Wages, Experience and Training of Women over the Lifecycle

DIAL Working Paper Series 3/2020

Richard Blundell, Monica Costa-Dias,

David Goll, Costas Meghir

Dynamics of Inequality Across the Lifecourse: structures and process

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 724363

NORFACE NETWORK

Wages, Experience and Training of Women over the Lifecycle

Richard Blundell[1], Monica Costas-Dias[2], David Goll[3], Costas Meghir[4]

Abstract

We investigate the role of training in reducing the gender wage gap using the UK-BHPS which contains detailed records of training. Using policy changes over an 18 year period we identify the impact of training and work experience on wages, earnings and employment. Based on a lifecycle model and using 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.

Acknowledgments

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 thankful to seminar participants at DIW/WZB Berlin for helpful comments. Financial support from the ESRC Centre for the Microeconomic Analysis of Public Policy at the IFS (ES00116) and from the European Union Horizon 2020 programme on Dynamics of Inequality Across the Life-Course (Dial, ES00142) is gratefully acknowledged. The usual disclaimer applies.

[1] Richard Blundell, University College London and Institute for Fiscal Studies, *r.blundell@ucl.ac.uk*

[2] Monica Costa-Dias, Institute for Fiscal Studies and University of Porto, *monica_d@ifs.org.uk*

[3] David Goll, University College London and Institute for Fiscal Studies, *david.goll.11@ucl.ac.uk*

[4] Costas Meghir, Institute for Fiscal Studies, Yale University, NBER, IZA, CEPR, *c.meghir@yale.edu*

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.

Human capital investments take the form of pre-labor market education, passive learning by doing while working, and active work related training. Here we focus on the latter two, taking education as given. Each of these activities responds to incentives in a different way, which poses interesting policy questions. For example, passive learning 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). Finally, the design of the subsidy may interact with the return to training in ways that may increase or reduce its return. Understanding the importance of work-related training for the growth of human capital and how it helps shape career earnings profiles is thus central to designing policy that could help reduce the earnings costs of children on women.

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 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. In addition to that we use information on the availability of training by industry and region over time and for different education groups as exogenous variation for the opportunity to train. Thus individuals of different cohorts and education groups and living in different parts of the UK face both different work incentives and training opportunities. 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, showing that it can have a role in reducing the wage gap that arises from part time work post children. Training as a way to improve human capital accumulation and earnings for individuals is found to be more important for those women who completed high school level education but did not go on to complete University. Thus the group that would benefit most from increases in training is not the college educated, even though they they receive the most work-related training, but those who leave education with high school qualifications only. Indeed policies to 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. Finally, while training can play some role in reducing the labor market costs of children to women this gap would remain quite large even after systematic training policies, particularly for college graduates. 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 a reduced form analysis to investigate how incentives related to the tax and welfare system as well as the availability of training affect, employment, hours of work and 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

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

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 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 4,298 women and 36,597 observations. We arrange them into three groups by highest level of completed education, corresponding to secondary qualifications and below, 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.

Wages are measured on a per-hour rate by dividing weekly earnings including paid overtime

[2] In the UK, these levels correspond, respectively, to GCSE qualifications and below, A-levels and equivalent, and 3-year University degree and higher.

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

	Education			
	Secondary	**High School**	**University**	**Total**
Family type (%)				
Single, no kids	13.9	19.5	22.0	17.2
Couple, no kids	36.0	33.3	35.6	34.9
Single, with kids	10.5	7.8	4.4	8.5
Couple, with kids	39.7	39.5	38.0	39.3
Nr of individuals	2,195	1,433	670	4,298
Nr of observations	17,500	13,038	6,059	36,597

by weekly hours including any overtime, with both hours' measures referring to the main job of the respondent. 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 at least 50 hours of training over the previous year. To build it, we first convert total training duration – which can be reported in months, weeks, days or hours – into hours of training assuming 8 and 4 hours in a day for full- and part-time workers, respectively. Of those women who received training in the prior 12 months, 36% report a training intensity of 50 hours or above. Of those receiving less than than 50 hours, the median training intensity is 16 hours. We exclude these training episodes from our main analysis, since they seem likely to include capture minor work-based certification programmes, such as first-aid training.[3]

Table 2 briefly describes training spells among women, by education. For all descriptives, we have excluded instances of education and training that the participants report as being unrelated to work. 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 with more education. In our analysis, the focus is on the relatively long training episodes of at least 50 hours in a year, which account to just over a third of all training spells. These are non-negligible investments, with a median length of between 120 and 140 hours per year, or between 3 and 4 full-time weeks.

Panel C and D in Table 2 narrow the sample to only those who received more than 50 hours of training over the last 12 months. When explicit fees are charged for training, these fees are paid by the employer in between 67 and 70 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. Secondary educated women are more likely to receive training at work (48%) than either High School educated women (35%) or University educated women (27%). University educated women are often trained at work, at private training centres or utilise a mix of providers, all of which are categorised as "Other" in Table 2. Around one-quarter of

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

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

	Education			
	Secondary	**High School**	**University**	**Total**
Panel A: Training rates for employed (%)				
Any training	17.2	33.3	37.0	27.3
Training >50hrs	4.7	12.6	14.1	9.6
Panel B: Median hours of training for trainees (hrs per year)				
Any training	24	36	40	32
Training >50hrs	120	140	120	120
Panel C: Where did training take place? (%)				
At work	48.9	35.1	27.1	34.9
College/university	24.7	29.1	27.7	28.0
Other	26.3	35.8	45.2	37.1
Panel D: Who paid explicit fees, if charged? (%)				
Fees paid by employer	67.7	70.2	70.5	69.9
No fees paid by employer	32.3	29.8	29.5	30.1

Notes: All panels above exclude instances of education or training spells that are not work-related. Panel A shows training rates conditional on working at least 5 hours per week. Panel C and Panel D are estimated using sample of women who report more than 50 hours of work-related training over the last 12 months.

training occurs at a university or further education college, across all three education groups.

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 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. (2016b).

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%.

