), 3536–3558.

[93] Haelermans, C. and L. Borghans (2012). Wage effects of on-the-job training: a meta-analysis. *British Journal of Industrial Relations 50*(3), 502–528.

[94] Hammond, P. (1998). The efficiency theorems and market failure. In A. Kirman (Ed.), *Elements of General Equilibrium analysis*, pp. 211–260. Blackwell, Oxford.

[95] Hansen, G. D. and S. Imrohoroglu (2009). Business cycle fluctuations and the life cycle: how important is on-the-job skill accumulation? *Journal of Economic Theory 144*(6), 2293–2309.

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

[97] Hanushek, E., S. Machin, and L. Woessmann (2011). *Handbook of the Economics of Education* (1 ed.), Volume 3. Elsevier, Amsterdam.

[98] Hara, H. (2014). The impact of firm-provided training on productivity, wages, and transition to regular employment for workers in flexible arrangements. *Journal of the Japanese and International Economies 34*, 336–359.

[99] Harrison, R. and Ö. Oomen (2010). Evaluating and estimating a DSGE model for the United Kingdom. BOE Working Papers 380, Bank of England.

[100] Heckman, J. and P. Carneiro (2003). Human capital policy. NBER working papers 9495, National Bureau of Economic Research.

[101] Heckman, J. J. (1976). A life-cycle model of earnings, learning, and consumption. *Journal of Political Economy 84*(4), 9–44.

[102] Heckman, J. J., R. J. Lalonde, and J. A. Smith (1999). The economics and econometrics of active labor market programs. In O. C. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 3A, Chapter 31, pp. 1865–2097. Elsevier, Amsterdam.

[103] Hornstein, A., P. Krusell, and G. L. Violante (2005). The Effects of Technical Change on Labor Market Inequalities. In P. Aghion and S. Durlauf (Eds.), *Handbook of Economic Growth*, Volume 1, Chapter 20, pp. 1275–1370. Elsevier, Amsterdam.

[104] Huggett, M., G. Ventura, and A. Yaron (2006). Human capital and earnings distribution dynamics. *Journal of Monetary Economics 53*(2), 265–290.

[105] Ionescu, F. (2009). The federal student loan program: quantitative implications for college enrollment and default rates. *Review of Economic Dynamics 12*(1), 205–231.

[106] Jann, B. (2008). The blinder-oaxaca decomposition for linear regression models. *Stata Journal 8*(4), 453–479.

[107] Jones, F. L. and J. Kelley (1984). Decomposing differences between groups: a cautionary note on measuring discrimination. *Sociological Methods & Research 12*(3), 323–343.

[108] Judge, T. A., C. J. Thoresen, J. E. Bono, and G. K. Patton (2001). The job satisfaction-job performance relationship: a qualitative and quantitative review. *Psychological Bulletin 127*(3), 376–407.

[109] Kahn, L. B. (2013). Asymmetric information between employers. *American Economic Journal: Applied Economics 5*(4), 165–205.

[110] Kamihigashi, T. (2001). Necessity of transversality conditions for infinite horizon problems. *Econometrica 69*(4), 995–1012.

[111] Katz, E. and A. Ziderman (1990). Investment in general training: the role of information and labour mobility. *The Economic Journal 100*(403), 1147–1158.

[112] Kim, D. and C.-I. 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*(3), 494–518.

[113] Konings, J. and S. Vanormelingen (2015). The impact of training on productivity and wages: firm-level evidence. *Review of Economics and Statistics 97*(2), 485–497.

[114] Koster, F., A. De Grip, and D. Fouarge (2011). Does perceived support in employee development affect personnel turnover? *The International Journal of Human Resource Management 22*(11), 2403–2418.

[115] Krusell, P., L. E. Ohanian, J.-V. Ríos-Rull, and G. L. Violante (2000). Capital-skill complementarity and inequality: a macroeconomic analysis. *Econometrica 68*(5), 1029–1053.

[116] Kunze, A. (2005). The evolution of the gender wage gap. *Labour Economics 12*(1), 73–97.

[117] Kydland, F. E. and E. C. Prescott (1982). Time to build and aggregate fluctuations. *Econometrica 50*(6), 1345–1370.

[118] Lechthaler, W. (2005). Firm training and public policy. Discussion Paper 21, University of St. Gallen.

[119] Lechthaler, W. (2009). The interaction of firing costs and firm training. *Empirica 36*(3), 331–350.

[120] Leeper, E. M., T. B. Walker, and S. C. S. Yang (2010). Government investment and fiscal stimulus. *Journal of Monetary Economics 57*(8), 1000–1012.

[121] Leuven, E. (2004a). Explaining international differences in male skill wage differentials by differences in demand and supply of skill. *The Economic Journal 114*(1997), 466–486.

[122] Leuven, E. (2004b, Dec). A review of the wage returns to private sector training. In *EC-OECD Seminar on Human Capital and Labour Market Performance*, Brussels, pp. 1–26.

