Wages, Experience and Training of Women over the Lifecycle
Richard Blundell, Monica Costa Dias, David A. Goll, and Costas Meghir
NBER Working Paper No. 25776
April 2019, Revised in August 2020
JEL No. H2,J16,J22,J24,J3,J31

ABSTRACT

We investigate the role of training in reducing the gender wage gap using the UK-BHPS. Based on a lifecycle model and using tax and welfare benefit reforms as a source of exogenous variation we evaluate the role of formal training and experience in defining the evolution of wages and employment careers, conditional on education. Training is potentially important in compensating for the effects of children, especially for women who left education after completing high school, but does not fundamentally change the wage gap resulting from labor market interruptions following child birth.

Richard Blundell
University College London
Department of Economics
Gower Street
London
England
r.blundell@ucl.ac.uk

Monica Costa Dias
Institute For Fiscal Studies
7, Ridgmount Street
London WC1E 7AE, UK
monica_d@ifs.org.uk

David A. Goll
University College London
Room G01-4
Drayton House
30 Gordon Street
London WC1H 0QB
United Kingdom
david.goll.11@ucl.ac.uk

Costas Meghir
Department of Economics
Yale University
37 Hillhouse Avenue
New Haven, CT 06511
and IZA
and also NBER
c.meghir@yale.edu

Wages, Experience and Training of Women over the Lifecycle*

Richard Blundell
University College London and Institute for Fiscal Studies

Monica Costa-Dias
Institute for Fiscal Studies and University of Porto

David Goll
University College London and Institute for Fiscal Studies

Costas Meghir
Yale University, NBER, IZA, CEPR and Institute for Fiscal Studies

June 2020

Abstract

We investigate the role of training in reducing the gender wage gap using the UK-BHPS. Based on a lifecycle model and using tax and welfare benefit reforms as a source of exogenous variation we evaluate the role of formal training and experience in defining the evolution of wages and employment careers, conditional on education. Training is potentially important in compensating for the effects of children, especially for women who left education after completing high school, but does not fundamentally change the wage gap resulting from labor market interruptions following child birth.

*We are very grateful to Michael Webb and William Elming for the many discussions and inputs at different stages of this project. We are also grateful to conference and seminar participants at DIW/WZB Berlin, Chicago Harris School, CREST(ENSAE) Paris, LISER (Luxembourg), University of Bristol, University of Essex, University of Sussex, University of York, Bank of Spain, Birkbeck and University College London for their many useful comments. Financial support from the Economic and Social Research Council through the ESRC Centre for the Microeconomic Analysis of Public Policy at IFS (reference ES/M010147/1) and from project grant ES/N015304/1, and from the European Union Horizon 2020 programme on Dynamics of Inequality Across the Life-Course (Dial, project file number 462-16-090) is gratefully acknowledged. Costas Meghir is grateful to the Cowles foundation and ISPS at Yale for financial support. The usual disclaimer applies.

1 Introduction

Women's careers are marked by interruptions related to childbirth and the resulting loss in labor market experience. This, together with the fact that women often work part time while children are growing up, underlies an increasing wage gap relative to men as well as to women who continue an uninterrupted career as full time workers. The question we address in this paper is whether work-related training has a role to play in reducing this wage gap and whether it can be used to help reintegrate women in the labor market following a long absence.

In this paper we specify a model of female labor supply over the lifecycle including the choice to obtain work-related training. In our model, women enter the labor market after completing education. In each period they face a working hours and savings choice. Marriage, separation and children arrive exogenously with a probability estimated from the data and depending on prior children, age and marital status. The evolving family structure over the lifecycle is a key feature because it affects the incentives and preferences of women for work and training. While working their human capital grows through experience, at a rate depending on whether work is part time or full time. Job separations imply a loss in human capital and hence earnings. During their working life they may also participate in work-related training, which is paid for by deductions from their earnings but increases human capital and therefore wages in future periods. While we recognise that part of the cost of training and part of the return may accrue to the firm, we do not explicitly model incidence. However, we do not impose that the worker enjoys the full return to training: we allow the data to determine the returns to training episodes for the worker based on wage data.

Our focus is on the two human capital enhancing activities, working and training. Each of these activities responds to incentives in a different way, which poses interesting policy questions. For example, passive learning in work is encouraged by any factor increasing the incentives to work, such as in-work benefits (EITC in the US, WFTC in the UK). By making working more desirable, these work conditioned policies may also mechanically increase the amount of active work-related training over the life-cycle. Perhaps more interestingly, by topping up low pay benefits can indirectly subsidise the cost of training associated with foregone earnings, see Heckman et al. (2002). The design of the subsidy may also interact with the return to training in ways that may increase or reduce its return. Understanding the importance of work-related training for human capital and wages is thus central to designing policy that could help reduce the earnings costs of children on women. In turn, this discussion also reveals that policy reforms changing incentives to work, and to work more, may also affect training rates. In such case, they can be used to identify the effects of training on future wages. We will exploit such variation together with our model to quantify these effects.

Our basic data source is the UK BHPS, a long panel running since 1991 with key labor market and household information. Importantly, it includes detailed information on the incidence and intensity of training. This information is similar to one of the first systematic analyses of work related training by Altonji and Spletzer (1991). We supplement this with information on welfare and tax systems in the UK over many years, which allows us to construct the precise budget constraint that an individual is facing in each year of work. This leads us to our identification strategy: our data includes multiple cohorts, entering the labor market at different times. Each is facing a different welfare and tax system implying changes in incentives. During their lifetimes they face reforms that affect a number of cohorts but at different ages. This generates exogenous variation in the incentives that people face at different parts of the distribution. Thus individuals of different cohorts and education groups face both different work and training incentives. This is the key idea that underlies our identification strategy and provides the variation we need to estimate the model.

Our findings point to a potentially important role for training women who completed high school level education but did not go on to complete University. We show that it can have a role in reducing the wage loss that arises from part time work post children. Moreover, policies that subsidize the training of recent mothers from this group can increase their disposable income (beyond the taxation required to fund it) as well as overall welfare. We also find that a modest subsidy pays for itself by incentivising full-time work both during the eligibility period and after it. Finally, while training can play some role in reducing the labor market costs of children, this cost remains quite large even after systematic training policies. Other policies that would reduce the incidence of part time work, such as better childcare availability, may have a more important role to play.

The paper proceeds as follows: In the next section we describe our data, followed by a description of the institutional framework. We then carry out an empirical analysis to investigate how incentives related to the tax and welfare system affect training. Having shown that training is indeed sensitive to such incentives we specify our model and describe our estimation approach, which uses the simulated method of moments. This section is followed by the description of the results including our counterfactual simulations. We then offer some concluding remarks.

2 Data

Estimation uses the 18 yearly waves of the British Household Panel Survey (BHPS), a longitudinal dataset following the lives of families and their offshoots from 1991 to 2008. The survey started with a representative sample of 5,050 households living in Great Britain; it was later replenished in 1997 and 2001 with 1,000 households from the former European Community Household Panel, and in 1999 with two samples of 1,500 households each from the Welsh and Scottish extensions.[1] Except for some attrition, all household members in the original samples remain in the sample until the end of the period. Other individuals have also been added to the sample, as they formed families with original members of the panel or were born into them.

The BHPS collects detailed demographic information that we use to characterise the dynamics of family formation, as well as socio-economic information mapping the education attainment, labour supply, earnings, training events, childcare expenditures and assets of all

[1] An additional sub-sample from Northern Ireland was added in 2001 but is not used here.

household members aged 16 and above. In 1992, 2001 and 2002, the BHPS contains an additional module on lifetime histories that we use to recover the employment history of adult respondents since they first started to work. Respondents also report retrospective information on family background, including measures of parental education, number of siblings, sibling order, whether they lived with parents when aged 16, books at home during childhood, etc. We synthesise this information into two indices of socio-economic background that will be used to qualify individual earnings capacity and choices.

Our observation unit is women who have completed education, are aged 19 to 60, and for whom we observe complete employment histories. The histories of women who return to full-time education to acquire additional qualifications are truncated. We also truncate the histories of those who become self-employed at any point during the sample period, from that moment onwards. Finally, we exclude women who are not UK citizens or who are ever observed claiming disability benefits. The records of women in the cleaned sample are then linked to information on a present partner and children as relevant.

Our final sample is an unbalanced panel of 7,359 women and 55,591 observations. We arrange them into three groups by highest level of completed education, corresponding to less than high-school, high-school qualifications and equivalent, and 3-year college degree and above.[2] Table 1 shows the sample composition by family type and education of the woman.

We consider both the extensive and intensive margins of labour supply and discretise the distribution of labor supply to 3 points: not working for pay, which we take to be 0 hours in paid work per week and corresponds empirically to the cases of workers doing less than 5 weekly hours of work; working part-time, which we take to be 20 hours of work per week and combines all those doing 5 to 20 hours; and full-time work, which we take to be 40 weekly hours and combines workers doing 21 or more hours per week. The underlying measure of weekly hours we use is for usual hours in main job, including paid and unpaid overtime. We also consider

[2] In the UK, these levels correspond, respectively, to GCSE qualifications (which are acquired at the end of secondary school, at age 16) and below, A-levels qualifications (obtained at the end of high-school, aged 18) and equivalent, and 3-year University degree and higher.

Table 1: Sample size and distribution of family types by education

	Education			
	Less than high school	**High school**	**University**	**Total**
Family type (%)				
Single, no kids	15.1	21.0	24.7	18.2
Couple, no kids	34.6	33.6	35.6	34.4
Single, with kids	11.1	7.9	4.6	9.2
Couple, with kids	39.2	37.5	35.1	38.1
Employment (%)				
Full-time (>20 hours)	53.2	68.9	77.3	61.2
Part-time (5-20 hours)	21.2	15.6	11.6	18.2
Nr of individuals	3,921	2,377	1,061	7,359
Nr of observations	30,802	17,419	7,370	55,591

Notes: BHPS data for the years 1991 to 2008.

only employees, and delete the paths of workers becoming self-employed, from that moment onwards. More details on data selection can be found in the Online Appendix.

Wages are measured on a per-hour rate by dividing weekly earnings in main job including paid overtime by weekly hours also in main job (including any overtime, as detailed above). Since our model does not deal with macroeconomic fluctuations, we net out aggregate wage growth from the wage rates and from all monetary values of the tax and benefit system, described below in Section 3. We also trim the wage rate distribution, on the 2nd and 98th percentiles, to limit the importance of measurement error in earnings and working hours.

Training Data. One distinctive feature of the BHPS is that it includes a detailed description of all work-related training taking place during the year prior to the interview among those currently employed. This measure of training is an umbrella to a wide variety of education activities meant to increase or improve skills in work and that can be pursued while working full or part time hours. It includes part-time college or university courses, evening classes, employer-provided courses either on or off the job, government training schemes, open university courses, correspondence courses and work experience schemes, but excludes full-time education. Work-related training amounts to over 80% of all recorded training episodes, of which 96% happen among those in paid work at the time of the interview. The data documents the purpose of the training (whether induction training in a new job, to gain skills for current job or to prepare for some new job in the future), its total duration, who paid for any direct costs, where it took place and whether it lead to any qualification.

Our measure of training is an indicator for whether the respondent has had strictly more than 40 hours of training over the previous year. In calculating the total time in training over the year, we have excluded instances of training for induction in a new job or where the participants report as it being unrelated to work. Specifically, we only consider training spells meant to increase the skills workers need in their current job (for example by learning a new technology), or to prepare for a new job; we exclude training meant to help workers getting started in their current job (induction training) or to develop skills generally (not work-related). We also exclude the 4% of cases where trainees are not working. For the remaining instances of training we first convert total duration – which can be reported in months, weeks, days or hours – into hours assuming 8 or 4 hours in a day for those in full-time or part-time hours, respectively. We then exclude all training episodes that result in 40 hours or less of training per week since they seem likely to capture minor work-based certification programmes, such as first-aid training.[3] Conditional on our selection, 76% of the training we account for leads to formal qualifications. This we take as suggestive evidence that the training considered here is human capital enhancing and transferable across jobs and firms.

Table 2 briefly describes training spells among women, by education. We show figures for our measure of training, labelled 'selected', and for a similar measure constructed on all work-

[3]In robustness checks, we have included induction related training and used a continuous training hours measure. The life-cycle patterns and our regression analysis shows (discussed below) are not qualitatively affected.

related training, labelled 'any'. Panel (a) of the Table shows that training is a common event, with between 17 and 37 percent of employed women receiving some form of training in each year. It is also much more common among those in the middle and top education groups. Our more demanding measure of training accounts for just over 40% of all training spells. These are non-negligible investments, with a median length of between 80 and 96 hours per year, or between 2 and 3 full-time weeks (panel (b)). In a working year of 48 weeks, the median training duration amounts to an average of about 2 hours of job-related training per week.

Panels (c) and (d) focus in Table 2 narrow the sample to include only trainees under our preferred definition. Women who have not completed high-school education are more likely to receive training at work (50%) than either High School educated women (36%) or University educated women (28%). University educated women are often trained at work, at private training centres. Around one-quarter of training occurs at a university or further education college, across all three education groups. When explicit fees are charged for training, these fees are paid by the employer in between 69 and 72 percent of instances. However, this measure does not account for additional costs of training, such as the loss of income that could result from fewer working hours.

3 Institutional background

The personal tax and welfare benefit systems operating in the UK during the 90s and 00s all consist of a small set of individual-based taxes and a larger set of benefits that are mostly means-tested on family income. Within the same structure, the period saw numerous reforms to the specific parameters determining entitlement to benefits and tax liabilities. The most significant was the sequence of reforms to the benefits of families with children that ocurred between Autumn 1999 and April 2002, which introduced the Working Families Tax Credit (WFTC) and changed the Income Support (IS) benefits for low-income families. We exploit these reforms in addition to other smaller changes in taxes and benefits to identify the returns

Table 2: Training descriptives for women by education (BHPS)

	Education			
	Less than High School	**High School**	**University**	**Total**
	***(a)* Training rates for employed (%)**			
Any training	17.1	33.4	37.0	27.4
Selected training	5.4	14.3	16.2	11.1
	***(b)* Median hours of training for trainees (hrs per year)**			
Any training	24	40	40	32
Selected training	80	96	88	88
	***(c)* Where did training take place (selected training, %)**			
At work	50.3	36.4	28.6	36.3
College/university	22.8	27.6	26.2	26.5
Other	26.9	35.9	45.2	37.2
	***(d)* Who paid explicit fees, if charged (selected training, %)**			
Fees paid by employer	69.3	71.0	71.5	70.9
No fees paid by employer	30.7	29.0	28.5	29.1

Notes: BHPS data for the years 1991 to 2008. All figures exclude instances of education or training spells that are not work-related. Training measured only for those in work at the time of the interview. 'Selected training' further excludes induction training and instances of training that add up to 40 or fewer hours of training in the course of one year.

to work experience and training and to study how welfare policy may affect training. We do so by modelling women and their families living through two tax and benefit systems that are representative of the main institutional features over the period of the data: that operating in April 1995, describing the policy environment of the 90s, and that finally implemented in April 2002, after the WFTC-IS reform was completed. Here we describe the main features of these systems; a more comprehensive discussion of the taxes and transfers in the UK can be found in Adam et al. (2010) and Blundell et al. (2016).

In terms of tax liabilities, the main instruments targeting families are the Income Tax and the National Insurance contributions. The basic structure of these taxes remained unaltered over the period. Income Tax is progressive, a step function over four income brackets. The 1995 system comprised of a personal income disregard that was not taxed, and rates 20% (starting), 25% (basic) and 40% (higher) that were gradually applied to additional fractions of personal income. The period saw a mild tax reduction, with a modest increase in the personal income disregard and some reduction of the rates to 10%, 22% and 40%. This was partly compensated by adjustments in the basic income threshold defining the brackets at which the starting and basic rates apply, and by a small increase in the main rate of National Insurance contributions, from 10% to 11%.

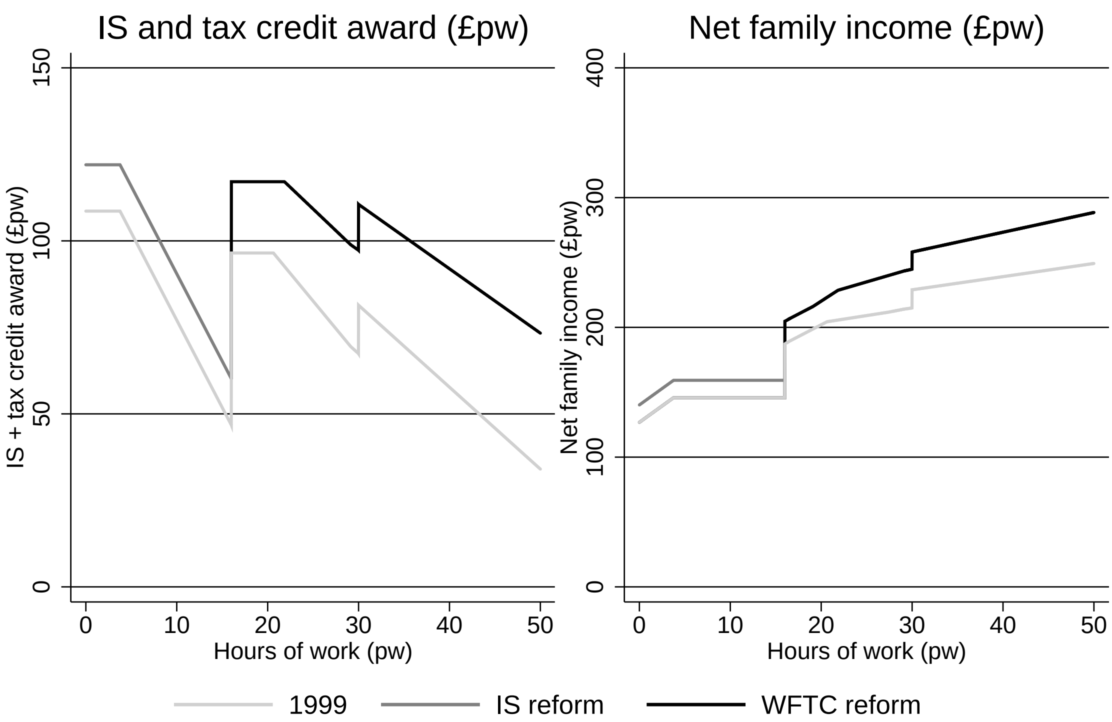
The UK benefit system is more complex. We model a range of benefits, including: Job-Seekers Allowance (JSA), which is the UK unemployment benefit; Income Support (IS), a minimum income floor that carries no work or job-search requirement; Tax Credits, a benefit for working families; Child Benefit, a universal benefit for families with children; Housing Benefit (HB), which subsidises housing costs for families who live in rented accommodation; and Council Tax Benefit, which subsidises the local property tax. These benefits interact in complex ways, so it is important to consider them together.