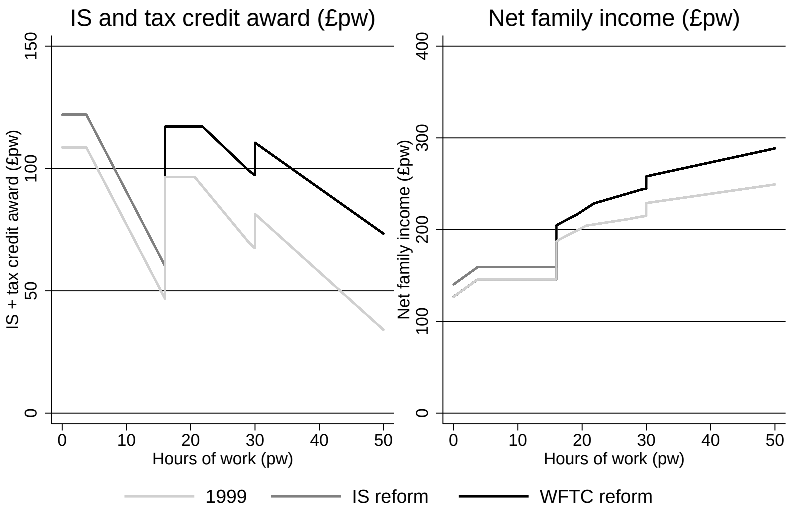
The UK benefit system is more complex. We model a range of benefits, including: Job-Seekers Allowance (JSA), which is the UK unemployment benefit; Income Support (IS), a minimum income floor that carries no work or job-search requirement; Tax Credits, a benefit for working families; Child Benefit, a 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.

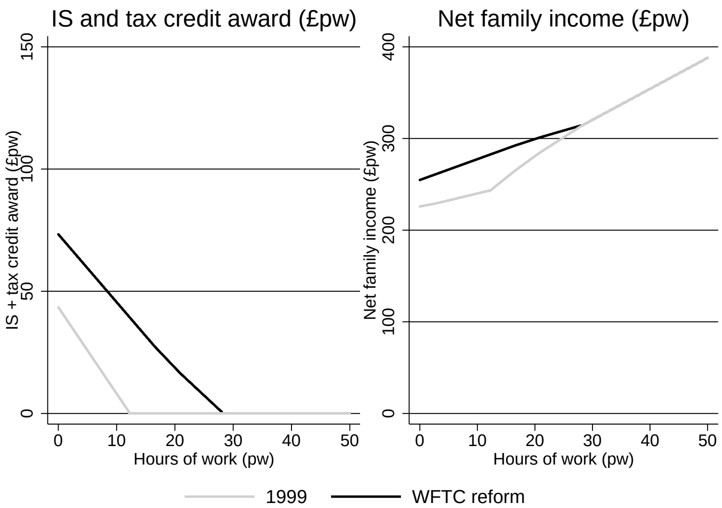
Figure 2 pictures the equivalent quantities for low-paid couples with one child aged 4 with one spouse working 40 hours per week at the 2004 minimum wage, by working hours of the second earner. Clearly, the reform had a much more modest effect on the disposable income of couples and, if anything, it 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. (2016b). 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.

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

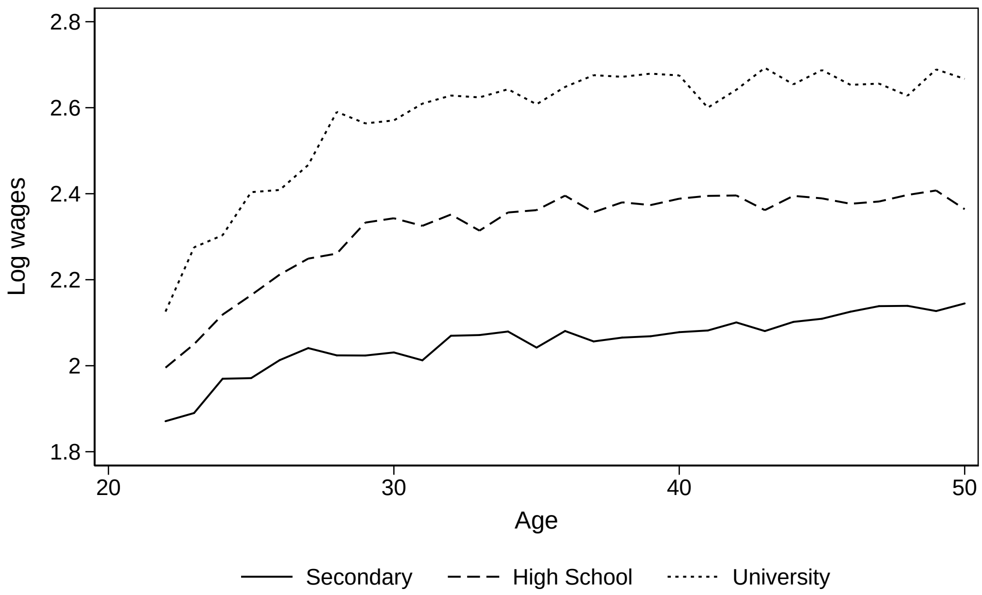


Notes: From Blundell et al. (2016b). Simulations from FORTAX for couple of one child aged 4, not paying housing rents or childcare, both spouses earning the 2004 minimum wage, one spouse working 40 hours per week. Graph on the right pictures the IS plus TC award, graph of the left pictures the disposable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

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, by education.[4] They exhibit the typical strong gradient by education and a steep upward profile early in the working life, particularly for the higher educated group. However, the wages of women flatten out quickly during their late 20s or early 30s, coinciding with the main fertility period. The flattening is permanent after that.

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



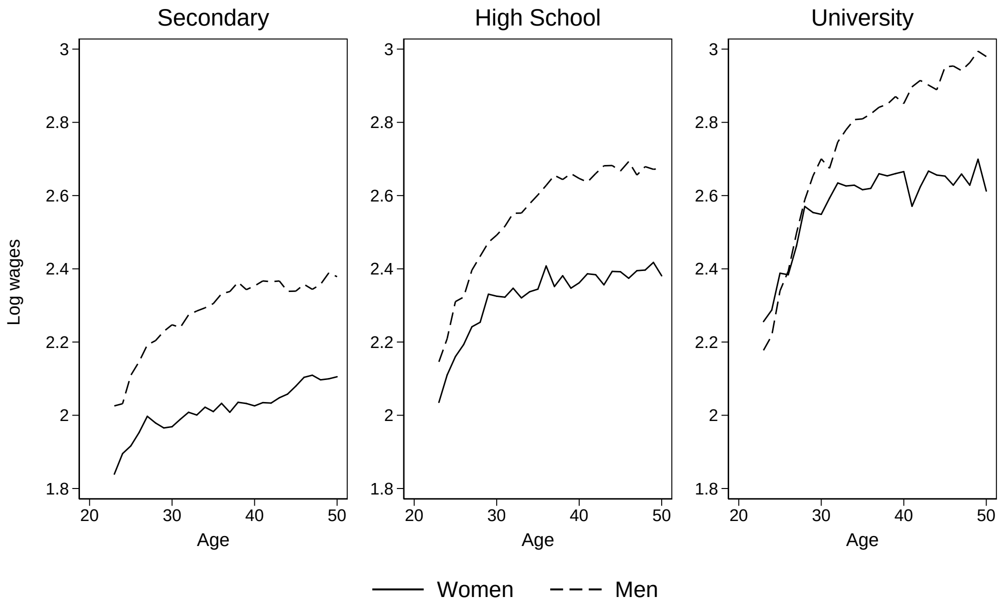
Notes: BHPS data. Wages measured on a per-hour rate, in logs. They are net of aggregate time effects and education-specific cohort effects.