[123] Leuven, E. and H. Oosterbeek (2002). A new approach to estimate the wage returns to work-related training. IZA Discussion Papers 1, IZA Institute of Labor Economics.

[124] Leuven, E. and H. Oosterbeek (2004). Evaluating the effect of tax deductions on training. *Journal of Labor Economics 22*(2), 461–488.

[125] Leuven, E., H. Oosterbeek, R. Sloof, and C. van Klaveren (2005). Worker reciprocity and employer investment in training. *Economica 72*(285), 137–149.

[126] Liebeskind, J. P. (1997). Keeping organizational secrets: protective institutional mechanisms and their costs. *Industrial and Corporate Change 6*(3), 623–663.

[127] Lillard, L. A. and H. W. Tan (1986). Private sector training: who gets it and what are its effects? Technical Report 3331, RAND Corporation.

[128] Lindquist, M. J. (2004). Capital-skill complementarity and inequality over the business cycle. *Review of Economic Dynamics 7*(3), 519–540.

[129] Lusardi, A. and O. S. Mitchell (2013). The economic importance of financial literacy: theory and evidence. NBER Working Paper 18952, National Bureau of Economic Research.

[130] Lynch, L. M. (1992). Private sector training and its impact on the earnings of young workers. *American Economic Review 82*, 299–312.

[131] Machin, S. and A. Vignoles (2001). The economic benefits of training to the individual, the firm and the economy: the key issues. CEE Technical Report, Department for Education and Skills, Centre for the Economics of Education.

[132] Manuelli, R. E. and A. Seshadri (2014). Human capital and the wealth of nations. *American Economic Review 104*(9), 2736–2762.

[133] Markowitsch, J., B. Käpplinger, and G. Hefler (2013). Firm-provided training in Europe and the limits of national skills strategies. *European Journal of Education 48*(2), 281–291.

[134] Marx, M. (2011). The firm strikes back: non-compete agreements and the mobility of technical professionals. *American Sociological Review 76*(5), 695–712.

[135] Mattox, J. R. and D. L. Jinkerson (2005). Using survival analysis to demonstrate the effects of training on employee retention. *Evaluation and Program Planning 28*(4), 423–430.

[136] McCall, B., J. Smith, and C. Wunsch (2016). Government-sponsored vocational education for adults. In E. A. Hanushek, S. Machin, and L. Woessmann (Eds.), *Handbook of the Economics of Education*, Volume 5, Chapter 9, pp. 479–652. Elsevier, Amsterdam.

[137] Mejía, D. and M. St-Pierre (2008). Unequal opportunities and human capital formation. *Journal of Development Economics 86*(2), 395–413.

[138] Melero, E. (2010). Training and promotion: allocation of skills or incentives? *Industrial Relations 49*(4), 640–667.

[139] 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. *Contemporary Economic Policy 34*(1), 107–118.

[140] Merz, M. (1995). Search in the labor market and the real business cycle. *Journal of Monetary Economics 36*(2), 269–300.

[141] Merz, M. and E. Yashiv (2007). Labor and the market value of the firm. *American Economic Review 97*(4), 1419–1431.

[142] Metcalfe, R. and P. J. Sloane (2007). Human capital spillovers and economic performance in the workplace in 2004: some British evidence. IZA Discussion Papers 2774, IZA Institute of Labor Economics.

[143] Mincer, J. (1958). Investment in human capital and personal income distribution. *The journal of political economy 66*, 281–302.

[144] Mincer, J. and H. Ofek (1982). Interrupted work careers: depreciation and restoration of human capital. *The Journal of Human Resources 17*(1), 3–24.

[145] Mo, J. (2016). UK results from the Programme for International Student Assesment (PISA 2015). Technical report, Organisation for Economic Cooperation and Development.

[146] Moen, E. R. and Å. Rosén (2004). Does poaching distort training? *Review of Economic Studies 71*(4), 1143–1162.

[147] Moreno-Galbis, E. (2012). The impact of TFP growth on the unemployment rate: does on-the-job training matter? *European Economic Review 56*(8), 1692–1713.

[148] Moretti, E. (2004a). Human capital externalities in cities. In J. V. Henderson and J.-F. Thisse (Eds.), *Handbook of Regional and Urban Economics* (1 ed.), Volume 4, Chapter 51, pp. 2243 – 2291. Elsevier, Amsterdam.

[149] Moretti, E. (2004b). Workers' education, spillovers, and productivity: evidence from plant-level production functions. *American Economic Review 94*(3), 656–690.

[150] Mortensen, D. T. and C. A. Pissarides (1999). New developments in models of search in the labor market. In *Handbook of Labor Economics*, Volume 3, Chapter 39, pp. 2567 – 2627. Elsevier.

[151] Muehlemann, S., H. Pfeifer, G. Walden, F. Wenzelmann, and S. C. Wolter (2010). The financing of apprenticeship training in the light of labor market regulations. *Labour Economics 17*(5), 799–809.