For mothers, the key components of the public transfer system are the IS and the Tax Credits. These were also the focus of the WFTC-IS reform of 1999-2002, an intervention aimed at improving the financial circumstances of low income families with children and keep mothers in work to protect their skills and labour market attachment. The reform implemented a significant increase in the generosity and coverage of IS and Tax Credits. For lone mothers, the IS award increased by over 10% relatively to wage levels over the period and remained taxed at 100% marginal rate. Since this subsidy is not work-contingent, this aspect of the reform reduced the incentives to work of mothers. The reform of the Tax Credit benefits, however, counteracted the increase in out-of-work benefits with a generous increase in subsidies for working mothers and an expansion of the target population to higher levels of family income. This was implemented by a 25% rise (in constant wage levels) in the maximum award for lone-mothers of one child, and a drop in the withdrawal rate from 70% to 55%. Over this period, Tax Credits kept the minimum working hours eligibility rule of 16 hours per week as well as the additional award for families working at or above the 30 hours threshold.

Figure 1 summarises the effects of these reforms on the take-home pay of single mothers. It shows, in 2008 prices and for a lone-mother on the minimum wage of April 2004, her entitlement (on the left) and disposable income (on the right) by working hours per week. The strong incentive to work part-time hours is clearly visible both before an after the reform. It is also apparent that the reform increased the incentive to work both part-time and more hours, by increasing the award at 16 hours by more than it increased out-of-work benefits and by reducing the rate at which in-work benefits are tapered away.

Figure 2 pictures the equivalent quantities for low-paid couples with one child aged 4 with one spouse working 40 hours per week at the 2004 minimum wage, by working hours of the second earner. Clearly, the reform had a much more modest effect on the disposable income of couples and, if anything, it reduced the incentives to work of the second earner in the family by taxing additional earned income more heavily.

Figure 1: Income Support and Tax Credit for minimum wage lone parent with 1 child



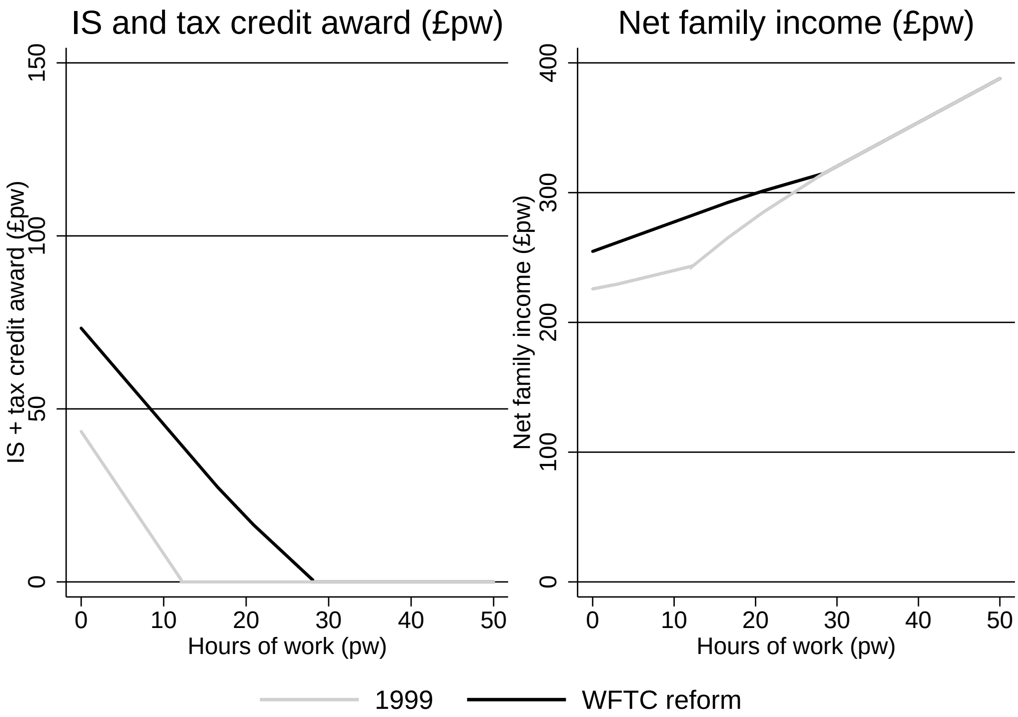
Notes: From Blundell et al. (2016). Simulations from FORTAX for lone-mother of one child aged 4, earning the 2004 minimum wage, not paying housing rents or childcare. Graph on the right pictures the IS plus TC award, graph of the left pictures the disposable income of the family; both in 2008 prices by working hours of the mother.

4 Life-cycle profiles of employment and training

The life-cycle patterns of wages, labour supply and training are suggestive of how these variables are linked for women, and of the motivations behind investments in training. Figure 3 shows the life-cycle profile of average log hourly wages of women and men, by education. The dashed lines for women exhibit the typical strong gradient by education and a steep upward profile early in the working life, particularly for high-school and university graduates. However, women's wages quickly flatten out during their late 20s or early 30s, coinciding with the main fertility period. The flattening is permanent after that.

The solid lines for men show wages increasing with education and growing rapidly in the

Figure 2: Income Support and Tax Credit for low-paid couple with 1 child

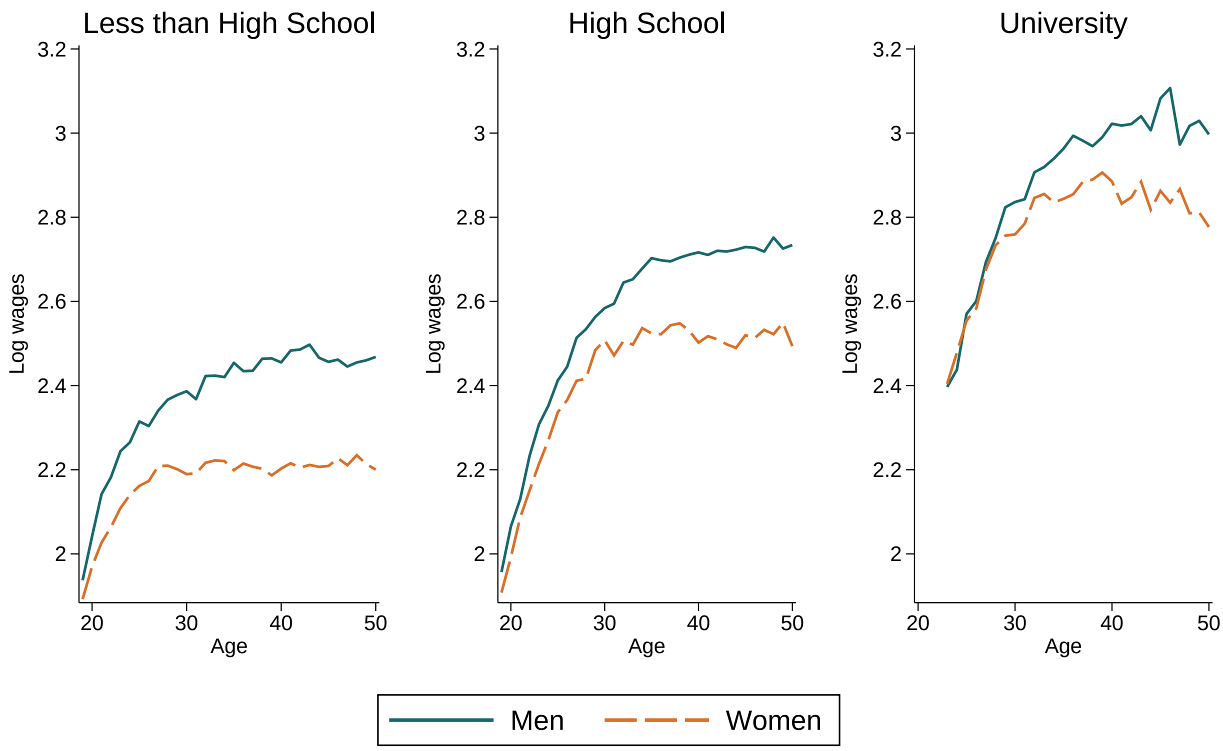


Notes: From Blundell et al. (2016). Simulations from FORTAX for couple of one child aged 4, not paying housing rents or childcare, both spouses earning the 2004 minimum wage, one spouse working 40 hours per week. Graph on the right pictures the IS plus TC award, graph of the left pictures the disposable income of the family; both in 2008 prices by working hours of the second earner.

early years of working life. However, the wages of men continue to grow far later into working life than the wages of similarly educated women, independently of education. The continued growth of men's wages compared to a flattening of women's wage profiles opens up a gender wage gap. For low educated women, this gap is already apparent by their early 20s. For higher educated women, the gap opens in their late 20s. These patterns coincide with differences across women by education in the timing of childbirth. For instance, 51% of women with less than high school qualifications in our sample have at least one child by age 23. This compares to 4% of University educated women. University educated women only reach comparable levels at age 32, where 50% of our sample have at least one child.

This wage profile is accompanied by strong changes in labour supply. Figure 4 shows, on

Figure 3: Average log wages of employed women and men over the life-cycle, by education

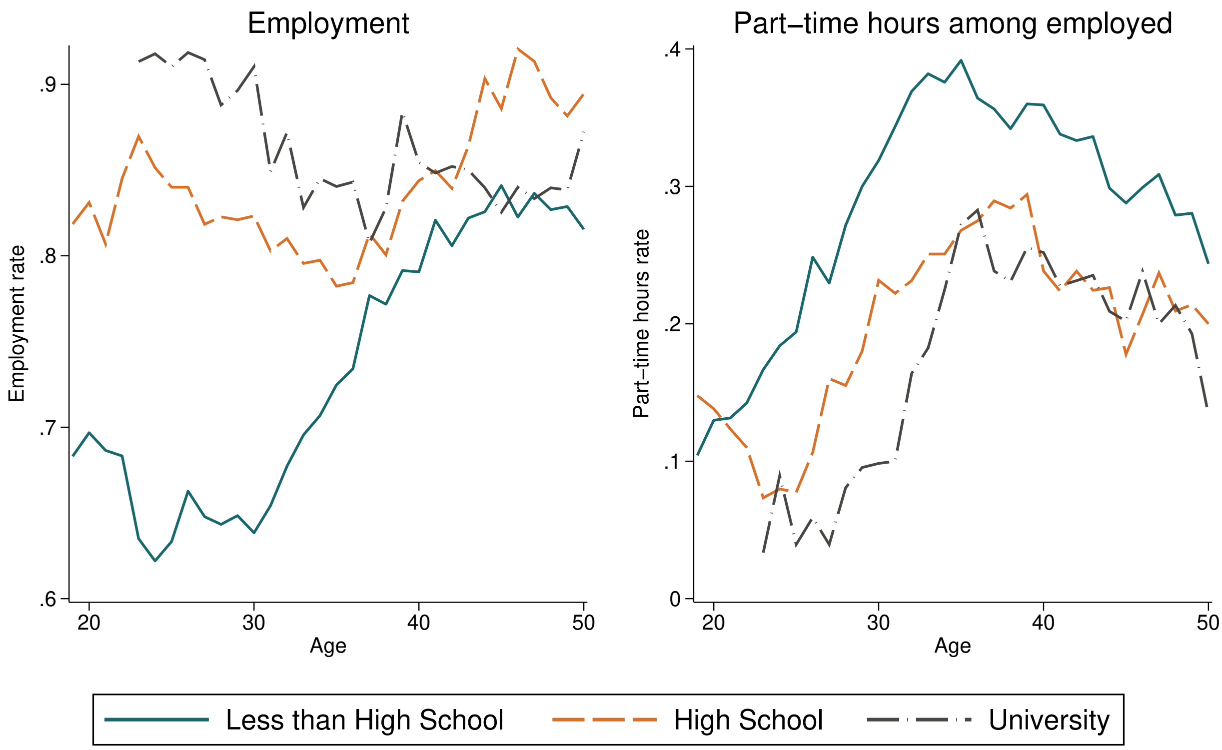


Notes: BHPS data for years 1991-2008. Real wages measured on a per-hour rate in logs.

the left, that the employment rates of women dip in the middle of their working lives. The dip happens earlier and is more pronounced for the lower educated. The right panel shows the proportion working part-time among women in work. The same period witnesses a strong growth in part-time hours that persists into late working life, particularly for those with high-school qualifications and less. Overall, employment and full-time working hours seem strongly complementary with education.

Blundell et al. (2016) documented these working patterns, related them to fertility episodes and quantified their consequences for the wage progression of women with different levels of completed education. What that paper did not consider, however, is how work-related training interacts with education, labour supply, work experience and wages. Here we see training as one

Figure 4: Employment and working hours over the life-cycle, by education



Notes: Notes: BHPS data for years 1991-2008. The graph on the left shows employment rates by age and education. The graph on the right shows the proportion of working women in part-time hours conditional on being in work, also by age and education.

element of human capital, together with education and work experience. Whether these three factors are complements or substitutes in the formation of wages will have consequences for the intensity and timing of training across different groups. For instance, if training can be used to offset human capital depreciation from non-working periods then it may be more prevalent among women returning to the labour market after a long fertility-related interruption than among men of similar age.

We start investigating this by contrasting the training patterns of women and men over the course of life in Figure 5. Panel A of this figure shows training rates by gender and education for all individuals, independently of work status (with training for those out of work always set to zero). Several features are noteworthy. First, on-the-job training is very common among High-

School and University graduates. There is a clear education gradient in training, with workers with less than high school qualifications being much less likely to invest. This suggests that, like work experience, the type of training that we measure is complementary with education instead of being used to compensate for the lack of academic skills.[4] Second, despite women being much more likely to interrupt their careers during the main child-rearing period, the training rates of women and men are surprisingly similar. This holds even at the start of working life, at which point women may foresee a long career interruption linked to fertility in the near future. Third, the overall pattern of training is downward slopping, as predicted by the classical Mincer/Ben-Porath human capital framework. Noticeably, however, the slope is not monotonic for women, particularly so for the more educated. Instead, training rates peak for a second time when women in these education groups are in their 40s or early 50s, a period that coincides with many of them returning to full-time work.

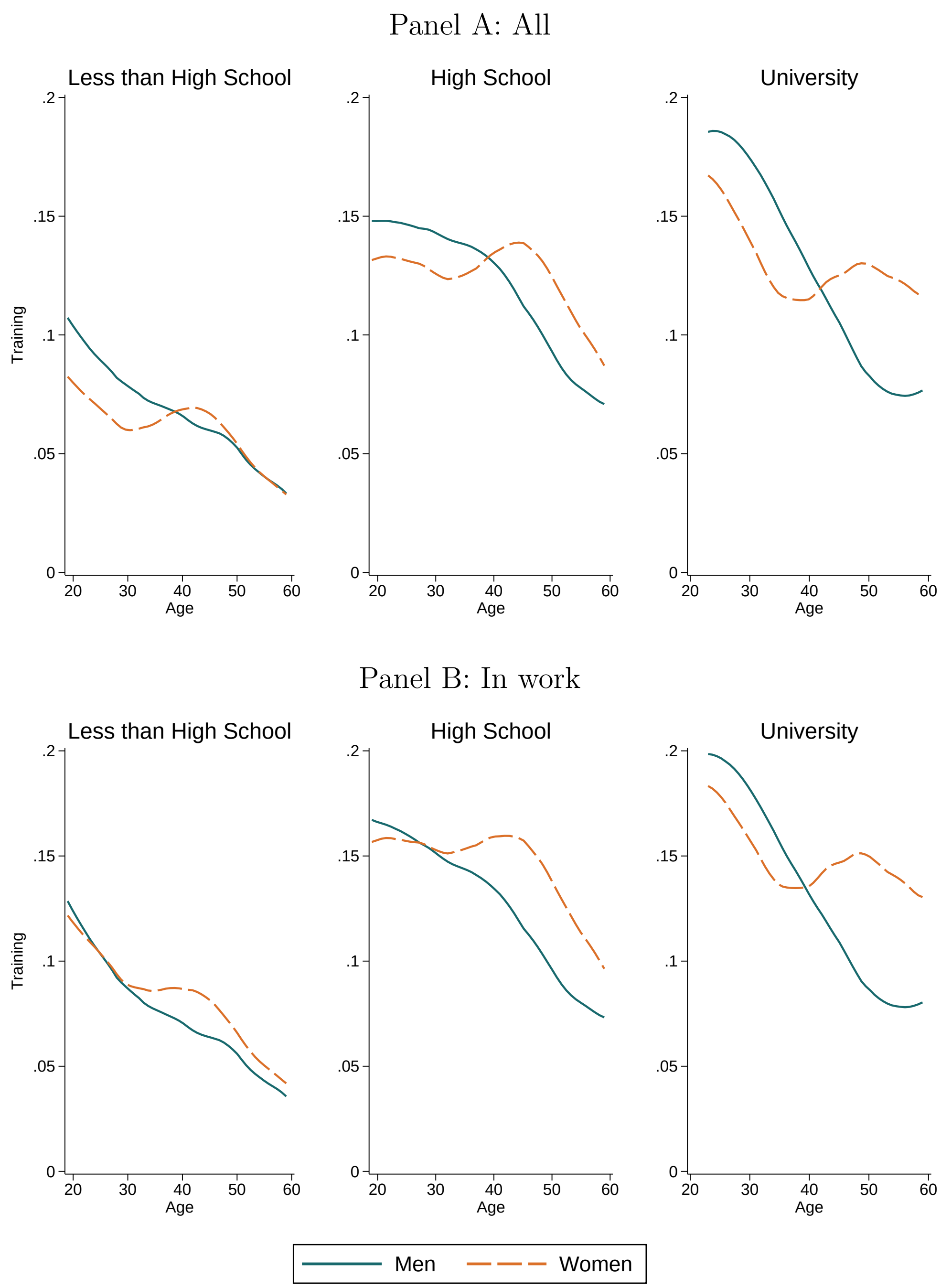
Conceivably, these patterns can be mechanically driven by the life-cycle of employment among women. Specifically, since female employment rates drop markedly during the main childrearing periods and recover once children are older, lower training rates at that stage and their subsequent pick up may just reflect that movement out and back into work. Panel B refutes that hypothesis by showing similar life-cycle variation in training rates among those in work.