Figure 4 compares wage profiles for employed men and women. Similar to women, men's wages increase with education and grow rapidly in the early years of working life. However, the wages of men, across all three education groups, continue to grow far later into working life than

[4] We use wages net of aggregate wage growth and education-specific cohort effects.

the wages of similarly educated women. 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 up around their early 30s. These pattens coincide with differences across women by education in the timing of childbirth. For instance, 51% of secondary educated women 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.

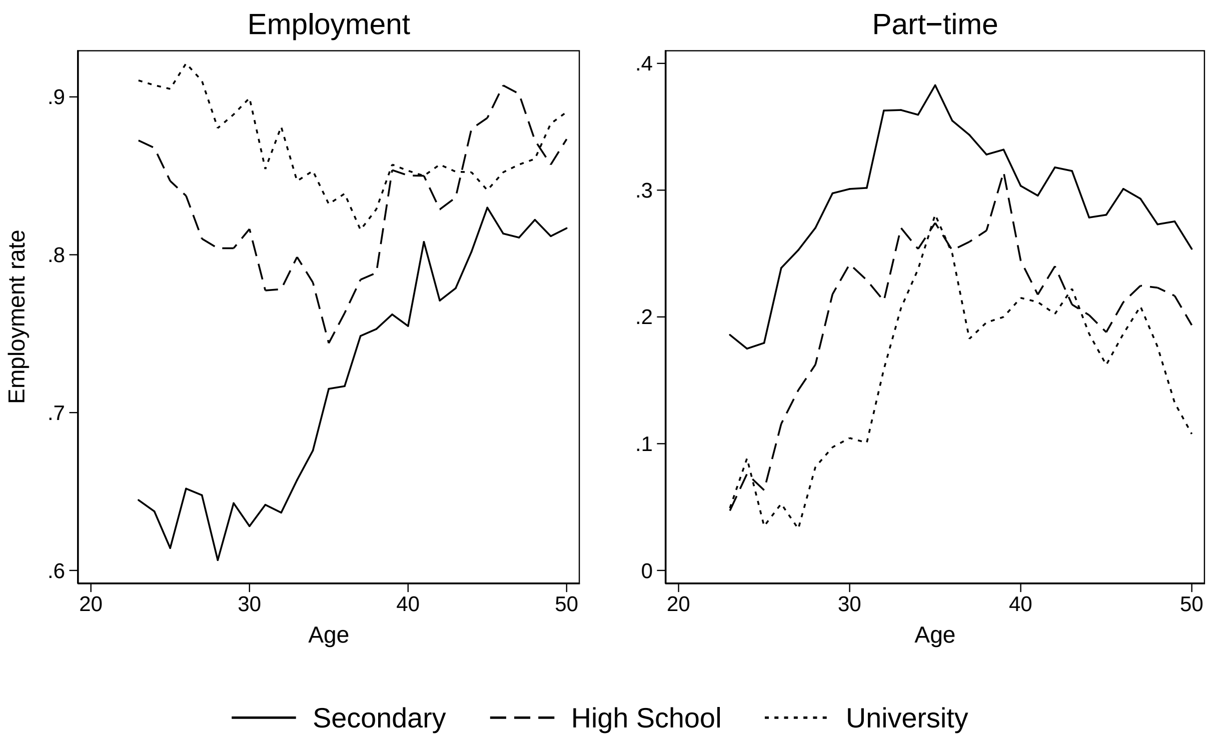
Figure 4: Average log wages of employed women and men over the life-cycle, by education



Notes: BHPS data. Wages measured on a per-hour rate, in logs. They are net of gender-specific aggregate time effects and gender- and education-specific cohort effects.

This wage profile is accompanied by strong changes in labour supply. Figure 5 shows, on 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 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 secondary and high-school qualifications. Overall, employment and full-time working hours seem strongly complementary with education.

Figure 5: Employment and working hours over the life-cycle, by education



Notes: BHPS data. The graph on the left shows employment rates by age. The graph on the right shows the proportion of working women in part-time hours conditional on being employed.

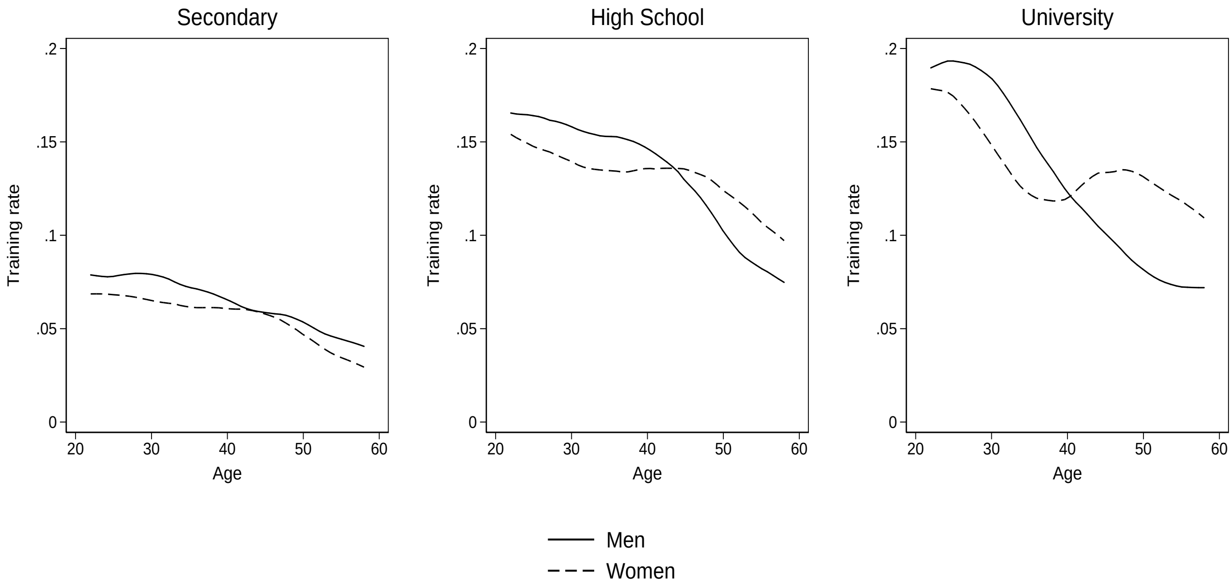
Blundell et al. (2016b) documented these working patterns, related them to fertility episodes and quantified their consequences for the wage progression of women with different levels of completed education. What that paper did not consider, however, is how work-related training interacts with education, labour supply, work experience and wages. Here we see training as one 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 in Figure 6. 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 Secondary educated workers being much less likely to invest. This suggests that, like work experience, training is complementary with education instead of being used to compensate for the lack of academic skills. 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 High-School and University educated women. Instead, training rates peak for a second time when women in these education groups are in their late 40s and 50s, a period that coincides with many of them returning to full-time work.