[152] Muehlemann, S., J. Schweri, R. Winkelmann, and S. C. Wolter (2007). An empirical analysis of the decision to train apprentices. *Labour 21*(3), 419–441.

[153] OECD (1995). The OECD jobs strategy. Technical report, Organisation for Economic Co-operation and Development.

[154] O'Mahony, M. (2012). Human capital formation and continuous training: evidence for EU countries. *Review of Income and Wealth 58*(3), 531–549.

[155] O'Mahony, M. and R. Riley (2012). Human capital spillovers : the importance of training. INDICSER Discussion Papers 23, National Institute of Economic and Social Research and LLAKES.

[156] Parent, D. (1999). Wages and mobility: the impact of employer-provided training. *Journal of Labor Economics 17*(2), 298–317.

[157] Persson, T. and G. Tabellini (1992). The politics of 1992: fiscal policy and European integration. *The Review of Economic Studies 59*(4), 689–701.

[158] Piketty, T. and E. Saez (2007). How progressive is the US federal tax system? a historical and international perspective. *Journal of Economic Perspectives 21*(1), 3–24.

[159] Pischke, J.-S. (2001). Continuous training in Germany. *Journal of Population Economics 14*(3), 523–548.

[160] Pischke, J.-S. (2005). Comments on "workplace training in Europe" by Bassanini *et al.* (2005). In *fRDB conference "Education and Training in Europe"*, Venice, pp. 1–16. London School of Economics.

[161] Pourpourides, P. M. (2011). Implicit contracts and the cyclicality of the skill-premium. *Journal of Economic Dynamics and Control 35*(6), 963–979.

[162] Rawls, J. (1971). *A theory of justice.* Belknap Press of Harvard University Press, Cambridge.

[163] Riksdag, S. (1999). Directive 106: Förutsättningarna för att införa ett system för individuell kompetensutveckling. Committee Report, Sweedish Parliament.

[164] Rogerson, R. and R. Shimer (2011). Search in macroeconomic models of the labor market. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics* (1 ed.), Volume 4A, Chapter 7, pp. 619–700. Elsevier.

[165] Rosholm, M. and L. Skipper (2009). Is labour market training a curse for the unemployed? evidence from a social experiment. *Journal of Applied Econometrics 24*(2), 338–365.

[166] Rubinstein, Y. and Y. Weiss (2006). Post schooling wage growth: investment, search and learning. In E. Hanushek and F. Welch (Eds.), *Handbook of the Economics of Education*, Volume 1, Chapter 1, pp. 1–67. Elsevier.

[167] Smith, J. C. (2011). The ins and outs of UK unemployment. *The Economic Journal 121*(552), 402–444.

[168] Sofer, C. (2004). *Human capital over the life cycle.* Edward Elgar Publishing, Cheltenham.

[169] Sousounis, P. and R. Bladen-Hovell (2010). Persistence in the determination of work-related training participation: evidence from the BHPS, 1991-1997. *Economics of Education Review 29*(6), 1005–1015.

[170] Stephen, M. and A. Vignoles (2005). *What's the good of education? the economics of education in the UK.* Princeton University Press, Oxford.

[171] Stevens, M. (1994). A theoretical model of on-the-job training with imperfect competition. *Oxford Economic Papers 46*(4), 537–62.

[172] Stokey, N. L. and S. Rebelo (1995). Growth effects of flat-rate taxes. *Journal of Political Economy 103*(3), 519–550.

[173] Tomaskovic-Devey, D. and S. Skaggs (2002). Sex segregation, labor process organization, and gender earnings inequality. *American Journal of Sociology 108*(1), 102–128.

[174] UK Department for Business Innovation and Skills (2013). Continuing vocational training survey: CVTS 4. Technical Report 102, Department for Business, Innovation and Skills.

[175] Veum, J. R. (1995). Sources of training and their impact on wage. *Industrial and Labor Relations Review 48*(4), 812–826.

[176] Wegge, J., K.-H. Schmidt, C. Parkes, and R. Dick (2007). Taking a sickie: job satisfaction and job involvement as interactive predictors of absenteeism in a public organization. *Journal of Occupational and Organizational Psychology 80*(1), 77–89.

[177] Wolter, S. and B. Stiftung (2018). Apprenticeship training in England: a cost-effective model for firms? Technical report, Education Policy Institute.

[178] Wooldridge, J. M. (2001). *Econometric Analysis of Cross Section and Panel Data*, Volume 1. The MIT Press, Cambridge.

[179] Wooldridge, J. M. (2016). *Introductory econometrics: A modern approach.* Nelson Education.

[180] Wößmann, L. and G. Schütz (2008). Efficiency and equity of European education and training policies. *International Tax and Public Finance 15*(2), 199–230.

[181] Yoong, J. (2010). Financial illiteracy and stock market participation: evidence from the RAND American life panel. PRC Working Paper 2010-29, Pension Research Council, Philadelphia.