Figure 6 provides further insight on the timing of training by plotting its frequency around the birth of the first child. It shows that the training rates are flat around the time of first birth for women with less than high school qualifications, seemingly unaffected by childbirth. In contrast, the training rates of women with high school or university qualifications vary significantly around chilbirth, first declining to reach a minimum while the child is very young and later partly recovering as the child moves to primary and secondary schools.

These patterns suggests a role for training in offsetting some of the losses in human capital

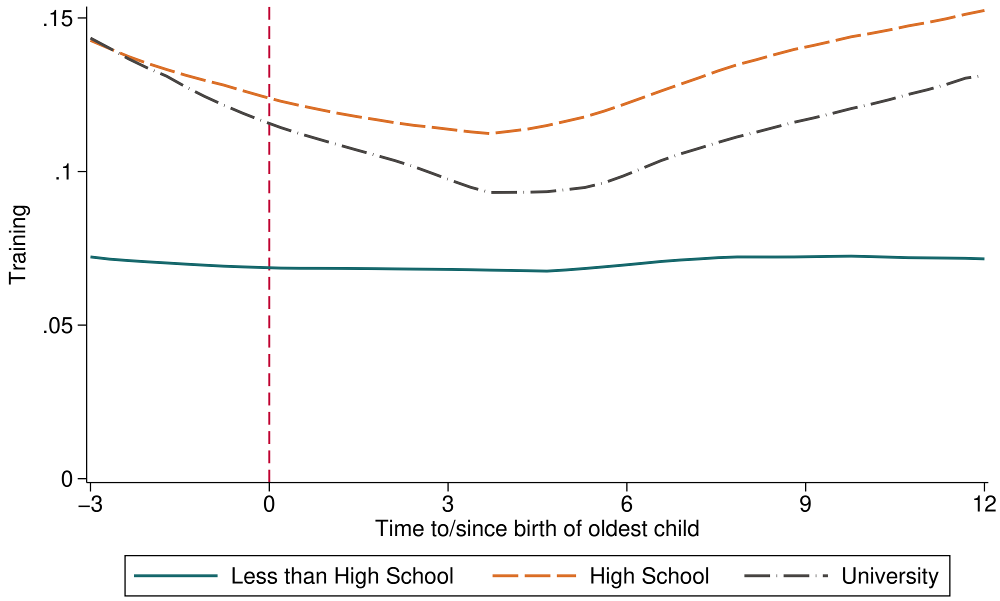
[4]One alternative explanation is that our measure favours training that is closer to the type that high educated people receive, and that other types of training (needed, for instance, for manual jobs) are not captured by our data.

Figure 5: Training rates over the life cycle, by gender and education



Notes: BHPS data for years 1991-2008. The training variable is an indicator for having had 40 or more hours of work-related training over the last 12 months. Panel A shows training rates for the entire population, by age, gender and education. Panel B additionally conditions on working at least 5 hours per week on an usual week, which is the measure of employment used in this paper. Lines are smoothed using a Epanechnikov kernel.

Figure 6: Training rates among mothers and mothers-to-be in paid work, by time to/since birth of first child and education



Notes: BHPS data for years 1991-2008. The training variable is an indicator for having had more than 40 hours of work-related training over the last 12 months.

and earnings capacity due to career interruptions, at least among mothers with High School qualifications or more. It is unlikely though that training alone will be enough to close the kind of gender differences in pay shown in figure 3. Even if the returns to training are similar to those from additional years of formal education, training spells are generally much shorter and so we would expect an effect that is proportionally adjusted. But training may, nevertheless, speed up gains in skills that women lose during working interruptions and make work more valuable for them.

The life-cycle patterns of training also suggest a role for public policies subsidising working mothers that has received little attention so far (one notable exception being Heckman et al. (2002). Specifically, working incentives targeting mothers – such as the UK Tax Credits that we described before or the US Earned Income Tax Credit – may have unforeseen effects on the take up of training through various channels. First, by making working more desirable they may mechanically increase the amount of training over the entire life-cycle. Second, by increasing the number of periods that women are in work, wage subsidies will also increase the number of periods over which women will reap the return from training, hence overall increasing the total return to the investment. Third, by topping up low pay, the benefits may indirectly subsidise the cost of training associated with foregone earnings. And finally, the design of the subsidy may interact with the return to training among subsidised women in ways that may increase or reduce its return.

5 Training responses to work incentives

One observation from the discussion in the previous sectin is that reforms in incentives to work may provide useful exogenous variation to identify the impact of training on the earnings of women. Existing studies have mostly focused on the impact of tax reforms on employment and hours. For instance, it has been shown that the WFTC reform affected the labour supply of lone-mothers (e.g. Brewer et al. (2006), Blundell et al. (2016)). Here we show that the various reforms to the tax and benefit system that happened in the UK over the 90s and 00s, of which the WFTC reform is a prominent example, also affected the probability that women take-up training.[5] This implies that tax and benefit variation can be used to help identify the returns to training in the context of a life-cycle model.

Our empirical specification is very simple. We estimate the following regression model of training T on a set of three simulated income variables that describe how working incentives change over time for different families, in response to policy changes:

$$T_{it} = \mathbf{1}\left[\gamma_0 + \gamma_1 \hat{Y}_{it}^{O} + \gamma_2 \hat{Y}_{it}^{P} + \gamma_3 \hat{Y}_{it}^{F} + \gamma_4' X_{it}' + \epsilon_{it}^{T} \geq 0\right] \quad (1)$$

In the above, the dependent variable T_{it} is an indicator for having had more than 40 hours

[5] We supplemented the variation in the monetary incentives to work with local variation in the availability of training captured by a Bartik instrument. We found that geographical variation to be too weak to drive training rates and dropped it.

of training over the last 12 months for woman i at time t, and $\left(\hat{Y}_{it}^{O}, \hat{Y}_{it}^{P}, \hat{Y}_{it}^{F}\right)$ are the respective simulated income variables. They measure family disposable income for three scenarios of female labour supply, respectively not working (superscript O), working part-time hours (respectively P) and working full-time hours (F). We use the tax system in place in period t to simulate these incomes based on predicted female wages (on her age and education) and details of the demographics of the family.[6] $\hat{Y}$ single out how policy reforms differentially affect the resources of families of different types depending on the labour supply of women. We also control for a set of other covariates X, which includes time dummies, a quadratic polynomial in age, indicators for family composition, two indices that summarise parsimoniously a set of observed variables characterising the socio-economic background of the woman.[7] These variables are meant to control for variation in the disposable income variables not induced by policy reforms.

Table 3 displays the results, focussing on the income variables. It shows that changes in incentives to work strongly affect the probability that women enrol in significant amounts of training. The F-statistics at the bottom of the table show that this is especially true for the two bottom education groups. This is not unexpected since public policies target the bottom of the income distribution and are, therefore, more effective in influencing choices at that margin.

Estimates in Table 3 are for all women, regardless of their employment status. Since the type of training that we are considering only happens among those in work, it could be thought that our estimates are effectively capturing the effects of monetary incentives to work on employment, and through employment on training. To check this possibility, we estimated the same regression model for the restricted sample of women in paid work. Results are shown in Table A1 in the Online Appendix. They demonstrate that this is not the case, particularly

[6] We use the IFS micro-simulation program Fortax, which provides a detailed description of the taxes and benefits operating at each time period.

[7] The indices are the first and second principle components of a set of observed retrospective variables on parental background, from when the woman was 16 years of age. They summarise information on the education of both parents (five levels each), number of siblings and sibling order (dummies for no siblings, three or more siblings, and whether respondent is the first child), books in childhood home (three levels) and whether lived with both parents when aged 16.

Table 3: Regression of training on simulated income

	(1) Less than High School	(2) High School	(3) Degree
Simulated income: no work ($\hat{Y}^O$)	-0.000254**	-0.000280	-0.000178
	(0.0000881)	(0.000145)	(0.000209)
Simulated income: part-time ($\hat{Y}^P$)	0.000606***	0.000524*	0.000751*
	(0.000146)	(0.000238)	(0.000376)
Simulated income: full-time ($\hat{Y}^F$)	-0.000705***	-0.000878***	-0.000960***
	(0.000105)	(0.000150)	(0.000223)
Observations	30383	17260	7328
Demographic Controls	Yes	Yes	Yes
Family Background Controls	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes
Age polynomial (2nd order)	Yes	Yes	Yes
F-Test on Instruments	20.93	15.53	8.312
F-Stat p-val	0.00	0.00	0.00

Notes: BHPS data for years 1991-2008. Outcome variable is an indicator for whether the woman has taken more than 40 hours of work-related training during the year that precedes the interview. Estimates show effects of simulated family disposable income for different levels of female labour supply on the probability of taking up training. The simulations are constructed using a detailed microsimulation model for the UK. We use the tax system in place in period t to simulate these incomes based on predicted female wages (on her age and education) and details of the demographics of the family. The regressions also control for year dummies, demographic characteristics (including a quadratic in age and dummies indicating family composition) and family background (including the first two principal components drawn from a collection of variables that describe the childhood household of each individual and an indicator for whether this information is missing). The F-statistics at the bottom of the table test the joint significance of the three simulated income variables. Standard errors, shown in parentheses under estimates, are clustered at the individual level.

for women in the middle education group. For them, the F-statistics that we estimate is still strong (at 8.8). The effect of the simulated income variation is weakest for college graduates (F-statistic of 6.4) and in between the two for the group of women with less than high school qualifications (7.3). Given the strength of the policy variation in affecting the training rates of high school graduates and the fact that training is very prevalent among women this group as well, our focus will be on this group for the remainder of the paper.

6 The model

We study training choices and their value for earnings through the lens of a life-cycle model of labour supply and human capital (HC) formation. Our model builds on the life-cycle model of female education, labour supply and experience capital of Blundell et al. (2016) by integrating on-the-job training in the process of HC formation and by adding a layer of heterogeneity that shapes the returns to HC investments. Here, however, we focus on a homogeneous education group.

6.1 Overview of the model and its key components

We consider the adult life of women, after completing education. Following our discussion of training incidence and training incentives, we focus on the key group of women who completed high school but did not complete a degree. Our model considers labour supply, training, consumption and savings choices of women from the moment they enter the working life at the age of 19. Adult life is split in two periods, the working period and the post-retirement period. Retirement is assumed to happen deterministically at the age of 60. Once retired, women stop working and live out of the savings they accumulated during working life (Fan et al. (2017)).

All women initiate their adult life as singles with no children. They are characterised by various dimensions of *ex-ante* permanent heterogeneity, some observed and others not. The observed heterogeneity is captured by two indices of family background, describing the socio-economic conditions of their parental home when they were aged 16. These affect their productivity in and preferences for work. The other component of observed heterogeneity is the cohort to which women belong. Different cohorts are affected by different sequences of work incentives shaped by the policy reforms, which may affect their working and training choices.

Ex-ante unobserved heterogeneity is two-dimensional. It includes one ability component, which directly affects wages, and one preference component, which drives the utility costs of working hours and training. We assumed that these two dimensions of heterogeneity are perfectly correlated. The structure of the unobserved heterogeneity terms is clearly specified below, when we set out preferences and wages.

During their working life, women decide in each period whether to work and for how many hours, whether to invest in training if they are working, and how much to consume today and save for the future. Labour supply is modelled in three hour-points, corresponding to not working, working part-time and full-time. Training is fixed at 2 hours per week, the median value of the distribution of training conditional on it exceeding 40 hours over the previous year, or 1 full-time working week worth of training.

Working has present and future returns, in the form of earnings and experience capital respectively. Earnings are proportional to the number of working hours net of time in training, with an hourly wage rate that depends on the stock of human capital, the woman's ability type and a persistent productivity shock. Human capital is represented by a single index, and is endogenous in our model. It accumulates over the life cycle through working experience and training episodes; it depreciates during out-of-work periods, formalising the idea that career interruptions carry long-term consequences for earnings capacity.

In a competitive labour market framework with general training, workers bear the full cost of training and capture its entire return. However, firm-specific training and labour market frictions may change this result, instead creating the grounds for firms and workers to share the costs and returns from training (Acemoglu and Pischke (1999), Lentz and Roys (2015)). In our model, we do not explicitly consider the role of firms and the labour market in determining how the cost and return to the investment is shared between workers. We assume that training carries a monetary cost equal to foregone earnings due to time taken away from work, and that it bears a return through HC that is reflected in future wages. However, we also allow training to carry a utility cost that may partly capture, in a reduced-form sense, incidence in the cost of training. It also captures other drivers of training, such as actual preferences, effort or congestion in training places. In the same vein, the contribution of training to the HC index also has a reduced-form interpretation. It represents a combination of its effect on the accumulation of skills and the sharing of their productive value with the firm. Training may also contribute to employer learning about productivity as in Altonji and Pierret (2001). They conclude that training has a mixed role, both as enhancing human capital and compensating for the depreciation of skills acquired in formal education, but also as a mechanism that supports employer learning. However, the nature of the data does not allow them to estimate the relative importance of these factors.

In our framework we give a pure human capital interpretation to the effects of training. Investments in training are driven by various mechanisms that also determine their timing and return. Crucially, if wages are concave in HC then the monetary cost of training is lower and returns are larger when HC is low. This creates stronger incentives to invest at the start of the working life – when there is also a longer period ahead to bear returns, as in a Ben-Porath model – and when returning to work after long separations, to compensate for the depreciation of skills.

Other key components of the model also create rich interactions with employment and training choices and their returns. One is the stochastic process of family formation and dissolution, which maps out the formation and dissolution of couples and fertility episodes. The model reproduces the empirical marital sorting patterns and fertility histories of women whose highest education qualification is high school (Chiappori et al. (2009), Chiappori et al. (2018)).

Finally, choices of consumption are restricted by liquidity constraints. The family budget is determined not only by the earnings of the woman but also by those of a present partner, tax liabilities and public transfers. In particular, the model embeds a detailed description of the personal taxes and benefits operating in the UK and how they change over the sample period. This is implemented using the micro-simulation tool FORTAX (Shaw (2011)).

6.2 Female wages and human capital

We consider the problem of a woman aged t and, for simplicity of notation, omit the individual index. If working, this woman draws a per-hour wage that depends on the human capital she accumulated so far (κ), indicators for whether the family background factors are above or below their median in the population (x_1, x_2), permanent ability type ω, and an idiosyncratic persistent productivity shock ν. The latter follows an AR(1) process with normal innovations ζ and initial value drawn from a normal distribution. Formally, the wage equation is

$$\begin{aligned} \ln w_t &= b_0 + b_1 x_1 + b_2 x_2 + (\gamma_0 + \gamma_1 x_1 + \gamma_2 x_2) \ln(\kappa_t + 1) + \omega + \upsilon_t \\ \text{where } \upsilon_t &= \rho \upsilon_{t-1} + \zeta_t \end{aligned} \tag{2}$$

We allow for classical measurement error in wages by defining observed wages w^m as follows

$$\ln w_t^m = \ln w_t + \xi_t \quad \text{where} \quad \xi_t \sim \text{iid}.$$

Gross pay y depends on workings hours h. Women can choose to work either 0 hours, 18 hours or 38 hours, representing out-of-work, part-time and full-time hours respectively. Total working time also depends on whether the woman takes time to train as follows

$$y_t = w_t(h_t - d_t \bar{h}_d) \tag{3}$$

where d is an indicator for training and $\bar{h}_d$ is training time, which is exogenously set to 2 hours per week.

Human capital κ is accumulated in work, at a rate that depends on working hours and training status, and depreciates at a constant rate δ per period. The human capital process is

$$\begin{aligned} \kappa_{t+1} &= \kappa_t (1 - \delta) + g_1(h_t) + g_2(h_t) k_t + \tau_1 d_t + \tau_2 d_t k_t \\ \kappa_{\underline{t}} &= 0 \end{aligned} \tag{4}$$

g_1 and g_2 define how human capital accumulates with work. We allow for the human capital gains from work to depend on the number of working hours and to vary linearly with human capital accumulated so far. Both g_1 and g_2 are set to 0 if the woman is not working, and g_1 is also set to 1 if she works full-time; other values are estimated. τ_1 and τ_2 measure the human capital return to training, which we also allow to vary linearly with the stock of human capital. The woman starts her working life at time $\underline{t}$ with an initial stock of human capital equal to zero.

Our model of wages and human capital formation implies that training is both cheaper and draws larger returns (if, as expected, $\gamma_0 + \gamma_1 x_1 + \gamma_2 x_2 < 1$ and $\tau_2 \leq 0$) when human capital is low. This reinforces the incentive to invest young in order to bear the returns for longer. It also makes training investments more valuable after the long career interruptions common among mothers of young children, if these interruptions carry a significant loss of skills that would be implied by a large depreciation rate δ.

The wage equation also exhibits complementarity between human capital and ability, implying that high ability workers have more to gain from training activities that enhance human capital. But since high ability workers also pay a higher cost in terms of foregone earnings, the overall effect of ability on training take-up is ambiguous.