Figure 7 provides further insight on the timing of training by plotting its frequency around the birth of the first child. It shows a mild downward pattern for Secondary educated women, seemingly unaffected by childbirth. In contrast, the training rates of women with higher levels of education 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 losses in human capital and earnings capacity due to career interruptions, at least among mothers with High School qualifications or more. They 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

Figure 6: Training rates over the life cycle, by gender and education

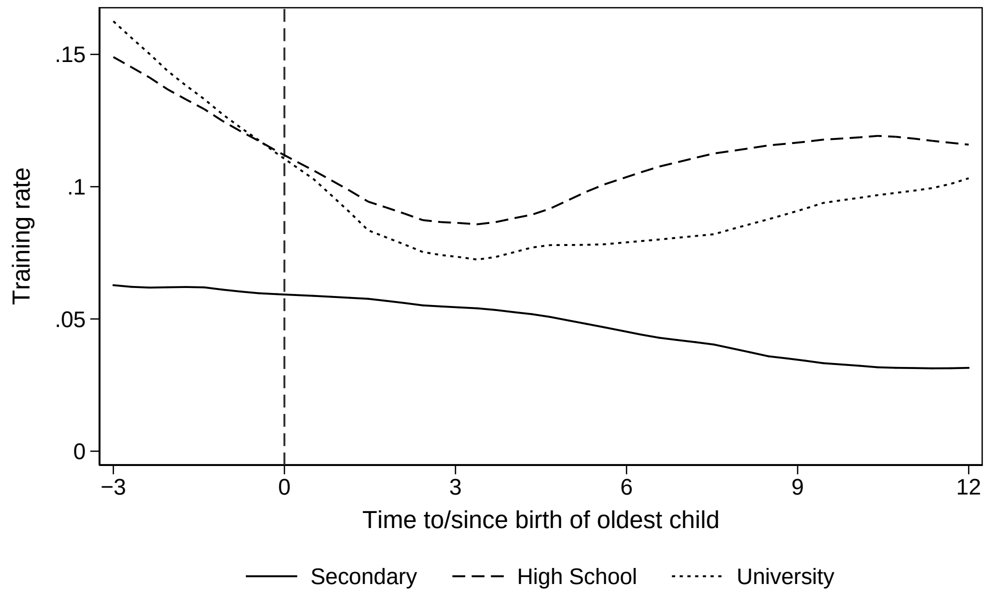


Notes: BHPS data. The training variable is an indicator for having had 50 or more hours of work-related training over the last 12 months. Lines are smoothed using a Epanechnikov kernel.

incentives targeting mothers – such as the UK Tax Credits described in the next section 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.

One observation from this discussion is that reforms in incentives to work may provide useful exogenous variation to identify the impact of training on the earnings of women. We will show

Figure 7: Training rates by time to/since birth of first child, by gender and education



Notes: BHPS data. The training variable is an indicator for having had 50 or more hours of work-related training over the year ending at the time of the interview, excluding induction training.

in section 5 that tax and welfare reforms in the UK do indeed affect the probability of training, and will then use such variation to identify the returns to training in the context of a life-cycle model.

5 Reduced form evidence

Existing studies show that the WFTC reform affected the labour supply of lone-mothers (e.g. Brewer et al. (2006), Blundell et al. (2016b)). Here we estimate a reduced-form selection model of employment, working hours and training, which are the three endogenous decisions that we will model later on. We supplement the institutional variation with geographical and time variation in training intensity among workers to show that, together, they are strong drivers of employment, working hours and the incidence of training among women.

We measure local training intensity with a Bartik instrument $\tilde{Q}$ that captures differential changes in training opportunities at the local level, driven by time-varying industry-level shocks affecting the demand for skills. Formally, it is a weighted average of the time-varying area $\times$ industry level training intensities, with weights reproducing the local industrial structure:

$$\tilde{Q}_{lt} = \sum_{j=1,\ldots,J} q_{jt}\pi_{ljt}$$

where l is travel-to-work area, of which there are 243 in Great Britain, j is industry at the 2-digit level from the SIC-92 code, of which there are 61 classes, π_{ljt} is the area l employment share of industry j at time t, and q is the industry $\times$ time index of training intensity. We construct q using the British Labour Force Survey – a large repeated cross-section representative sample of the adult population that supports more precise industry-level measures of training than the smaller BHPS.[5] The weights π are obtained from the Business Structure Database, a live register of all active firms that contains compatible information about their industry and location. We finally construct the training instrument Q as an indicator for whether $\tilde{Q}_{lt}$ is above or below its overall average. The index is merged to the BHPS dataset at the area and year level.