6.3 The employment and earnings of the spouse

Let $m_t = 0, 1$ be an indicator for the presence of a partner at time t. We denote his characteristics and outcomes by adding a 'tilde' to his variables. Although his labour supply choices and human capital process are not endogenously modelled, we adopt a stochastic specification that captures the main features of the richer female model.

The spouse at time t is characterised by his education $\tilde{s}_t$ and his productivity level $\tilde{\upsilon}$. The distribution of his education reproduces that observed empirically among spouses of high-school graduated women. To limit the size of the state space, his age is assumed to equal that of the woman, t. If working, his wage rate is

$$\begin{aligned} \ln \tilde{w}_t &= \tilde{b}_{\tilde{s}} + \tilde{\gamma}_{\tilde{s}} \ln(t-18) + \widetilde{\upsilon}_t & (5) \\ \text{where } \tilde{\upsilon}_t &= \tilde{\rho}_{\tilde{s}} \tilde{\upsilon}_{t-1} + \tilde{\zeta}_t. & (6) \end{aligned}$$

$\widetilde{\upsilon}$ is the productivity shock, initially drawn from a $\tilde{s}_t$-specific normal distribution when the couple is formed and later modelled as a $\tilde{s}$-specific auto-regressive process with normal iid innovations $\tilde{\zeta}$. As for women, we interpret transitory wage shocks as measurement error and specify the observed wages of the spouse as

$$\ln \tilde{w}_t^m = \tilde{w}_t^m + \tilde{\xi}_t \quad \text{where } \tilde{\xi} \sim \text{iid}.$$

In line with the empirical evidence, we consider only two labour supply points for men in couples: they are either not working, in which case their working hours $\tilde{h}$ are set to zero, or working full-time hours, with $\tilde{h} = 40$. Their employment process is

$$\begin{aligned} &\text{In new couples:} && \text{Prob}\left[\tilde{h}_t = 40 \mid t, \tilde{s}_t, m_{t-1} = 0\right] = \psi_0(t, \tilde{s}_t) & (7) \\ &\text{In existing couples:} && \text{Prob}\left[\tilde{h}_t = 40 \mid t, \tilde{s}_t, \tilde{h}_{t-1}, m_{t-1} = 1\right] = \psi_1\left(t, \tilde{s}_t, \tilde{h}_{t-1}\right) & (8) \end{aligned}$$

6.4 The budget constraint

Family resources include both the earnings of the woman, those of a present partner and net public transfers. Let a_t represent the stock of assets that the family brings into period t. Each period choices are limited by a liquidity constraint ruling-out borrowing. The budget constraint is formalised in terms of the evolution of assets:

$$\begin{aligned} a_{t+1} &= (1+r)a_t + y_t + m_t \tilde{h}_t \tilde{w}_t - T(w_t, h_t, X_t) \\ a_{t+1} &\geq 0 \quad \text{and} \quad a_{\underline{t}} = 0 \quad \text{and} \quad a_{\bar{t}+1} = 0 \end{aligned} \qquad (9)$$

In the above expression, r is the risk-free interest rate, $\underline{t}$ is the start of working life, and $\bar{t}$ is the last period of life, set at 10 years after the retirement age of 60. We assume that women enter their working life with no assets, which is consistent with empirical evidence, and that any remaining assets have no value after $\bar{t}$.

T is the tax and benefit function. It depends on the wage rate of the woman, her working hours (because the UK tax credits have an hours rule) and on all other state variables characterising the demographic and financial circumstances of the family, summarised in X. In particular, X includes presence of children and age of the youngest child, marital status, whether present partner is working and his wage rate. We use the detailed microsimulation tool, Fortax, to calculate T.[8]

6.5 The dynamics of family formation

We adopt a flexible Markov model to capture the dynamics of fertility, marriage and divorce. To preserve computational tractability while representing the key drivers of female labour supply, we only keep track of the age of the youngest child but allow for multiple fertility events. Let t^k denote the age of the youngest child in the family. Childbirth is represented by re-setting t^k to zero and happens at a rate that depends on the woman's age, whether she has other children (denoted by the indicator n^k) and the age of the youngest, and whether she is married (m)

$$\text{Prob}\left[t^k = 0 \mid t, n^k_{t-1}, t^k_{t-1}, m_{t-1}\right] \tag{10}$$

It is assumed that a child lives with her parents until turning 19, at which point she deterministically leaves her parents' home.

The probability that a woman marries or remains married to a man of education $\tilde{s}$ depends

[8] Fortax describes most of the UK personal taxes and benefits and how they changed over the period we model, including income tax, social security contributions, and the main subsidies for working-age families, namely income support, job-seekers allowance, tax credits, housing benefit, council tax benefit, child benefit.

on her past marital circumstances, her age, whether she has children, and the education of her spouse if he is present in the previous period,

$$\text{if single at } t-1: \qquad \text{Prob}\left[m_t = 1, \tilde{s} \mid t, m_{t-1} = 0, n^k_{t-1}\right] \tag{11}$$

$$\text{if married to man } \tilde{s} \text{ at } t-1: \qquad \text{Prob}\left[m_t = 1, \tilde{s} \mid t, m_{t-1} = 1, \tilde{s}, n^k_{t-1}\right] \tag{12}$$

Otherwise she will be single at time t.

6.6 Utility and value functions

In each period t of her working life, the woman decides about total family consumption (c), savings (a), her own labour supply and training investments to maximise her lifetime utility. Working life starts at $\underline{t} = 19$ for our sample of High School graduates. It ends deterministically at 60 when the woman retires, after which family savings fund an additional 10 years of consumption.

We assume intertemporal separability in preferences. The per-period utility of her choices depends on her preference type, θ, and a subset of the state variables X_t that characterise her circumstances at age t:

$$u\left(c_t, h_t, d_t; \, \theta, X_t\right) = \frac{\left(c_t/n_t\right)^{\mu}}{\mu} \exp\left\{U\left(h_t, d_t, \theta, X_t\right)\right\}. \tag{13}$$

In the above expression, n is the equivalence scale, factoring in family size,[9] and μ is the parameter determining both the degree of risk aversion and the elasticity of intertemporal substitution.

The function U reflects how the value of additional consumption varies with working hours and training status by family composition for women of different θ types. We decompose it into two additive terms, one relating only to working hours, U_h, and the other driving the utility

[9] $n = 1$ for singles, 1.6 for couples 1.4 for mother with child and 2 for a couple with children.

cost of training, U_T:

$$U(h,d,\theta,X) = U_h(\theta,X_1) + d \times U_T(h,\theta,X_2) \tag{14}$$

with (U_h, U_T) defined as follows

$$U_h(X_1) = \begin{cases} 0 & \text{for } h=0 \\ l_h(\theta) + \alpha_h X_1 & \text{for } h = 18, 38 \end{cases} \tag{15}$$

$$U_T(h,\theta,X_2) = l_T(\theta) + \alpha_T X_2 + \alpha_{T,h}. \tag{16}$$

In the above, we denote by X_1 and X_2 the two relevant subsets of state variables (not mutually exclusive) that directly affect preferences for working hours and training, respectively, and by (α_h, α_T) their associated parameters. X_1 includes a full set of interactions between marital status and whether she is a mother, indicators for age of youngest child in bands (0-2, 3-5, 6-10), and the background factors (x_1, x_2). X_2 includes indicators for whether or not she is a mother and age of youngest child in bands. Equation (16) also includes an interaction term between working hours and training status $(\alpha_{T,h})$. Heterogeneity in preferences θ takes two values, for low and high preferences for work, and is assumed perfectly correlated with heterogeneity in ability ω. The terms $(l_h(\theta), l_T(\theta))$ measure the importance of unobserved preferences for work and training in driving choices.

The intertemporal problem of the woman can now be formalised. Let β be the discount factor. Her problem in period t of her working life is

$$V_t(\omega,\theta,X_t) = \max_{(a_\tau,c_\tau,h_\tau,d_\tau)_{\tau=t,\dots,\bar{t}}} E_t\left[\sum_{\tau=t}^{\bar{t}} \beta^{\tau-t} u(c_\tau,h_\tau,d_\tau;\,\omega,\theta,X_\tau) + \beta^{\bar{t}-t} b(\kappa_{\bar{t}}) \middle|\, \omega,\theta,X_t\right] \tag{17}$$

The term $b(\kappa_{\bar{t}})$ represents the value of human capital at retirement. It is meant to capture the fact that human capital will have some value post age 59, both because some women will remain active in work and because human capital is valuable outside work as well. This values is specified as follows,

$$b(\kappa_{\bar{t}}) = \phi_1 \frac{(\phi_2 + \kappa_{\bar{t}})^{\mu}}{\mu}$$

The maximisation problem in 17 is conditioned by the budget constraint (9), the female wage and human capital processes (2)-(4), the dynamics of employment and wages of a present partner (5)-(8) and the dynamics of family formation (10)-(12). The woman starts her working life as a single woman with no children.

7 Estimation

We estimate the subset of model parameters driving female wages, human capital formation and preferences for working hours and training using the method of simulated moments. The values for all other parameters are taken from Blundell et al. (2016). These include the subset of parameters defining the pre-determined family dynamics, male employment and male wages. A description of their estimation procedure and the full set of estimates can be found in their Web Appendix B. Three other parameters are set at typical values in the literature: the parameter regulating the curvature of the utility function μ is set at -0.56, implying a risk aversion coefficient of 1.56; the risk-free interest rate r is set at 0.015 and the discount factor β at 0.98, together implying that agents are mildly impatient (Blundell et al. (1994), Attanasio and Weber (1995), Attanasio et al. (2008)).

Estimation relies on a set of 139 moments capturing various aspects of lifecycle behavior and wages.[10] We construct the simulated moments to reproduce their data counterparts, based on

[10]The moments include full- and part-time employment and training rates by age, family demographics, socio-economic background, and interactions between calendar time and demographics; employment and hours transition rates by family demographics and past wages; the mean, variance and percentiles of the wage distribution over the course of life and at entrance into working life; the correlation between wages and socio-economic background, years of work, working hours, training and past wages; the growth rate of wages by past working hours, training and socio-economic background.

the simulation of 5 lifetime profiles for each of the 1,443 high school educated women who are observed in BHPS with observed socio-economic background and life histories of employment. From the resulting 7,215 profiles we select a window that exactly matches the observation window of the corresponding woman in the survey data. This way, we exactly reproduce the time, age and socio-economic structure of the data.

Our estimation procedure uses the exogenous variation in the labour supply and training incentives from policy reforms. Using regression analysis, we showed in section 5 that such exogenous variation was important for high school graduates and may play an important role in driving the results for them (Andrews et al. (2017)).

Within the model we use the policy variation by considering four tax and benefit systems, namely the ones operating in April 1995, 1999, 2002 and 2004. The reforms are unannounced.

Our moments include pre- and post-2002 measures of employment, working hours and training that explicitly capture the variation induced by the reform. Responses to the reform are likely to vary by cohort, as they are differently exposed to the reform, and individual permanent characteristics. We exploit these interactions to identify the value of working and training for future wages, by explicitly modeling the differential exposure to the reforms of different cohorts and by allowing responses to depend on socio-economic background.

The estimates of the model parameters are the set of parameter values Θ that minimise the following expression

$$\sum_{\kappa=1,\ldots,K} \frac{\left(M^d_{\kappa,N} - M^s_{\kappa,S}\left(\Theta\right)\right)^2}{\mathrm{Var}\left(M^d_{\kappa,N}\right)} \tag{18}$$

where K is the total number of moments used in estimation, $M^d_{\kappa,N}$ is the estimate of moment κ from N observations of observed data and $M^s_{\kappa,S}$ is the corresponding moment calculated on S model simulations for parameter values Θ.[11] We calculate asymptotic standard errors following

[11]It is implicit in the maximisation criterion that we are not using the optimal asymptotic weighting matrix, following the suggestion of Altonji and Segal (1996). Instead, we use the diagonal matrix of inverse variances of the moments, which are bootstrapped using 1,000 replications.

Gourieroux et al. (1993).

8 Parameter estimates and implications for behaviour

8.1 Wages, human capital and the return to training

Table 4 shows estimates of the female wage process. Estimates in panel A of the Table are for the wage rates at the start of working life (b_0) and the return to human capital (γ_0). Socio-economic background has a relatively small (but statistically significant at conventional levels) effect on starting wages; in turn, the return to human capital does not vary significantly with socio-economic background. Our estimate of the return to human capital in wages (γ) is, as expected, smaller than 1. Combined with the log linear specification of the wage equation, this implies that the return to one additional unit of human capital decrease with the stock already accumulated. Finally, unobserved heterogeneity in the wage rates (ω) is important (see estimates in Panel B of the Table). Our estimates indicate that being high ability raises the wage rate by 24 log points compared to the average.

Uncertainty in wages is characterised by the persistent unobserved productivity process ν. Our estimates in Panel C suggest that though this process is highly persistent, with autocorrelation coefficients of around 0.95, there is a high level of wage uncertainty. There is also substantial heterogeneity in initial wages.

Training affects wages through its impact on human capital. Our estimates show the incremental effect of training over work experience for the duration of training; i.e. they show how much more human capital workers gain if they choose to take time away from working and use it to train instead. The top row of Table 5 shows the estimate of this effect for women at the start of working life, when they have not yet accumulated human capital from work (τ_1). Our estimate suggests that, at that stage of the working life, training increases human capital by 16 percent of the return to one year of full-time work (which is normalised to 1). We allow for more flexibility in how training affects human capital, and hence wages, by adding

Table 4: Wage parameters

	Parameter Value	St. Error
Panel A: Wage coefficients		
Intercept, $\exp(b_0)$	6.86	(0.065)
increment: high factor 1, $\exp(b_1)$	0.64	(0.093)
increment: high factor 2, $\exp(b_2)$	-0.31	(0.028)
Return to human capital, γ_0	0.27	(0.004)
increment: high factor 1, γ_1	-0.04	(0.005)
increment: high factor 2, γ_2	0.03	(0.004)
Panel B: Unobserved heterogeneity in ability, ω		
ω type I: wage effect	0.24	(0.012)
ω type I: probability	0.79	(0.002)
Panel C: Distribution of persistent productivity shock ν		
Persistence of productivity, ρ	0.95	(0.002)
St. dev. of productivity innovation, ζ_t	0.12	(0.003)
St. dev. of initial productivity, υ_0	0.27	(0.007)

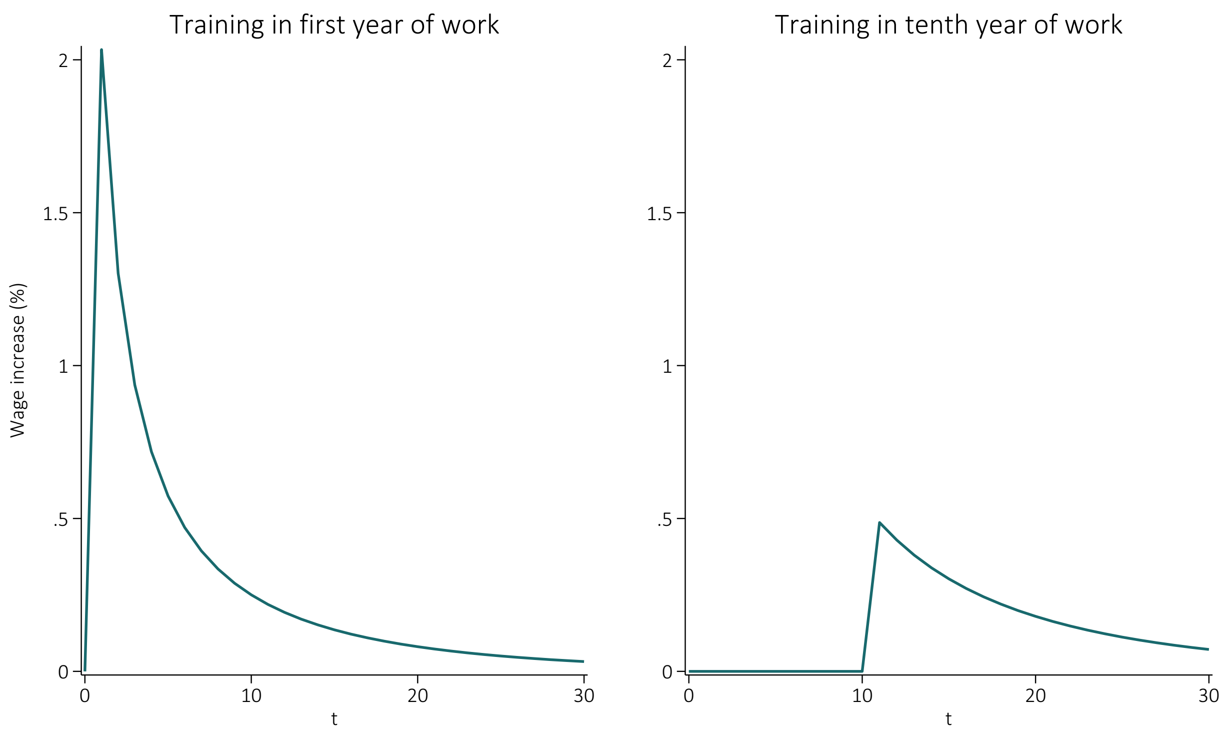
an interaction term with the stock of human capital (τ_2 in the second row of the table). Our estimates, however, suggest that this term is not needed.

Table 5: Parameters in the human capital accumulation process

	Parameter Value	St. Error
training, τ_1	0.16	(0.008)
training $\times$ human capital, τ_2	0.00	(0.004)
part-time, $g_1(18)$	0.13	(0.009)
part-time $\times$ human capital, $g_2(18)$	0.00	(0.003)
full-time $\times$ human capital, $g_2(38)$	-0.02	(0.005)
depreciation rate (δ)	0.08	(0.002)

The magnitude of the effect of training is slightly larger than the human capital return from working part-time hours, which are estimated to be 13 percent of the full-time return at the start of working life ($g_1(18)$ in second row of the Table). We also allow for an interaction term with the stock of human capital ($g_2(18)$), and again find no evidence of the need to allow for more flexibility in how part-time hours affect human capital and wages. The only interaction of the stock of human capital that is statistically significant at conventional levels is that with full-time hours ($g_2(38)$), but even there the effect is small. Our estimate shows that one additional unit of human capital reduces the human capital return to full time hours by 2%. Since human capital never increases beyond 12 in simulations, at the maximum this parameter is responsible for a 24% drop in the human capital return to one year of full-time work.