The reduced form model is a system of three equations, for employment, working hours and training, the two latter being conditional on working:

$$E_{it} = \mathbf{1}\left[\alpha_0 + \alpha_1 \hat{Y}^O_{it} + \alpha_2 \hat{Y}^P_{it} + \alpha_3 \hat{Y}^F_{it} + \alpha_4 Q_{it} + \alpha_5' X_{it}' + \epsilon^E_{it} \geq 0\right] \quad (1)$$

$$F_{it} = \mathbf{1}\left[\beta_0 + \beta_1 \hat{Y}^P_{it} + \beta_2 \hat{Y}^F_{it} + \alpha_3 Q_{it} + \alpha_4' X_{it}' + \epsilon^F_{it} \geq 0\right] \quad (2)$$

$$T_{it} = \mathbf{1}\left[\gamma_0 + \gamma_1 \hat{Y}^P_{it} + \gamma_2 \hat{Y}^F_{it} + \gamma_3 Q_{it} + \gamma_4' X_{it}' + \epsilon^T_{it} \geq 0\right] \quad (3)$$

In the above, the dependent variables (E, F, T) are indicators for employment, working full-time hours and having had training (50+ hours) over the last 12 months. The key variables whose explanatory power we are investigating are , the simulated disposable income of the family for

[5]The training information in the Labour Force Survey is not as detailed as that in the BHPS. In particular, we do not observe the duration of training or its purpose. Our measure is the average of an indicator variable for whether employed workers had any training during the quarter leading to the interview.

varying working hours of the woman, respectively not working, working part-time and full-time hours, $\left(\hat{Y}^O, \hat{Y}^P, \hat{Y}^F\right)$ as well as the training availability instrument Q_{it}. We use the tax system in place in period t to simulate family disposable income at the three working hours points for based on predicted wages for women.[6] We then regress the simulated incomes on indicators for family composition and age of the youngest child and net out their effects. The residuals $\hat{Y}$ are the instruments used in the regressions; they single out how policy reforms differentially affect the resources of families of different types.

The out-of-work disposable income residual $\hat{Y}^O$ is excluded from the hours and training equations, which are conditional on employment: this is the instrument that controls for selection into work. All regressions also include include in X a quadratic polynomial in age, indicators for family composition, age of the youngest child and two indices that summarise parsimoniously a set of observed variables characterising the socio-economic background of the woman.[7] We assume that the residuals $\epsilon = \left(\epsilon^O, \epsilon^F, \epsilon^T\right)$ are jointly normal with variances equal to 1 and unrestricted correlations.[8] The model is estimated by maximum likelihood.

Table 3 displays the results, focussing on the income variables and the availability of training, which are our sources of exogenous variation. Out of work income $\hat{Y}^O$ significantly reduces employment for all education groups. Part time income encourages employment again as expected. Similarly, hours of work depend on part time income and full time income as expected.

Turning now to training it tends to be lower when part time income is high, implying that training is more prevalent for full time workers. The availability of training is important for the High school and University groups, but not much for the lowest skill group. Thus overall the three exogenous instruments affect work and training decisions in the direction we expect.

[6]We use the IFS micro-simulation program Fortax. The wage equations depend on age and education.

[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.

[8]The errors can be serially correlated; ignoring it does not affect the consistency of our estimates, but can affect the standard errors.

Table 3: Estimated coefficients on instruments

	Secondary		High School		University	
	coeff	se	coeff	se	coeff	se
	Panel A: Employment					
$\hat{Y}^O$	-1.49**	(0.11)	-1.35***	(0.13)	-0.63***	(0.19)
$\hat{Y}^P$	0.70***	(0.10)	1.23***	(0.14)	0.43**	(0.20)
$\hat{Y}^F$	0.21**	(0.09)	-0.60***	(0.09)	-0.23*	(0.12)
Q	0.11***	(0.02)	0.04	(0.03)	-0.14***	(0.05)
	Panel B: Hours, conditional on employment					
$\hat{Y}^P$	-0.24**	(0.09)	0.10	(0.08)	-0.34***	(0.12)
$\hat{Y}^F$	0.17*	(0.10)	-0.42***	(0.11)	-0.19	(0.12)
Q	0.07***	(0.03)	0.07**	(0.03)	0.01	(0.05)
	Panel C: Training, conditional on employment					
$\hat{Y}^P$	-0.33**	(0.14)	-0.27***	(0.09)	-0.03	(0.11)
$\hat{Y}^F$	0.03	(0.16)	0.01	(0.10)	-0.17	(0.12)
Q	-0.03	(0.04)	0.08**	(0.03)	0.11**	(0.05)

The next question is how strong they are as instruments.

Table 4: Tests for instruments

		Secondary		High School		University	
		stat	p-val	stat	p-val	stat	p-val
(1)	Employment: all instruments	281.1	0.000	176.7	0.000	28.6	0.000
(2)	Hours: all instruments	15.9	0.001	28.8	0.000	24.7	0.000
(3)	Training: all instruments	25.3	0.000	27.3	0.000	9.1	0.027
(4)	Rank test: hours & training	7.4	0.001	4.9	0.007	2.8	0.065

Notes: (1), (2) and (3) report F-statistics under null that the coefficients associated with all instruments in the relevant equation are zero; (4) reports F-statistic for the rank as described in text.

Table 4 shows the results from a set of tests on the strength of our instruments. The first three rows detail F-statistics and p-values for the joint significance of all instruments in each of the three equations. These show that, independently across equations, the instruments are strong. Since the model relies on a common set of instruments for hours and training, we also require that they affect the two choices independently.[9] This requires that the matrix of coefficients for the instruments in the hours and training equations has rank 2. The results of our rank test are shown in row 4 of the Table.[10] The statistic of the test follows a F distribution. The test strongly rejects proportionality for Secondary and High-School educated women. However, for college graduate women the instruments seems to be overall weak, with a low F-statistic and a p-value of 6.5%.[11]

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. (2016b) 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. In turn, we simplify one aspect of that model by conditioning on (rather than considering the endogenous choice of) education.

6.1 Overview of the model and its key components

We consider the adult life of women, after completing education. Our model considers choices of labour supply, training, consumption and savings choices of women from the moment they enter the working life at the age of 19, for Secondary and High-School graduates, or 22, for University graduates. Adult life is split in two periods, the working period and the post-retirement period.

[9] We can exclude the employment regression from this argument because the excluded instrument, $\hat{Y}^O$, is a strong driver of employment, as demonstrated by the estimates in the first row of Panel A.

[10] We test the rank of the matrix of coefficients for the instruments in the hours and training equations using a likelihood ratio test. It contrasts our model with a null that imposes proportionality in the coefficients for $\left(\hat{Y}^P, \hat{Y}^F, Q\right)$ in equations (2) and (3). We then divide the statistic of the test by the degrees of freedom (equal to 2: the 6 free parameters in our model, associated with the three instruments in the two equations, are reduced to 4 parameters in the restricted model) to calculate the F-statistic shown in Table 4.

[11] There is no clear rule of thumb developed for this type of nonstandard model on the size of the F-statistic that would prevent large biases in small samples associated with weak instruments.