Figure 7: Wage return to one episode of training while working full-time, by education



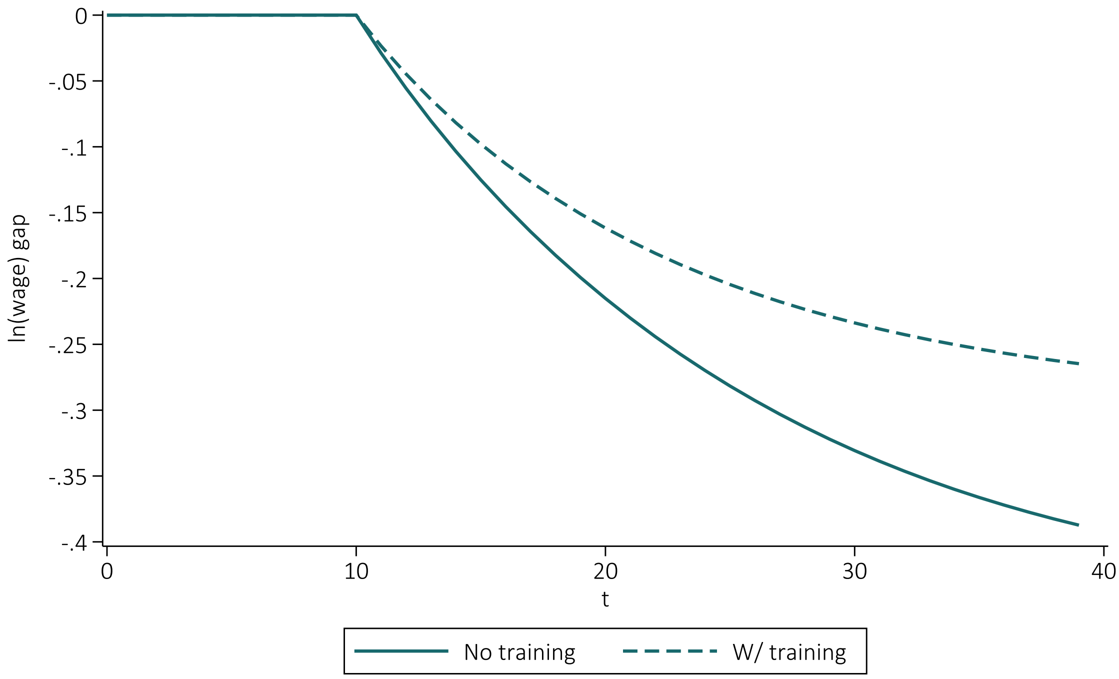
Notes: Percentage change in wage rates due to single episode of training in years 1 (LHS panel) and 10 (RHS panel) of full-time work. Agent is assumed to have no human capital at t = 0 except for that acquired through formal education and is working full-time over the entire period.

The size of the impact of training on wages depends on the interactions between its impact on human capital (determined by τ_1) and its wage return (determined by a combination of $(\gamma_0, \gamma_1, \gamma_2)$ for different groups), the depreciation rate (δ), and the stock of human capital at the time of training. Figure 7 illustrates the overall short- and long-term wage effects of one episode of training taking place at different stages of the working life. The plot on the left shows the impulse response to one training episode in year 1 of working life for women in full-time hours; the plot of the right shows the equivalent figure if training happens after 10 years of full-time work.

There is a modest but not insignificant initial effect on wage rates that, however, declines quickly as the additional human capital depreciates over time. The initial effect is much more pronounced if training is taken earlier in the working life, prior to the building up of human capital with working experience and consistent with decreasing marginal returns to investments in human capital. For instance, training increases the wage rate by 1.5% if taken in the first period of work, but only by 0.4% if taken after 10 years of working full-time. The falling returns to training with accumulated human capital is an important determinant of the timing of training in our model.

Our estimates of the wage impact of training can be compared with estimates of the impact of one additional year of education found in the broader literature once adjusted for the relatively small number of hours spent in training. Assuming that school requires thirty hours of study per week and takes place over forty weeks, a year of schooling requires 1,200 hours of time investment. This is approximately 12 times longer than the 100 hours corresponding to a training episode within our model. Card (1999) surveys the vast literature on returns to education and finds estimates implying increases in wages of between 5% and 15% associated with an additional year of high school, or approximately 0.4% to 1.3% per 100 hours invested. Blundell et al. (2005) estimate a wage return of 24% for the two years of education differentiating High School graduates from those who leaving school at 16 (with less than high school qualifications) in the UK context, or approximately 1% per 100 hours invested. Our estimates of the initial return from training at the start of working life fall on very similar values.

Figure 8: Training and the wage penalty from working part-time hours, by education



Notes: Solid lines represent the wage penalty, in log points, from moving to continuous part-time work after 10 years of continuous full-time work. The dotted lines factor in continuous training starting in year 10, together with part-time working hours.

In Figure 8 we document the extent by which training can offset the part time penalty in wages. The diagram compares the loss in wages that results from a shift from full-time work to (a) part-time work (solid line) or (b) part-time work plus training (dashed line). It represents how the impact of training compares with that of part-time hours. The solid lines in the figure show that part-time work is associated with a large wage penalty. The dashed lines show that taking training together with part-time hours offsets almost one third of the part-time penalty.

8.2 Utility parameters and the cost of training

Tables 6 and 7 show estimates of the parameters driving the utility cost of work and training as defined by the index functions U_h and U_T in equations 15 and 16. In both Tables, a positive parameter reflects higher costs of working or training.

Table 6: Parameters determining utility cost of working

	Parameter Value (1)	St. Error (2)	Parameter Value (3)	St. Error (4)
	Utility Parameters in U_h			
	Full-Time Employment (α_{38})		Part-Time Employment (increment: $\alpha_{18} - \alpha_{38}$)	
Singles, no children	0.56	(0.006)	-0.37	(0.004)
Single mothers	0.47	(0.011)	-0.22	(0.009)
Married, no children	0.33	(0.014)	-0.23	(0.015)
Married mothers	0.34	(0.013)	-0.24	(0.012)
Child aged 0-2	0.16	(0.009)	-0.07	(0.008)
Child aged 3-5	0.11	(0.010)	-0.05	(0.009)
Child aged 6-10	0.06	(0.010)	-0.04	(0.006)
Spouse working	-0.07	(0.013)	0.08	(0.012)
High background factor 1	0.02	(0.008)	0.00	(0.005)
High background factor 2	0.03	(0.008)	-0.02	(0.005)
$l_h(\theta)$ type I	-0.38	(0.178)	0.00	(0.005)

In order to rationalise the observed employment rates at the given monetary incentives to work, the model requires working to carry a utility cost for all groups (see estimates in columns 1 and 2 of Table 6). The costs are lower for married women than for single women, partly offsetting differences in incentives to work between the two groups due to spouse's income and benefit entitlement. Moreover, a working spouse brings down the utility cost of working, a result in line with past research showing complementarity in spouses' leisure (Blundell et al. (2016)). Mothers of young children, particularly of pre-school age, also face higher costs of working. Columns 3 and 4 of the Table report estimates for the incremental effects of working part-time hours, showing that part-time is less onerous in utility terms than full-time hours.

Estimates for the parameters governing the utility cost of training are shown in Table 7. We have fixed the monetary cost of training to equal the foregone wage for 2 hours of training per week, or 104 hours per calendar year, which corresponds in the data to the median level of training among trainees undergoing more than one week of training over the year. The utility cost of training is identified from the discrepancy between the predicted take up of training (if

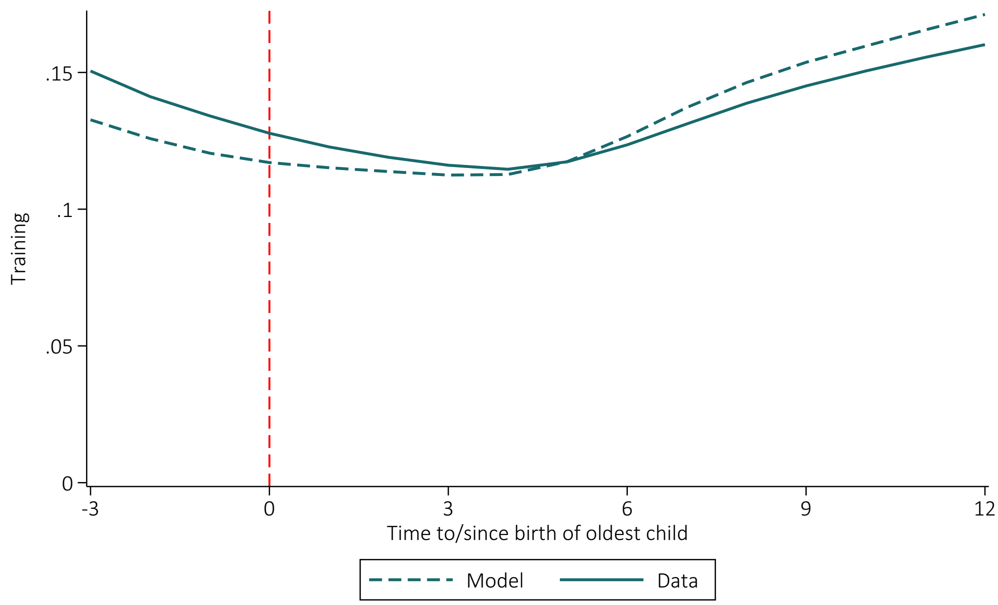
Table 7: Parameters determining utility cost and benefits of training and the terminal value of human capital

		Parameter Value	St. Error
	Utility Parameters in U_T, (α_T, α_{Th})		
(1)	Single, no children	0.002	(0.010)
(2)	Single mothers	0.007	(0.008)
(3)	Married, no children	-0.002	(0.015)
(4)	Married mothers	-0.002	(0.016)
(5)	Child aged 0 to 2	0.010	(0.025)
(6)	Child aged 3 to 5	0.004	(0.014)
(7)	Child aged 6 to 10	0.003	(0.008)
(8)	Spouse working	0.009	(0.014)
(8)	High background factor 1	0.004	(0.007)
(8)	High background factor 2	0.002	(0.005)
(9)	Part-time interaction	0.016	(0.006)
(10)	$l_T(\theta)$ type I	-0.028	(0.003)
	Terminal Value of Human Capital		
(11)	Scale parameter, ϕ_1	0.05	(0.009)
(12)	Curvature parameter, ϕ_2	0.21	(0.137)

costs were zero) and the actual take up.

Most parameters in the utility of training are small and mostly not statistically significant at conventional levels: the utility cost of training does not seem to depend on the demographic structure of the household or even on the family background factors. Perhaps this is not surprising since most of the cost associated to the household structure relates to the decision to work or not and once that has been paid it is no longer relevant for the training decision itself. However, the interaction with part-time hours (row 9) shows that training is more costly when women are doing short working hours; and the unobserved heterogeneity term (row 10) shows that the group with higher preferences for work also has a positive preferences for training (which mirrors a higher training cost for those with lower preferences for work). Thus, given our estimated returns to training our model rationalizes observed training cost as a preference for training among higher ability women (which constitute 80% of the population according to estimates in Panel B of Table 4). A model that admits search frictions or other imperfections could provide a structural interpretation of this since in that case the firm and the worker share the costs of training.

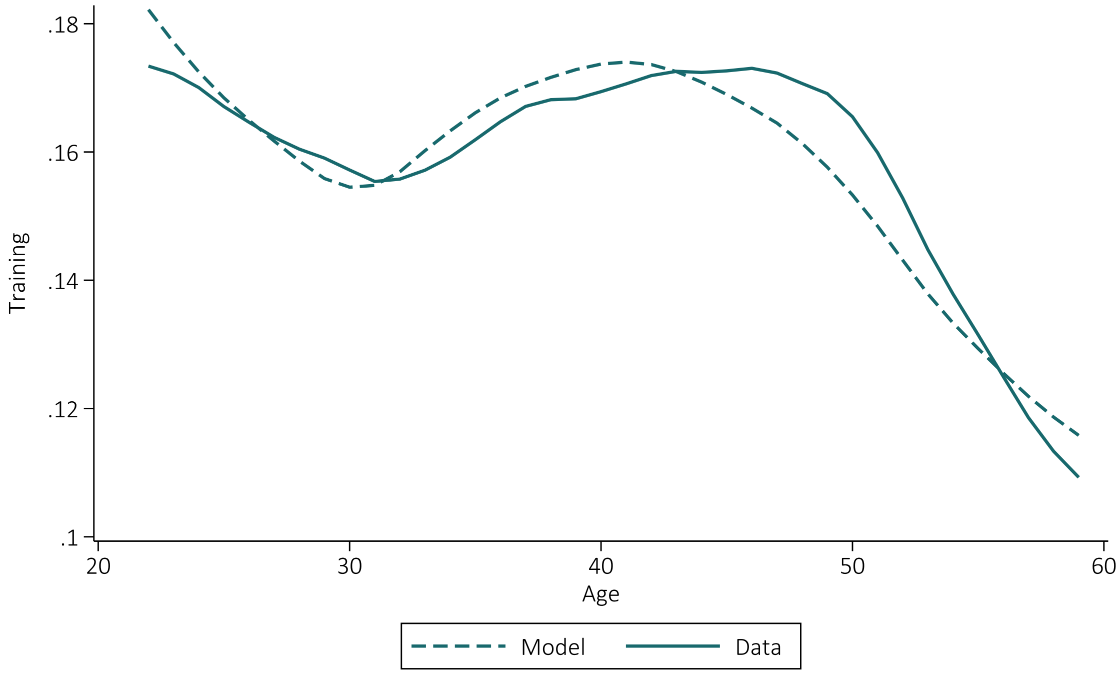
Figure 9: Model versus data – Training incidence among working mothers, by time since/to birth of oldest child and maternal education



Our model implicitly points to two additional mechanisms explaining the life-cycle patterns of training. First, families with children have higher needs and may be more likely to face liquidity constraints. In those circumstances, the foregone earnings associated with training may be an especially high cost to pay that could drive training rates down during that period of life. And second, the expected return to training may be negatively affected by motherhood as higher career intermittency limits women's ability to reap its full return before depreciation eventually washes out the human capital gains from training.

Figure 11 plots age profiles for the average total cost of training on the left, including both the monetary cost associated with lost labour time and the monetized direct utility cost. We

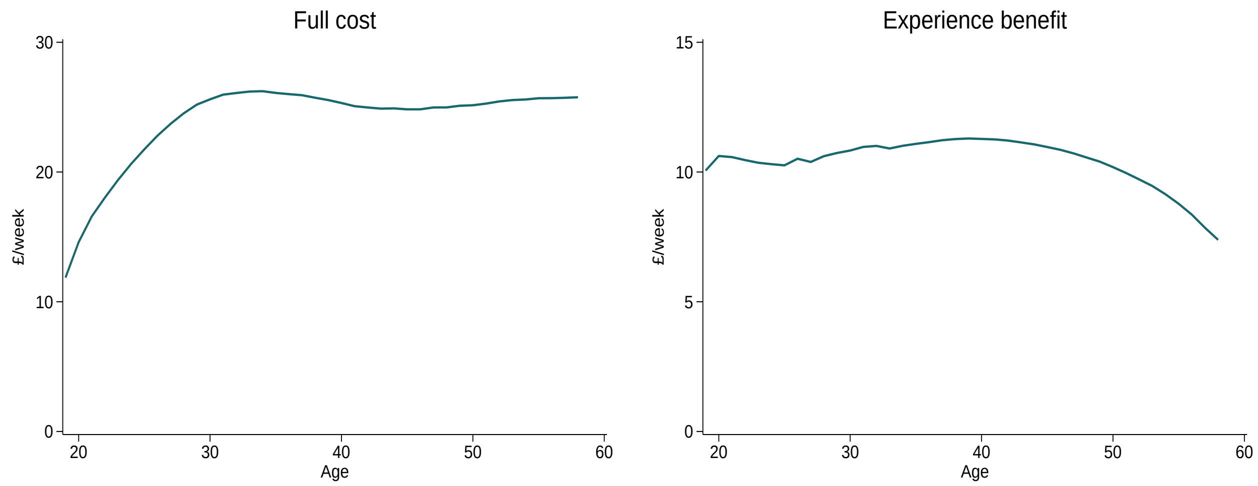
Figure 10: Model versus data – Training incidence over the life-cycle among working women, by maternal age and education



compare this to the consumption value of the additional human capital acquired through one episode of training on the right. In line with the observed training rates, average cost exceeds average return by a factor of 2 for most age groups. Figure 12 plots similar figures but by time to/from the birth of the first child. The life-cycle variation is strongly associated with the dynamics of family demographics through employment behaviour rather than through the utility cost of training. The returns to training also change around childbirth but by a much more modest amount, and then slowly recover as the child grows up.

Finally, the last two rows in table 7 show the parameters associated with the terminal value of human capital at the time of retirement. The scale parameter ϕ_1 is positive, which implies that human capital is valuable in retirement.

Figure 11: Monetised total cost of and experience return to training across whole population, by age and education



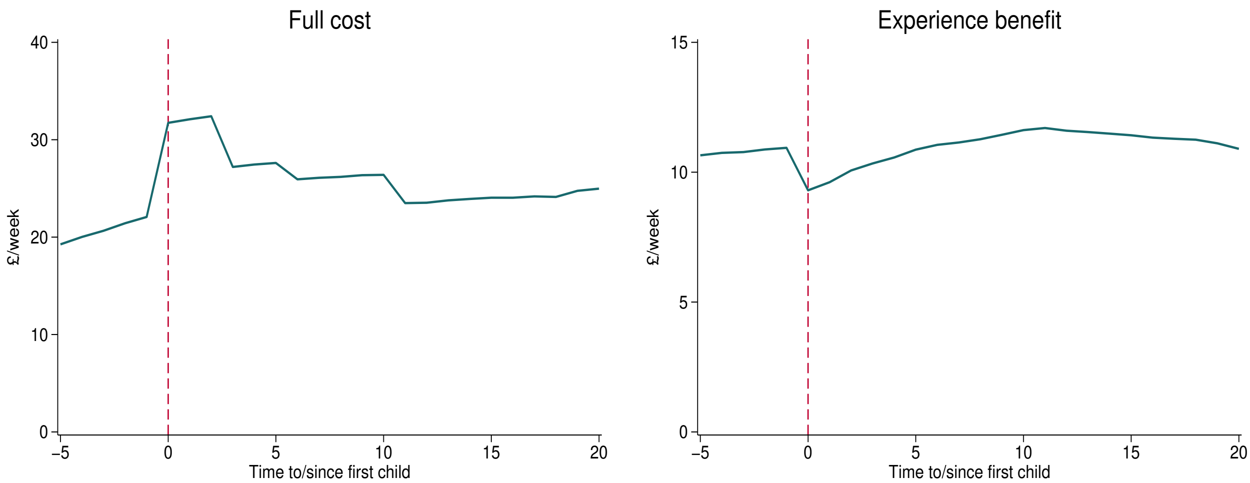
(a) Monetized cost of training (b) Consumption value of extra human capital

8.3 Responses of employment and training to changes in prices

We use the model to quantify responses to changes in the monetary incentives to work and train. Table 8 shows responses in employment rates (Panel A) and training rates among employed women (Panel B) to changes in the wage rates (column 2) and in the earnings foregone while training (column 3). Column 1 provides a sense of scale by displaying the simulated levels of employment and training by family demographics. All simulations are run under the 2002 tax system.

Column 2 reports average immediate response to an unanticipated and permanent 5% decline in the post tax wage rate starting at each age in the 23 to 50 interval. Overall, this change leads to a 2% decline in employment on a base of 85.7% displaying the dominance of the wealth effect. The response is larger for mothers, particularly single mothers, than it is for women without children, reflecting their larger labour supply elasticities. While training responses to changes in the wage rates are smaller than those of employment, they are nevertheless important given current training rates. A permanent drop in wages reduces future returns to training,

Figure 12: Monetised cost of and experience return to training across whole population, by time to/since first birth and education



(a) Monetised cost of training (b) Consumption value of extra human capital

Notes: Left-hand panels show average monetary compensation required to equalise period utility between (1) working full-time and not training and (2) working full-time and training. Right-hand panel shows the average monetary deduction required for an individual to be indifferent to receiving additional human capital equivalent to one unit of training.

but that is offset by the negative impact it has on the current cost of training. We find that the latter dominates leading to a small overall increase in training rates particularly for single mothers. Finally, column 3 in the Table shows that the training responses to a drop in the cost of training are large, particularly for single women. In turn, employment does not respond to changes in training incentives.

The parameters in Table 8 are key to inform policy as they reflect the potential responses in employment and training to reforms changing the monetary incentives to do so. They are consistent with the observed effects of the WFTC reform on employment and training. These are described in the set of moments we used to identify the model, and displayed at the bottom 8 rows of tables A2, A3 and A4. We can see that the model closely fits the employment and training rates before and after the reform for all family types.

Table 8: Model simulations – Employment and training responses to changes in wages and the monetary cost of training

	level (%)	5% permanent decrease in net earnings	5% permanent decrease in training cost
		(a) Employment	
All women	85.7	-2.0	0.0
By family demographics			
Singles, no kids	93.3	-1.4	0.0
Single mothers	69.7	-4.2	0.0
Couples, no kids	94.3	-0.8	-0.1
Mothers in couples	79.0	-2.6	0.0
		(b) Training conditional on employment	
All women	16.7	0.3	2.1
By family demographics			
Singles, no kids	16.3	0.0	3.0
Single mothers	10.3	0.7	2.9
Couples, no kids	18.4	0.3	1.6
Mothers in couples	16.6	0.2	1.9

Notes: Calculations based on model simulations. Column 2 shows effects of an unanticipated and uncompensated 5% permanent decrease in net earnings, on the employment and training rates in the period the change in earnings is first realised. Column 3 shows effects of an unanticipated and uncompensated 5% permanent decrease in the foregone earnings cost of training, on the employment and training rates in the period the change in costs is first realised. In all case, responses are averages of effects for women aged 23 to 50.

9 Counterfactual simulations and discussion

9.1 Subsidized training for mothers

We now investigate the long-term impacts of subsidizing training for mothers of young children, who may have especially loose links to the labor market. The policy could impact the labor market outcomes of these mothers in two ways. First, by increasing training rates among eligible mothers, it may help recover some of the losses in productive human capital associated with career interruptions once mothers return to work. Second, the subsidy may also reduce the duration of career breaks by indirectly promoting employment during the early stages of motherhood. The results from the previous sections suggest that mothers are especially sensitive to the cost of training and that training has modest but positive effects on wages, so the question is whether subsidizing training could help close the cost of child-rearing for mothers.

We compare outcomes under the 2002 tax and benefit system with three modified regimes that introduce training subsidies. In all three cases, mothers of children aged 7 or younger are entitled to subsidies of different levels of generosity if they decide to take up training.

Our simulations quantify the long-term effects of these policies for women living through the new regimes over their entire lives. All effects are calculated under revenue neutrality, with any costs being recovered through adjustments in the basic tax rate from the tax liabilities net of benefit entitlements of this group of women and their partners. The way one achieves revenue neutrality is relevant since, for example, changing the tax rate to fund subsidies has its own incentive effects.

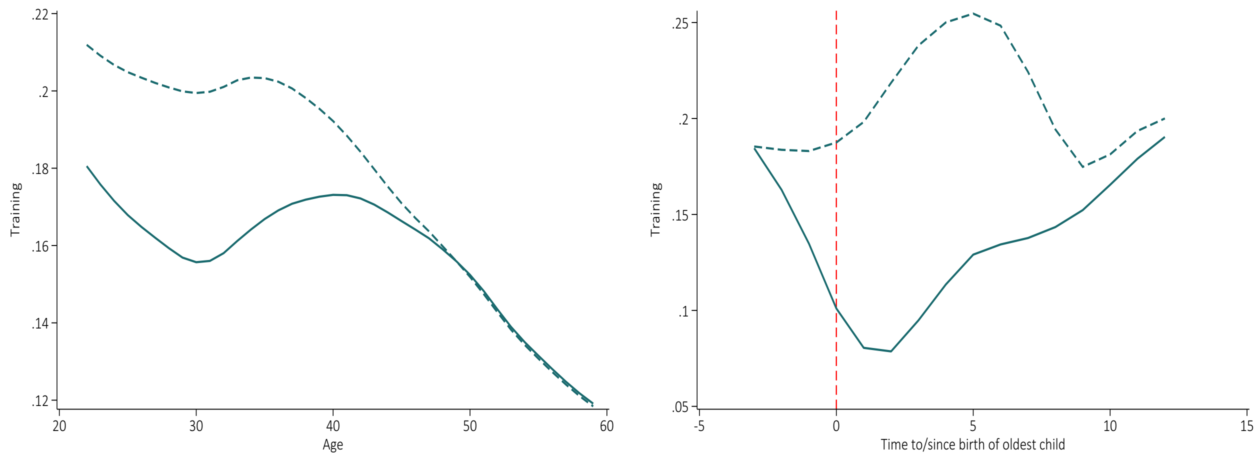
Table 9 shows model predictions of the effects of subsidized training on training rates, employment, hours, wages, savings, income and welfare. The first column displays the effects of a £500 lump-sum subsidy for mothers of children aged 0 to 7 in training. The second column increases this to £1500, and in the final column the subsidy provides full compensation for the monetary cost of training, which includes foregone earnings. The subsidy policy is made revenue neutral with a change in the basic tax rate.

Under our assumption of standard training units of 2 hours per week, the £500 annual subsidy amounts to approximately £5 per hour. This is not a trivial subsidy, making up about 40% of the average hourly wage rate of eligible mothers. However, it is more modest than other work related subsidies such as Tax Credits because it only supports a limited amount of training.

The results in the first column show that training rates respond strongly to the subsidy during the eligibility period (panel (a)), as suggested by the responses to a drop in the monetary cost of training described in table 8. Moreover, the effect quickly fades to 0 in later periods when mothers lose eligibility (panels (c) and (e)).

The subsidy is timed to coincide with the fall in training we observe around the birth of first child. Figure 13 shows, as an example, the impact of the subsidy on the prevalence and timing of training. The fall in training at the time of childbirth, which is observed in the data and replicated by our baseline model, is completely offset by the subsidy. As a result, training rates decline gradually over the lifecycle, resembling the male training profiles discussed above (see Figure 5).

Figure 13: Training over lifecycle for High School educated under £500 subsidy



(a) Training rate by age (b) Training rate by time since birth of oldest child

The least generous subsidy has a small impact on full time employment, increasing it by 0.64 percentage points during the first 7 years of the child, which corresponds to the period of entitlement (panel (a)). All of this extra time in paid work comes from those who were previously doing part-time work, resulting in a net effect on employment close to zero. The small net response in employment is aligned with predictions of how employment responds to changes in the cost of training, detailed in Table 8.

Panel (b) shows that the cumulative effect of the additional training and full time work on the wage rates of women at the end of the eligibility period is positive, with mothers benefiting from a 0.31% increase in wages. This demonstrates that the policy has a small but not negligible impact on the human capital of mothers at the end of the eligibility period. The subsidy also reduces savings by a modest 0.12% when the child reaches 8 years of age. This suggests that the additional human capital the woman accumulated over this period will make future work more likely and she will need to rely less on savings.

Indeed we find that the small impacts on human capital and assets at the end of the eligibility period drive similarly small dynamic effects. Panels (c) and (e) confirm that that the policy slightly increases employment after the eligibility period, and that all increase is on the full-time margin. These responses drive an increase in the lifetime disposable income of the families of these women by 0.24% (panel (d)) and a larger increase in equivalised consumption of 0.83%.[12] Since the counterfactual simulation is revenue neutral, all these responses are net of the tax adjustment. In the case of this less generous policy, we find that it pays for itself. By bringing more women into full-time work for an extended period, the government raises in extra taxes the funds required to implement the subsidy (panel (g)).

Column 2 of the Table shows similar results for a more generous lump-sum subsidy of £1,500 per year, or about 120% of the pay of eligible mothers during training episodes. The additional generosity comes with a high price, requiring an increase of 0.5pp in the basic tax rate to balance the public budget. For comparison, Blundell et al. (2016) calculations suggest that funding for the 2002 Tax Credit scheme in the UK adds 0.9pp to the basic tax rate. Despite its cost, which is fully borne by this population of women and their partners, our simulations show that this policy is welfare increasing and drives up disposable income by more than the less generous policy. These effects result from the strong impact that the policy has on the training rates of

[12]The value of the consumption compensation (ι) is the solution to:

$$EV_0 = E\sum_t \beta^{t-\underline{t}} \frac{((1-\iota)\, c_{1t}/n_{1t})^{\mu}}{\mu} \exp\{U(h_{1t}, d_{1t}, \theta, \omega, X_{1a})\}$$

where the index 0/1 stands for the pre/post-reform solutions and the value function is evaluated at different stages in life for different rows. The equation can be solved for ι, yielding: $\iota = 1 - \left(\frac{EV_0}{EV_1}\right)^{\frac{1}{\mu}}$.

eligible mothers, which increase by 41 percentage points, and their employment rates, which also increase by 3 percentage points. The combined effect of these responses result in higher wages at the end of the eligibility period, by 2.6% (panel (b)) that drive later employment gains and persistent increases in wages (by 0.72% when the child reaches 19 years of age).

The lump-sum subsidies provide a stronger incentive for those in low pay, who may also benefit less from training if they are from the low ability group or on a flatter wage trajectory induced by low (persistent) productivity shocks. We therefore re-designed the subsidy to exactly cover the foregone earnings of trainee mothers of children aged 0 to 7. Results for this policy are displayed in column 3 of Table 9. Because this type of design incetivises training among higher paid mothers, it also ends up being more expensive for each trainee than the generous £1,500 lump sum subsidy, costing £1,600 per trainee. However, it draws fewer women into training than the lump-sum benefit because it is less generous for lower paid women. So in the end the cost of such policy is smaller than that of the more generous lump-sum transfer, requiring an increase of 0.15pp in the basic tax rate to balance the public accounts. Its effects lie between the figures in the first two columns of the table, for the two lump-sum subsidies.

10 Conclusions

We have estimated a lifecycle model of female labor supply, and human capital accumulation through work experience and training. Our main aim has been to understand the role that job training can have in offsetting the loss of experience resulting from having children, which leads to an increasing wage gap for women with children.

Training can be important for wages and we show that it can partly offset the wage gap attributable to the prevalence of part time work and non-employment following a return to the labor market after having children.

Finally, we evaluate a policy of subsidizing training for mothers with children younger than 8. All policies are revenue neutral and funded by increasing taxes. A fixed modest subsidy of

£500 increases the take up of training substantially and leads to small but persistent gains in wages, lifetime disposable income and welfare. It also pays for itself. We also consider other less effective and more expensive approaches.

This paper has ignored the all important question of incidence for the costs of training as well as for the returns. In a classical competitive labor market workers pay for general training and wages fully reflect returns to investment (Becker 1964). But in the presence of frictions this may not occur; firms and workers may share both the returns and the costs of training. While here we measure correctly the returns to the individual and attribute some of the costs to them we have not considered the returns to the firm of individuals being trained or how the firms and the workers may share the costs. This is a central question, all the more so if we are to understand why college graduates have such high levels of job training but little or no observed return. In a follow up paper we are investigating this issue based on a model inspired by Acemoglu and Pischke (1999) (see also Flinn et al. (2017), Lentz and Roys (2015)).

Table 9: Impact of training subsidies

	Annual training subsidy		
	£500	£1500	Full compensation
	(*a*) Child aged 0-7		
Training	9.33	40.95	19.84
Employment	0.02	3.14	0.01
Full-time	0.64	1.37	1.35
Part-time	-0.62	1.77	-1.34
	(*b*) Child aged 8		
Assets (%)	-0.12	0.51	0.69
Wages (%)	0.31	2.60	1.36
	(*c*) Child aged 8-18		
Training	-0.19	-0.38	-0.31
Employment	0.10	0.52	0.15
Full-time	0.41	-0.30	0.20
Part-time	-0.31	0.82	-0.05
	(*d*) Child aged 19		
Assets (%)	-0.05	0.24	0.29
Wages (%)	0.14	0.72	0.44
	(*e*) Child aged 19+		
Training	0.00	-0.03	0.03
Employment	0.21	0.26	0.23
Full-time	0.72	0.26	0.62
Part-time	-0.51	0.00	-0.39
	(*f*) Lifetime outcomes		
Disposable income (%)	0.24	0.35	0.23
Consumption equivalent (%)	0.83	0.74	0.77
	(g) Revenue neutrality adjustment		
Basic income tax change	-0.02	0.5	0.15

Notes: Calculations based on model simulations. Column 1 shows the effects of a £500 yearly subsidy, while columns 2 shows similar calculations for an yearly subsidy of £1,500. Column 3 shows simulated figures for a subsidy that exactly covers foregone earnings of trainee mothers. In all cases, only mothers of children aged 0 to 7 are entitled to the subsidy if taking training. Age of the child in panels (a) to (e) refers to the youngest child in the family. The change in disposable income (panel (f)) is net of the tax adjustment. The consumption equivalent in the same panel is calculated at the start of working life to keep expected lifetime utility constant. All changes are in percentage points unless otherwise stated.

References

Acemoglu, D. and J.-S. Pischke (1999, February). Beyond becker: Training in imperfect labour markets. *The Economic Journal 109*(453), 112–142.

Adam, S., J. Browne, and C. Heady (2010). Taxation in the uk. In J. Mirrlees, S. Adam, T. Besley, R. Blundell, S. Bond, R. Chote, M. Gammie, P. Johnson, G. Myles, and J. Poterba (Eds.), *Dimensions of Tax Design: The Mirrlees Review*. Oxford University Press.

Altonji, J. G. and C. R. Pierret (2001, February). Employer learning and statistical discrimination. *Quarterly Journal of Economics*, 313–350.

Altonji, J. G. and L. M. Segal (1996, July). Small-sample bias in GMM estimation of covariance structures. *Journal of Business & Economic Statistics 14*(3), 353.

Altonji, J. G. and J. R. Spletzer (1991, October). Worker characteristics, job characteristics, and the receipt of on-the-job training. *Industrial and Labor Relations Review*, 58–79.

Andrews, I., M. Gentzkow, and J. M. Shapiro (2017, June). Measuring the sensitivity of parameter estimates to estimation moments. *The Quarterly Journal of Economics 132*(4), 1553–1592.

Attanasio, O., H. Low, and V. Sánchez-Marcos (2008, August). Explaining changes in female labor supply in a life-cycle model. *American Economic Review 98*(4), 1517–1552.

Attanasio, O. P. and G. Weber (1995, December). Is consumption growth consistent with intertemporal optimization? evidence from the consumer expenditure survey. *Journal of Political Economy 103*(6), 1121–1157.

Blundell, R., M. Browning, and C. Meghir (1994, January). Consumer demand and the life-cycle allocation of household expenditures. *The Review of Economic Studies 61*(1), 57–80.

Blundell, R., L. Dearden, and B. Sianesi (2005, July). Evaluating the effect of education on earnings: models, methods and results from the national child development survey. *Journal of the Royal Statistical Society: Series A (Statistics in Society) 168*(3), 473–512.

Blundell, R., M. C. Dias, C. Meghir, and J. Shaw (2016). Female labor supply, human capital, and welfare reform. *Econometrica 84*(5), 1705–1753.