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. The cost of and return from training is affected by both dimensions of unobserved heterogeneity, through their effects on wages and preferences for working and training. These links are 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 50 hours over the previous year.

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 education, the stock of HC, the woman's ability type and a persistent productivity shock. HC is represented by an education-specific 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 on future wages. However, we weaken this assumption by allowing training to carry a utility cost. This we see as partly capturing, 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 dynamics of demographic variables are education-specific and the model reproduces the empirical marital sorting patterns by education.[12]

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.[13]

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 her education (s), human capital 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 mixture of normal distributions with means that depend on the unobserved heterogeneity in preferences, θ. Formally, the wage equation is

$$\begin{aligned} \ln w_t &= b_{s,0} + b_{s,1}x_1 + b_{s,2}x_2 + (\gamma_{s,0} + \gamma_{s,1}x_1 + \gamma_{s,2}x_2)\ln(\kappa_t + 1) + \omega + \upsilon_t \\ \text{where } \upsilon_t &= \rho\upsilon_{t-1} + \zeta_t \end{aligned} \tag{4}$$

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

[12] See Chiappori et al. (2009); Chiappori et al. (2018)

[13] See Shaw (2011).

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

$$y_t \quad = \quad w_t(h_t - d_t\bar{h}_d) \tag{5}$$

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

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

$$\begin{aligned} \kappa_{t+1} &= \kappa_t \left(1 - \delta_s\right) + g_s\left(h_t\right) + d_t \tau_s \\ \kappa_{\underline{t}} &= 0 \end{aligned} \tag{6}$$

$g_s(h)$ defines the human capital gain from working h hours when education attainment is s; it is set to 0 if the woman is not working, to 1 if she works full-time, and the value corresponding to part-time work is estimated. τ_s is the human capital return to training. 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, $\gamma_{s,0} + \gamma_{s,1}x_1 + \gamma_{s,2}x_2 < 1$) 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.

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 $\tilde{s}_t$, which is correlated with the education of the woman through the assortative marital matching described below, and his productivity level $\tilde{\upsilon}$. 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

$$\ln \tilde{w}_t = \tilde{b}_{\tilde{s}} + \tilde{\gamma}_{\tilde{s}} \ln (t - 18) + \widetilde{\upsilon}_t \tag{7}$$

$$\text{where } \tilde{\upsilon}_t = \tilde{\rho}_{\tilde{s}} \tilde{\upsilon}_{t-1} + \tilde{\zeta}_t. \tag{8}$$

$\widetilde{\upsilon}$ 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

$$\text{In new couples:} \quad \text{Prob}\left[\tilde{h}_t = 40 \mid t, \tilde{s}_t, m_{t-1} = 0\right] = \psi_0 \left(t, \tilde{s}_t\right) \tag{9}$$

$$\text{In existing couples:} \quad \text{Prob}\left[\tilde{h}_t = 40 \mid t, \tilde{s}_t, \tilde{h}_{t-1}, m_{t-1} = 1\right] = \psi_1 \left(t, \tilde{s}_t, \tilde{h}_{t-1}\right) \tag{10}$$

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:

$$\begin{aligned} a_{t+1} &= (1+r)a_t + y_t + m_t\tilde{h}_t\tilde{w}_t - T\left(w_t, h_t, X_t\right) \\ a_{t+1} &\geq 0 \quad \text{and} \quad a_{\underline{t}} = 0 \quad \text{and} \quad a_{\bar{t}+1} = 0 \end{aligned} \tag{11}$$

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.[14]

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 and education, whether she

[14]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.

has other children (denoted by the indicator n^k) and the age of the youngest, and whether she is married (m)

$$\text{Prob}\left[t^k = 0 \mid t, s, n^k_{t-1}, t^k_{t-1}, m_{t-1}\right] \tag{12}$$

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

$$\text{if single at } t-1: \qquad \text{Prob}\left[m_t = 1, \tilde{s} \mid t, s, m_{t-1} = 0, n^k_{t-1}\right] \tag{13}$$

$$\text{if married to man } \tilde{s} \text{ at } t-1: \qquad \text{Prob}\left[m_t = 1, \tilde{s} \mid t, s, m_{t-1} = 1, \tilde{s}, n^k_{t-1}\right] \tag{14}$$

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 Secondary and High-School graduates and at $\underline{t} = 22$ for University 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 ability and preference type, (ω, θ), and a subset of the state variables X_t that characterise her circumstances at age t:

$$u\left(c_t, h_t, d_t;\, s, \theta, \omega, X_t\right) = \frac{\left(c_t/n_t\right)^{\mu}}{\mu} \exp\left\{U\left(h_t, d_t, \theta, \omega, X_t\right)\right\}. \tag{15}$$

In the above expression, n is the equivalence scale, factoring in family size,[15] 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, ω) types and education. We decompose it into two additive terms, one relating only to working hours, U_h, and the other driving the utility cost of training, U_T:

$$U(h, d, \theta, \omega, X) = U_h(\theta, X_1) + d \times U_T(h, \omega, \theta, X_2) \tag{16}$$

with (U_h, U_T) defined as follows

$$U_h(\theta, X_1) = \begin{cases} 0 & \text{for } h = 0 \\ \theta_h + \alpha_h X_1 & \text{for } h = 18, 38 \end{cases} \tag{17}$$

$$U_T(h, \omega, \theta, X_2) = \theta_T + l(\omega) + \alpha_T X_2 + \alpha_{Th,s}. \tag{18}$$

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 the woman's education, marital status and whether she is a mother, indicators for age of youngest child in bands (0-2, 3-5, 6-10) interacted with education, and the background factors (x_1, x_2). X_2 includes indicators for whether or not she is a mother, age of youngest child in bands, and an indicator for high prevalence of training at the local level, Q, all interacted with education. The variable Q captures market variability in training opportunities. We formalise its role in the utility cost of training but could have, equivalently, introduced frictions in training availability by adding its variation to an offer rate of training. Equation (18) also includes an education-specific interaction term between working hours and training status $(\alpha_{Th,s})$.

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

Unobserved heterogeneity in preferences is modelled as a multivariate vector $\theta = (\theta_{18}, \theta_{38}, \theta_T)$ for the full-time, part-time and training components. For computational tractability we assume it can take two values, for low and high preferences for working hours and training. Heterogeneity in ability ω also takes two values, for low and high ability, and is assumed to be independent from θ. The term $l(\omega)$ captures how this ability term, which drives productivity, affects preferences for training.

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

$$V_t\left(s,\theta,\omega,X_t\right) = \max_{(a_\tau,c_\tau,h_\tau,d_\tau)_{\tau=t,\ldots,\bar{t}}} E_t\left[\sum_{\tau=t}^{\bar{t}} \beta^{\tau-t} u\left(c_\tau, h_\tau, d_\tau;\, s,\theta,\omega,X_\tau\right)\middle| X_t\right] \tag{19}$$

where maximisation is conditioned by the budget constraint (11), the female wage and human capital processes (4)-(6), the dynamics of employment and wages of a present partner (7)-(10) and the dynamics of family formation (12)-(14). The woman starts her working life as a single woman with no children and draws her initial productivity shock from a distribution that is correlated with unobserved heterogeneity in preferences θ.

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. (2016b). These include the subset of parameters defining the pre-determined family dynamics, male employment and male wages, which are estimated in that study for exactly the same specification, population and data that we use here. 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.[16]

Estimation relies on a set of 348 moments capturing various aspects of lifecycle behavior and wages.[17] We construct the simulated moments to reproduce their data counterparts, based on the simulation of 5 lifetime profiles for each of the 4,338 women who are observed in BHPS with initial conditions on education, socio-economic background and year of birth. From the resulting 21,690 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, education and socio-economic structure of the data.

Our estimation procedure uses the exogenous variation in the labour supply and training incentives from policy reforms and the local prevalence of training. Based on the reduced form results we showed that such exogenous variation was important for the two lower education groups and may play an important role in driving the results.[18] However, it is much weaker for the University group; clearly functional form restrictions and the structure of the model play an important role and more so for the higher education group.

Within the model we use the policy variation by considering two tax and benefit systems, namely the ones operating in April 1995 and 2002, with the reform happening unannounced in 2000. Our moments include pre- and post-reform 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 by 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

[16] See also Blundell et al. (1994), Attanasio and Weber (1995), Attanasio et al. (2008).

[17] The moments include for each of the three education groups: the full- and part-time employment and training rates by age, family demographics, socio-economic background, local training intensity and calendar time; 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.

[18] See the discussion in Andrews et al. (2017).

reforms of different cohorts and by allowing responses to depend on socio-economic background. Variation in local training intensity is captured by a set of moments describing its effect on the take up of training.

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

$$\sum_{\kappa=1,\ldots,K} \frac{\left(M^d_{\kappa,N} - M^s_{\kappa,S}\left(\Theta\right)\right)^2}{\operatorname{Var}\left(M^d_{\kappa,N}\right)} \tag{20}$$

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.[19] We calculate asymptotic standard errors following Gourieroux et al. (1993).

8 Parameter estimates and implications for behaviour

8.1 Wages, human capital and the return to training

Table 5 shows estimates of the female wage process by education. Estimates in panel A of the Table show that there is a strong gradient with education in both the wage rates at the start of working life ($b_{s,0}$) and the return to human capital ($\gamma_{s,0}$). In most cases, socio-economic background has a small effect on starting wages and their progression with human capital. In contrast, unobserved heterogeneity in the wage rates (ω) is important, particularly for women who leave education with only Secondary qualifications (see estimates in Panel B of the Table). Our estimates indicate that being high ability raises the wage rate of women in this education group by 22 log points compared to the average. The effect is smaller for more educated women, dropping to only 15 log points among University graduates.

[19]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.

Table 5: Wage parameters

	Education					
	Secondary		High School		University	
	value	se	value	se	value	se
Panel A: Wage coefficients						
Intercept, $\exp(b_{s,0})$	5.17	(0.04)	5.19	(0.06)	7.42	(0.06)
increment: high factor 1, $\exp(b_{s,1})$	0.00	(0.05)	0.01	(0.07)	0.06	(0.06)
increment: high factor 2, $\exp(b_{s,2})$	0.01	(0.04)	-0.19	(0.06)	0.05	(0.06)
Return to human capital, $\gamma_{s,0}$	0.13	(0.01)	0.27	(0.01)	0.28	(0.01)
increment: high factor 1, $\gamma_{s,1}$	0.07	(0.01)	0.02	(0.01)	-0.01	(0.01)
increment: high factor 2, $\gamma_{s,2}$	-0.01	(0.01)	0.03	(0.01)	-0.02	(0.01)
Panel B: Unobserved heterogeneity in ability, ω						
ω type I: wage effect	0.22	(0.01)	0.17	(0.03)	0.15	(0.02)
ω type I: probability			0.72	(0.01)		
Panel C: Distribution of persistent productivity shock ν						
Persistence of productivity	0.93	(0.01)	0.92	(0.01)	0.88	(0.01)
St. dev. of productivity innovation	0.12	(0.00)	0.15	(0.01)	0.14	(0.01)
Mean initial productivity for θ type I	0.14	(0.01)	0.11	(0.01)	0.31	(0.02)
St. dev. of initial productivity	0.15	(0.01)	0.20	(0.02)	0.22	(0.02)

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 0.9, there is a high level of wage uncertainty. There is also substantial heterogeneity in initial wages, which is correlated with heterogeneity in preferences. In particular, being Type I in preference heterogeneity θ is associate with initial wages that are 30 log points higher than the group average for University graduates, and 10 to 15 log points higher for those with less than University education.