Blundell, R., L. Pistaferri, and I. Saporta-Eksten (2016, February). Consumption inequality and family labor supply. *American Economic Review 106*(2), 387–435.

Brewer, M., A. Duncan, A. Shephard, and M. J. Suarez (2006, December). Did working families tax credit work? the impact of in-work support on labour supply in great britain. *Labour Economics 13*(6), 699–720.

Card, D. (1999). The causal effect of education on earnings. In *Handbook of Labor Economics*, pp. 1801–1863. Elsevier.

Chiappori, P.-A., M. C. Dias, and C. Meghir (2018, October). The marriage market, labor supply, and education choice. *Journal of Political Economy 126*(S1), S26–S72.

Chiappori, P.-A., M. Iyigun, and Y. Weiss (2009, December). Investment in schooling and the marriage market. *American Economic Review 99*(5), 1689–1713.

Fan, X., A. Seshadri, and C. Taber (2017, September). Understanding Earnings, Labor Supply, and Retirement Decisions. Working Papers wp367, University of Michigan, Michigan Retirement Research Center.

Flinn, C., A. Gemici, and S. Laufer (2017, April). Search, Matching, and Training. *Review of Economic Dynamics 25*, 260–297.

Gourieroux, C., A. Monfort, and E. Renault (1993). Indirect inference. *Journal of Applied Econometrics 8*, S85–S118.

Heckman, J., L. Lochner, and R. Cossa (2002, July). Learning-by-doing vs. on-the-job training: Using variation induced by the EITC to distinguish between models of skill formation. Technical report.

Lentz, R. and N. Roys (2015, November). Training and search on the job. Technical report.

Shaw, J. (2011, May). FORTAX: UK tax and benefit system documentation.

Wages, Experience and Training of Women over the Lifecycle: Web Appendix

Richard Blundell
University College London and Institute for Fiscal Studies

Monica Costa-Dias
Institute for Fiscal Studies and University of Porto

David Goll
University College London and Institute for Fiscal Studies

Costas Meghir
Yale University, NBER, IZA, CEPR and Institute for Fiscal Studies

August 2020

Estimation is based on all 18 yearly waves of the British Household Panel Survey (BHPS), covering the period from 1991 to 2008. Apart from those who are lost through attrition, all families in the original 1991 sample and subsequent booster samples remain in the panel from then onwards. Other individuals have been added to the sample in subsequent periods, sometimes temporarily, as they formed families with original interviewees or were born to them. All members of the household aged 16 and above are interviewed. We select the sample of women in all types of family arrangement observed while aged 19 to 59.

Some definitional and data preparation procedures should be mentioned for clarity. Employment is determined by present labor-market status and excludes self-employment. The paths of women who report being self-employed are deleted from that moment onwards. This partly eliminates the trajectories of 889 women of the original sample of 7,755 women, dropping 6,569 individual-year observations. Similarly, we truncate the paths of women who report returning to full-time education after they have entered the labor market. 764 individuals are observed returning to full-time education, for whom we drop their pathways from that moment onwards, amounting to 4,737 individual-year observations. We start with 67,399 and end with 56,093 individual-year observations after the cleaning trajectories after they cross self-employment or full-time education. We also drop observations for women for whom the age of children are missing or look wrong. This leaves 55,591 individual-year observations in our final dataset.

Only women working 5 or more hours per week are classified as employed. We consider employment choices from the age of 19 for women with secondary and high school education, and from the age of 22 for women with university education. Working hours refer to the usual hours in main job including overtime. We discretized labor supply using a three-point distribution: not working (0 to 4 hours per week, modeled as 0 hours), working part-time (5 to 20 hours per week, modeled as 18 hours), and working full-time (21 hours or more per week modeled as 38 hours). The employment status and working hours observed at one point in the year are assumed to remain unaltered over the entire year. Earnings are the usual gross weekly earnings in the main job. (Hourly) wage rates are the ratio of weekly earnings to weekly hours capped at 70. The wage distribution is trimmed at percentiles 2 and 99 from below and above, respectively, and only for women working at or above 5 hours per week to reduce the severity of measurement error in wage rates.

Wage rates are detrended using the aggregate wage index and all other monetary parameters in the model, including all monetary values in the annual sequence of tax and benefit systems, were deflated using the same index. To construct this index, we run three regressions, one for each education level, of trimmed wages on time dummies and dummies of Scotland and Wales. We create three education-specific wage indices from the coefficients in time. Then we aggregate these indices using the distribution of education for the entire population of workers aged 25-59 in the sample to form the wage index. Any real monetary values (using the CPI) are then rescaled using this index.

Family type includes four groups: single women and couples without children, lone mothers, and couples with children. Women are assumed to have children only after finishing education, once entering the labor market. Cumulated work experience is measured in years. Individual assets at the beginning of adult life are the total of savings and investments net of debts. They are truncated at zero, never allowed to be negative.

Our full data set remaining after the sample selection procedure described above, used for the descriptive graphs and tables, is an unbalanced panel of 7,359 women observed for some varying period during the years 1991 to 2008. A great deal of information is collected for them, including family demographics, employment, working hours and earnings as well as those of a present partner, women's demographics such as age and education, demand for childcare and its cost.

Within the model, we focus on "high-school" educated women. These women have completed A-levels or equivalent qualifications, which are acquired at the end of high school at the age of 18. They do not possess a first degree level or post-graduate level qualification. High-school educated women are 32% of the individuals in our full sample and 31% of observations. Moreover, for inclusion in our model sample we require observation of historical data on the characteristics of their parental home when they were aged 16, including whether lived with parents, parent's education, employment status, number of siblings and sibling order, books at home. Of the 2,377 individuals with high-school education, we observe the family background for 1,443 (60.7%). These individuals form the basis for our moment estimates and the initial conditions of the model.

Figure A1 and A2 replicate Figure **??** from the main text, using a slightly different method. Rather than smoothing training rates using a local polynomial, we have binned individuals into five year age ranges and presented the average training rate for each bin alongside the 95% confidence interval. Figure A1 presents the training rates of men, while Figure A2 presents the training rates of women. In each case, Panel A presents the unconditional training rates and Panel B presents training rates conditional on working. The training rates of men appear to decline steadily over the lifecycle, whereas training rates of women decline at first but increase somewhat during their 40s.

For completeness, we have included below an alternative version of Table **??** in the main text. Whereas the table presented in the main text includes all individuals in our sample, including those who are not employed, Table A1 conducts the same regression conditioning on working more than 5 hours per week. The sample is significantly smaller and the instruments lose some power, particularly for low education individuals. However, simulated full-time income retains strong explanatory power for training among the employed.

We also present some additional graphs showing the fit of the model in terms of employment and part-time hours over the lifecycle (Figure A3) and over age of the oldest child (Figure A4).

Tables A2 to A18 display the full list of data moments used in estimation, together with their simulated counterparts and the normalized (by the data standard error) differences between the two. Estimation used 139 moments, which fall into the following categories:

- Mean employment, part-time hours and training conditional on demographics (Table A2, A3 and A4)
- Mean employment and training conditional on age band (Table A5 and A6)
- Transition rates from unemployment to employment conditional on demographics (Table A7)
- Transition rates from employment to unemployment conditional on demographics and wage decile (Table A8)
- Mean, variance and quantiles of log wage at entrance to working life (Table A9)
- Log wage regression in first differences on training dummy and change in log experience (Table A10)
- Log wage regression on lagged wage, family background, log years of work experience and lagged log years of work experience (Table A11)
- Log wage regression on training, experience and working status last period (Table A12)

- Log wage regression on age and family background (Table A13)
- Mean yearly change in wages conditioning on working status last period (Table A14)
- Mean wages and proportion of population with wages below pre-defined empirical wage deciles, conditional on working hours and training (Table A15, A16 and A17)
- Mean log wages conditional on family background (Table A18)

All moments are constructed from the BHPS and are education-specific. Among the 139 simulated moments, 19 fall outside the 95% confidence interval for the respective data moment, but many amongst these are very similar to their BHPS counterparts.

Table A1: Regression of training conditional on employment

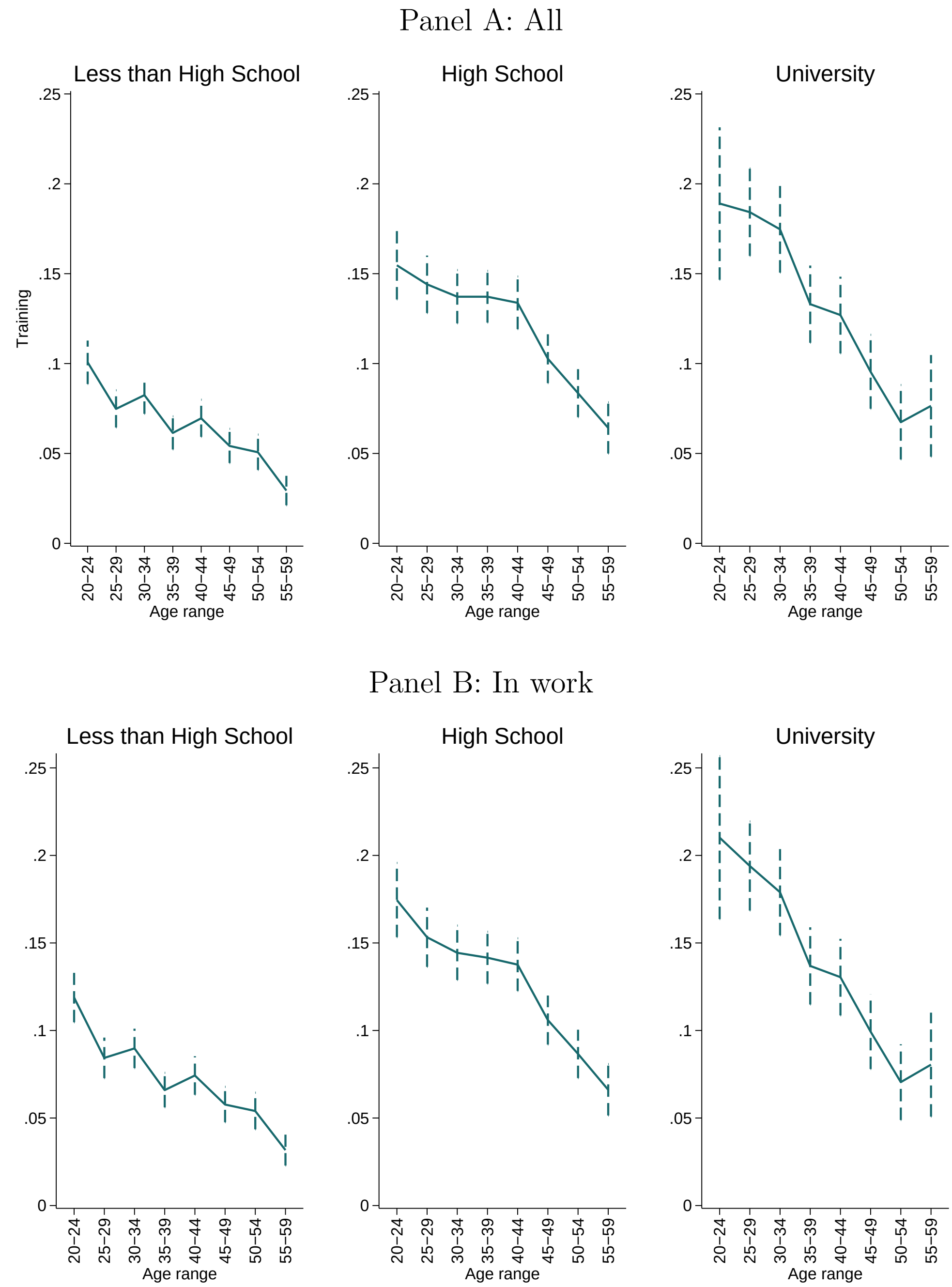
	(1) Secondary	(2) High School	(3) Degree
Sim Income: 0 hours	-0.0000393	-0.000134	-0.000133
	(0.000123)	(0.000171)	(0.000237)
Sim Income: 20 hours	0.000436**	0.000389	0.000711
	(0.000198)	(0.000288)	(0.000434)
Sim Income: 40 hours	-0.000668***	-0.000793***	-0.000922***
	(0.000148)	(0.000181)	(0.000255)
Observations	22739	14658	6537
Demographic Controls	Yes	Yes	Yes
Family Background Controls	Yes	Yes	Yes
Wave Dummies	Yes	Yes	Yes
F-Stat	7.259	8.787	6.369
F-Stat p-val	0.0000751	0.00000862	0.000282

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: BHPS data. Outcome variable indicates whether the individual is observed in more than 40 hours of work-related training. Sample is conditioned on working at least 5 hours a week. Standard errors are clustered at the individual level. Demographic controls include a quadratic in age and dummies indicating family composition. Family background controls include the first two principal components drawn from a collection of variables that describe the childhood household of each individual and an indicator for whether this information is missing.

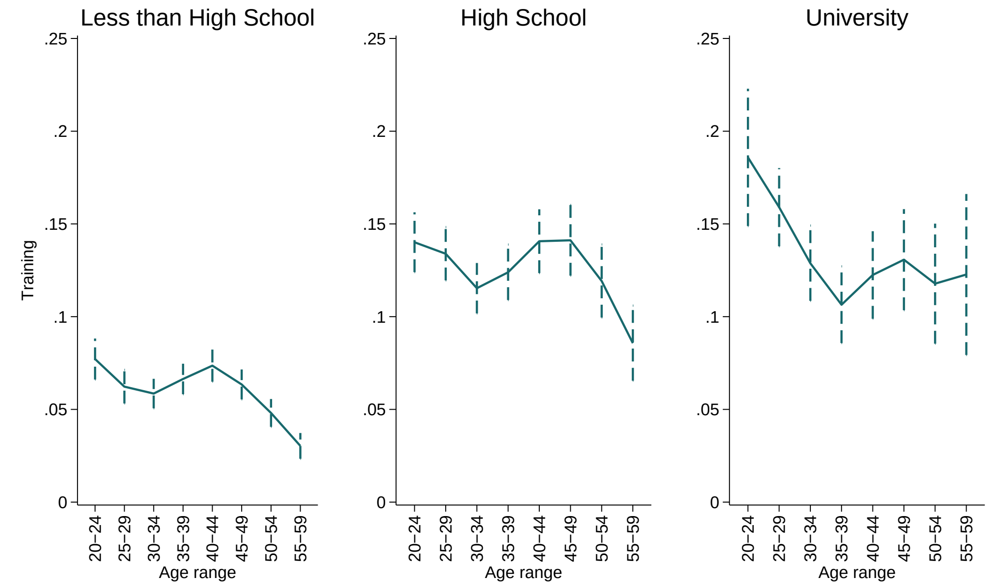
Figure A1: Training rates for men over the lifecycle



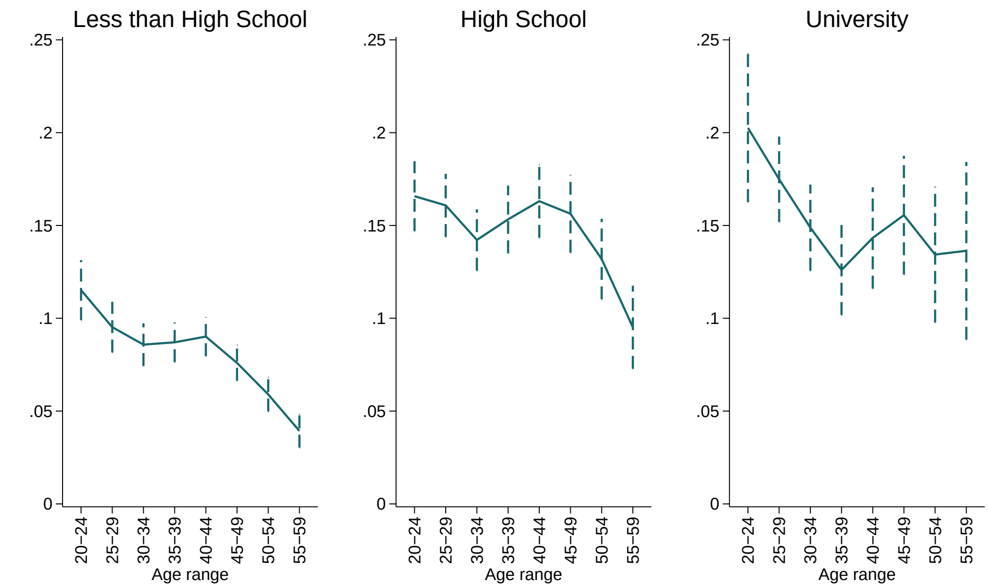
Notes: BHPS data for years 1991-2008. The training variable is an indicator for having had 40 or more hours of work-related training over the last 12 months. Panel A shows training rates for the entire population, by age and education. Panel B additionally conditions on working at least 5 hours per week on an usual week, which is the measure of employment used in this paper. Dashed line shows the 95% confidence interval.

Figure A2: Training rates for women over the lifecycle

Panel A: All



Panel B: In work



Notes: BHPS data for years 1991-2008. The training variable is an indicator for having had 40 or more hours of work-related training over the last 12 months. Panel A shows training rates for the entire population, by age and education. Panel B additionally conditions on working at least 5 hours per week on an usual week, which is the measure of employment used in this paper. Dashed line shows the 95% confidence interval.

Figure A3: Employment over life-cycle

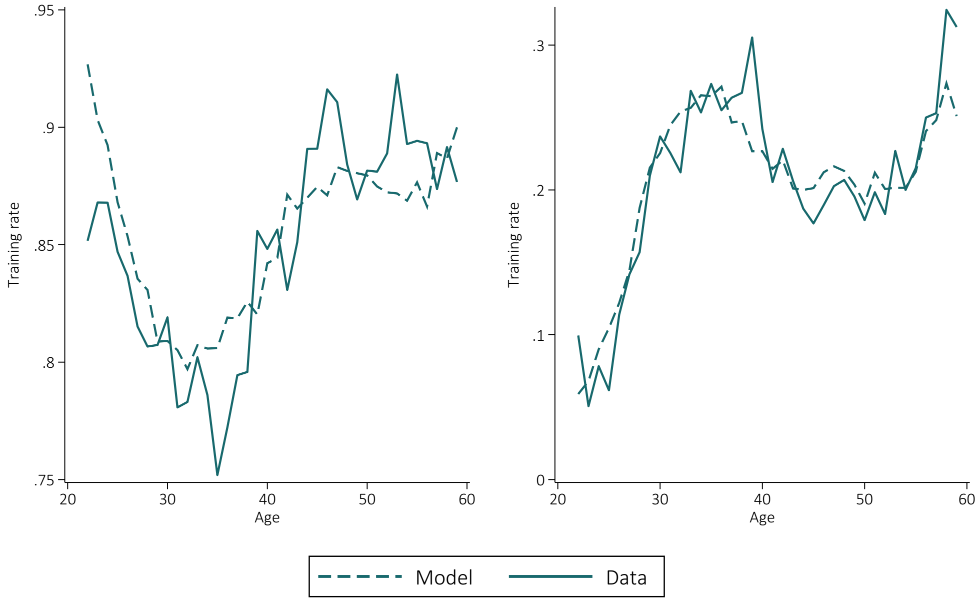


Figure A4: Employment of mothers

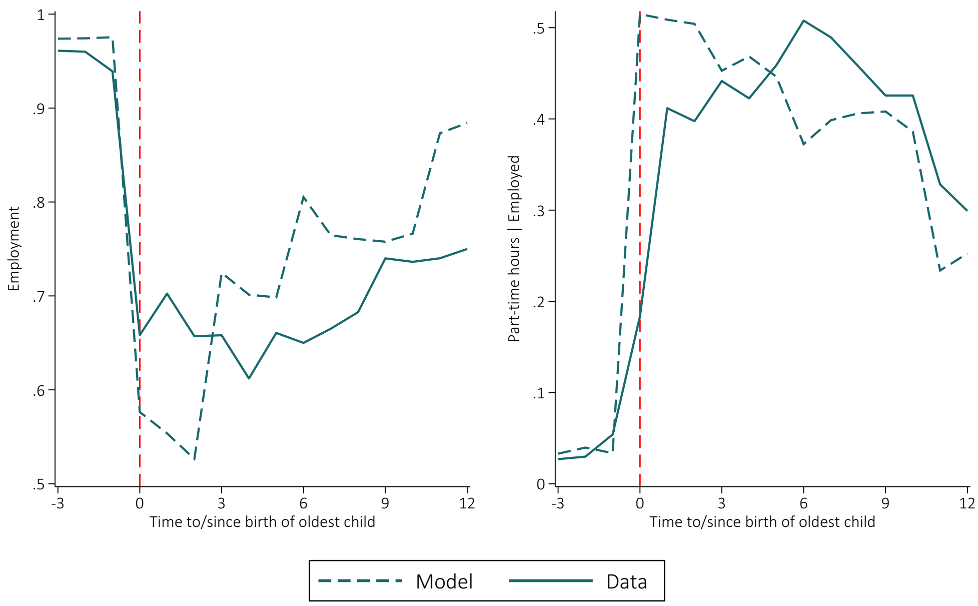


Table A2: Mean employment during working life

Moment	Data	Simulated	SE data	Norm. SE diff
Panel A: Averages by demographics				
All	0.837	0.829	0.010	0.791
Single women, no child	0.914	0.920	0.011	0.505
Married women, no child	0.938	0.935	0.010	0.300
Lone mothers	0.682	0.644	0.041	0.948
Married mothers	0.731	0.730	0.019	0.039
Partner working	0.837	0.847	0.012	0.883
Youngest child 0-2	0.602	0.570	0.025	1.309
Youngest child 3-5	0.713	0.714	0.024	0.047
Youngest child 6-10	0.777	0.762	0.025	0.616
Youngest child 11+	0.854	0.870	0.023	0.683
Family background: factor 1	0.825	0.834	0.014	0.633
Family background: factor 2	0.841	0.850	0.014	0.612
Panel B: Impact of benefit reform				
Pre-1999: single women, no child	0.084	0.076	0.018	0.439
Pre-1999: married women, no child	0.110	0.098	0.013	0.917
Pre-1999: lone mothers	-0.251	-0.217	0.053	0.648
Pre-1999: married mothers	-0.101	-0.079	0.016	1.398
Post-1999: single women, no child	0.071	0.104	0.015	2.291
Post-1999: married women, no child	0.093	0.114	0.013	1.657
Post-1999: lone mothers	-0.081	-0.156	0.044	1.699
Post-1999: married mothers	-0.111	-0.119	0.015	0.551

Notes: Moments in Panel A are average employment rates (measured as working five or more hours per week) among individuals with the listed demographic and background characteristics. Moments in Panel B are the deviation in percentage points of each of the family types employment rates from average employment rates in the period indicated.

Table A3: Mean part-time employment during working life

Moment	Data	Simulated	SE data	Norm. SE diff
Panel A: Averages by demographics				
All	0.164	0.170	0.009	0.723
Single women, no child	0.062	0.074	0.011	1.104
Married women, no child	0.096	0.100	0.012	0.348
Lone mothers	0.173	0.205	0.036	0.901
Married mothers	0.277	0.272	0.016	0.317
Partner working	0.191	0.193	0.011	0.160
Youngest child 0-2	0.257	0.280	0.020	1.163
Youngest child 3-5	0.321	0.316	0.024	0.200
Youngest child 6-10	0.281	0.280	0.024	0.010
Youngest child 11+	0.196	0.165	0.027	1.116
Family background: factor 1	0.158	0.135	0.012	1.928
Family background: factor 2	0.179	0.187	0.013	0.590
Panel B: Impact of benefit reform				
Pre-1999: single women, no child	-0.115	-0.106	0.018	0.502
Pre-1999: married women, no child	-0.068	-0.070	0.014	0.158
Pre-1999: lone mothers	-0.054	-0.010	0.035	1.264
Pre-1999: married mothers	0.128	0.115	0.015	0.805
Post-1999: single women, no child	-0.089	-0.086	0.015	0.201
Post-1999: married women, no child	-0.068	-0.070	0.012	0.169
Post-1999: lone mothers	0.059	0.075	0.044	0.350
Post-1999: married mothers	0.100	0.088	0.014	0.813

Notes: Moments in Panel A are average rates of part-time hours (measured as working between 5 and 20 hours a week) among employed individuals with the listed demographic and background characteristics. Moments in Panel B are the deviation in percentage points of each of the family types part-time hours rates from average part-time hours rates in the period indicated.

Table A4: Mean training during working life

Moment	Data	Simulated	SE data	Norm. SE diff
Panel A: Averages by demographics				
All	0.160	0.169	0.007	1.291
Single women, no child	0.173	0.185	0.014	0.901
Married women, no child	0.173	0.174	0.010	0.087
Lone mothers	0.162	0.177	0.028	0.545
Married mothers	0.138	0.141	0.011	0.334
Partner working	0.152	0.155	0.008	0.370
Youngest child 0-2	0.086	0.080	0.013	0.436
Youngest child 3-5	0.127	0.119	0.016	0.503
Youngest child 6-10	0.159	0.149	0.017	0.606
Youngest child 11+	0.192	0.196	0.018	0.212
Family background: factor 1	0.158	0.148	0.009	1.053
Family background: factor 2	0.165	0.163	0.010	0.188
Part-time hours	0.069	0.069	0.008	0.018
Panel B: Impact of benefit reform				
Pre-1999: single women, no child	0.026	0.030	0.017	0.252
Pre-1999: married women, no child	0.009	0.005	0.010	0.343
Pre-1999: lone mothers	0.032	0.035	0.044	0.072
Pre-1999: married mothers	-0.030	-0.034	0.012	0.330
Post-1999: single women, no child	0.000	0.009	0.014	0.674
Post-1999: married women, no child	0.015	0.010	0.009	0.538
Post-1999: lone mothers	-0.006	-0.014	0.025	0.340
Post-1999: married mothers	-0.017	-0.019	0.010	0.190

Notes: Moments in Panel A are average training rates (measured as spending more than 40 hours in training over the last 12 months) among employed individuals with the listed demographic and background characteristics. Moments in Panel B are the deviation in percentage points of each of the family types training rates from average training rates in the period indicated.

Table A5: Employment by age

Moment	Data	Simulated	SE data	Norm. SE diff
20 - 24 years	0.850	0.919	0.015	4.552
25 - 29 years	0.822	0.840	0.016	1.106
30 - 34 years	0.794	0.806	0.018	0.641
35 - 39 years	0.792	0.813	0.020	0.997
40 - 44 years	0.855	0.850	0.020	0.231
45 - 49 years	0.895	0.865	0.020	1.470
50 - 54 years	0.893	0.868	0.022	1.099
55 - 59 years	0.886	0.884	0.026	0.099

Notes: Moments are average employment rates (measured as working five or more hours per week) for individuals in each of the age bands indicated.

Table A6: Training by age

Moment	Data	Simulated	SE data	Norm. SE diff
20 - 24 years	0.168	0.214	0.015	2.994
25 - 29 years	0.165	0.152	0.013	1.016
30 - 34 years	0.152	0.158	0.013	0.449
35 - 39 years	0.163	0.165	0.015	0.136
40 - 44 years	0.170	0.160	0.014	0.704
45 - 49 years	0.176	0.161	0.018	0.887
50 - 54 years	0.160	0.144	0.018	0.883
55 - 59 years	0.113	0.111	0.020	0.121

Notes: Moments are average training rates (measured as spending more than 40 hours in training over the last 12 months) for employed individuals in each of the age bands indicated.

Table A7: Transition rates from unemployment to employment

Moment	Data	Simulated	SE data	Norm. SE diff
All	0.251	0.275	0.017	1.372
Single women, no child	0.408	0.301	0.048	2.220
Married women, no child	0.187	0.218	0.035	0.869
Lone mothers	0.212	0.283	0.019	3.805

Notes: Moments are average transitions from unemployment (working less than 5 hours) to employment (working at least 5 hours) for individuals with the listed demographic characteristics.

Table A8: Transition rates from employment to unemployment

Moment	Data	Simulated	SE data	Norm. SE diff
All	0.050	0.051	0.003	0.528
Single women, no child	0.027	0.025	0.003	0.852
Married women, no child	0.086	0.148	0.016	3.809
Lone mothers	0.082	0.080	0.007	0.339
w_{t-1} below 1st decile	0.124	0.102	0.014	1.505
w_{t-1} below median	0.068	0.072	0.005	0.726
w_{t-1} below 9th decile	0.050	0.055	0.003	1.609

Notes: Moments are average transitions from employment (working at least 5 hours) to unemployment (working less than 5 hours) for individuals with the listed demographic characteristics or with wages in the previous period below the indicated quantile.

Table A9: Log wage at entrance to working life

Moment	Data	Simulated	SE data	Norm. SE diff
Mean	2.123	2.134	0.038	0.295
Variance	0.149	0.137	0.015	0.784
Mean: high background factor 1	2.137	2.151	0.045	0.316
Mean: high background factor 2	2.055	2.117	0.059	1.051
w_t below 1st quartile	0.250	0.223	0.043	0.625
w_t below median	0.500	0.487	0.051	0.251
w_t below 3rd quartile	0.750	0.813	0.045	1.391

Notes: Moments are mean and variance of wages at entrance to working life, mean of wages at the entrance to working life conditional on the indicated background characteristic, and the proportion of individuals with wages below specific quantiles of the empirical wage distribution at entrance to working life.

Table A10: Log wage regression in first differences

$$\Delta \ln(w_t) = \beta_0 + \beta_1 \Delta \ln(\kappa_t + 1) + \beta_2 d_{t-1} + \epsilon_t$$

Moment	Data	Simulated	SE data	Norm. SE diff
Diff in log years of work exp: $\Delta \ln(\kappa_t + 1)$	0.189	0.235	0.019	2.431
Lagged training dummy: d_{t-1}	-0.000	0.005	0.006	0.843

Notes: Moments are coefficients of the regression shown above, where κ_t is the observed years of full-time work experience and d_{t-1} is a dummy for spending more than 40 hours in training over the last year.

Table A11: Log wage regression on accumulated experience and lagged wages

$$\ln(w_t) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 \ln(w_{t-1}) + \beta_4 \ln(1+\kappa_t) + \beta_5 \ln(1+\kappa_{t-1}) + \epsilon_t$$

Moment	Data	Simulated	SE data	Norm. SE diff
Constant	0.424	0.400	0.032	0.773
High background factor 1: x_1	0.016	0.012	0.006	0.648
High background factor 2: x_2	-0.001	0.005	0.006	0.959
Lagged log wages: $\ln(w_{t-1})$	0.802	0.810	0.010	0.828
Log years of work exp: $\ln(1+\kappa_t)$	0.174	0.221	0.055	0.836
Lagged log years of work exp: $\ln(1+\kappa_{t-1})$	-0.139	-0.192	0.048	1.103
Variance of ϵ_t	0.053	0.055	0.002	0.907
First-order auto-corr of ϵ_t	-0.011	-0.014	0.001	3.587

Notes: Moments are coefficients of the regression shown above, where x_1 and x_2 are dummy variables indicating above median family background factors and κ_t is the observed years of full-time work experience. Sample for regression is conditional on being employed last period, since we cannot observe w_{t-1} for unemployed individuals.

Table A12: Log wage regression on lagged experience, working hours and training

$$\ln(w_t) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 \ln(1+\kappa_{t-1}) + \beta_4 1(h_{t-1} = 38) + \beta_5 1(h_{t-1} = 18) + \beta_6 d_{t-1} + \epsilon_t$$

Moment	Data	Simulated	SE data	Norm. SE diff
Constant	1.939	1.878	0.038	1.575
High background factor 1: x_1	0.059	0.062	0.022	0.135
High background factor 2: x_2	0.020	0.028	0.022	0.332
Log years of work exp: $\ln(1+\kappa_t)$	0.162	0.157	0.011	0.446
Lagged full-time dummy: $1(h_t = 38)$	0.241	0.284	0.025	1.709
Lagged part-time dummy: $1(h_t = 18)$	-0.023	-0.015	0.030	0.251
Lagged training dummy: d_{t-1}	0.089	0.120	0.013	2.433

Notes: Moments are coefficients of the regression shown above, where x_1 and x_2 are dummy variables indicating above median family background factors, κ_t is the observed years of full-time work experience, h_{t-1} indicates full-time or part-time working hours last period and d_{t-1} is a dummy for spending more than 40 hours in training over the last year.

Table A13: Log wage regression on age and family background

$$\ln(w_t) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 t + \epsilon_t$$

Moment	Data	Simulated	SE data	Norm. SE diff
Constant	2.109	2.229	0.041	2.949
High background factor 1: x_1	0.053	0.047	0.023	0.232
High background factor 2: x_2	0.015	0.020	0.021	0.274
Age: t	0.114	0.078	0.009	3.995

Notes: Moments are coefficients of the regression shown above, where x_1 and x_2 are dummy variables indicating above median family background factors and t is age in years.

Table A14: Mean yearly change in log wages given working hours at $t-1$

Moment	Data	Simulated	SE data	Norm. SE diff
Working full time at $t-1$	0.030	0.021	0.002	4.313
Working part-time at $t-1$	-0.016	0.009	0.006	4.227
Not working at $t-1$	-0.002	-0.006	0.012	0.378

Notes: Moments are coefficients of regression of mean yearly change in wages on dummies variables indicating working hours last period. Mean yearly change in wages is measured as wages this period minus wages when last observed in employment divided by number of years since last observed in employment. It is therefore observed for any individual who has been employed in at least one previous period.

Table A15: Other moments in log wages conditional on full-time work

Moment	Data	Simulated	SE data	Norm. SE diff
Mean log wages	2.603	2.597	0.011	0.568
w_t below 1st decile	0.100	0.102	0.006	0.438
w_t below 1st quartile	0.250	0.258	0.010	0.787
w_t below median	0.500	0.523	0.013	1.793
w_t below 3rd quartile	0.750	0.763	0.011	1.139
w_t below 9th decile	0.900	0.892	0.007	1.164

Notes: Moments are mean log wages for individuals working full-time (measured as working more than 20 hours a week) and the proportion of individuals with wages below specific quantiles of the empirical wage distribution of full-time workers.

Table A16: Other moments in log wages conditional on part-time work

Moment	Data	Simulated	SE data	Norm. SE diff
Mean log wages	2.382	2.342	0.020	2.019
w_t below 1st decile	0.100	0.083	0.009	1.934
w_t below 1st quartile	0.250	0.212	0.015	2.431
w_t below median	0.500	0.493	0.022	0.302
w_t below 3rd quartile	0.750	0.830	0.019	4.150
w_t below 9th decile	0.900	0.972	0.012	5.763

Notes: Moments are mean log wages for individuals working part-time (measured as working between 5 and 20 hours a week) and the proportion of individuals with wages below specific quantiles of the empirical wage distribution of part-time workers.

Table A17: Other moments in log wages conditional on training

Moment	Data	Simulated	SE data	Norm. SE diff
Mean log wages	2.660	2.675	0.014	1.103
w_t below 1st decile	0.100	0.093	0.009	0.784
w_t below 1st quartile	0.250	0.268	0.015	1.278
w_t below median	0.500	0.547	0.019	2.545
w_t below 3rd quartile	0.750	0.764	0.016	0.887
w_t below 9th decile	0.900	0.871	0.010	3.071

Notes: Moments are mean log wages for individuals in training (measured as spending more than 40 hours in training over the last 12 months) and the proportion of individuals with wages below specific quantiles of the empirical wage distribution of trainees.

Table A18: Mean log wages by family background

Moment	Data	Simulated	SE data	Norm. SE diff
High background factor 1	2.552	2.554	0.014	0.082
High background factor 2	2.578	2.571	0.015	0.487

Notes: Moments are mean log wages for individuals with the indicated background characteristics.