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 much

Table 6: Human capital parameters

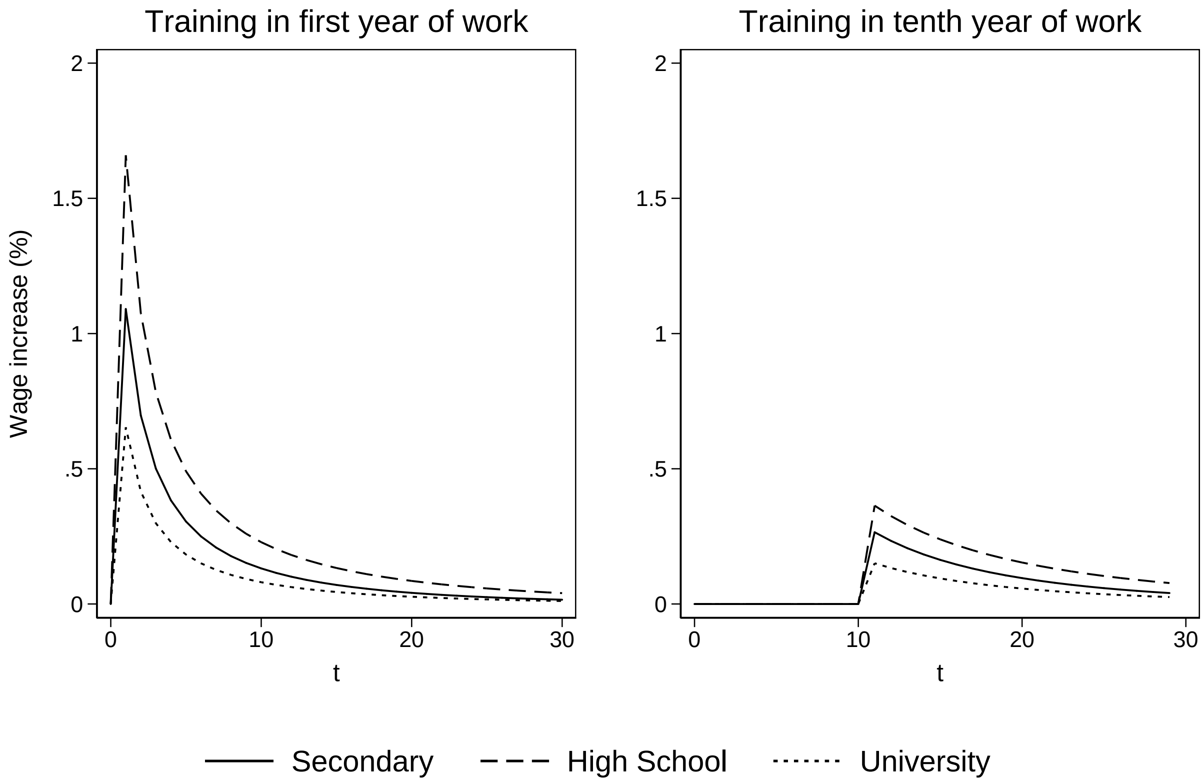
	Education					
	Secondary		High School		University	
	value	se	value	se	value	se
from training (τ)	0.17	(0.01)	0.12	(0.03)	0.05	(0.02)
from part-time work ($g(18)$)	0.04	(0.02)	0.05	(0.02)	0.12	(0.03)
depreciation rate (δ)	0.08	(0.01)	0.06	(0.01)	0.07	(0.01)

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 6 shows that estimates of these effects (τ) lie between 5 and 17 percent of the return to one year of full-time work, which is normalised to 1. Moreover, the magnitude of these effects is similar to that of the human capital returns from working part-time hours, which vary between 4 and 12 percent of the full-time return ($g(18)$ in second row of the Table). Interestingly, the human capital return to training is higher for women with Secondary qualifications and falls with education, and a similar ranking holds for the return in human capital from moving from part-time to full-time hours (which equals $1 - g(18)$).

The size of the impact of training on wages depends on the interactions between its impact on human capital (τ) and its wage returns (γ), the depreciation rate (δ), and the stock of human capital at the time of training. Figure 8 illustrates the overall short- and long-term wage effects of one episode of training taking place at different stages of the working life. Panel (a) shows the impulse response to one training episode in year 1 of working life for women in full-time hours continuously for each of the education groups; panel (b) 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

Figure 8: Wage return to one episode of training while working full-time, by education

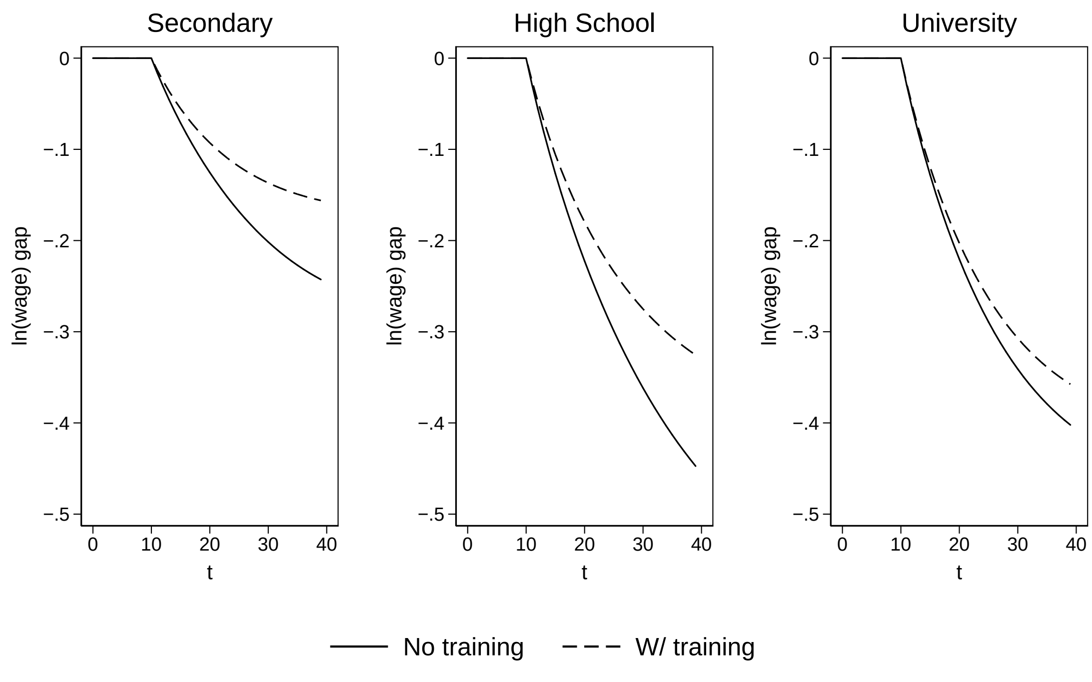


Notes: Percentage change in wage rates due to single episode of training in years 1 (LHS panel) and 10 (RHS panel) of full-time work. Agent is assumed to have no human capital at t = 0 except for that acquired through formal education and is working full-time over the entire period.

in human capital. For instance, training increases the wage rate of High School graduates in the period after training 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 also imply that High-School graduates benefit the most from training, and University graduates benefit the least. The intermediate effect for Secondary educated women reflects the combination of a comparatively high effect of training on human capital (high τ) and a low return to human capital on wages (low γ).

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 schooling, 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 and Secondary graduates 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 9: 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 9 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). The solid lines in the figure show that part-time work is associated with a large wage penalty between 0.25 and 0.45 log points, with higher losses among High School and University graduates. The dashed lines show that taking training together with part-time hours offsets almost one third of the part-time penalty among Secondary and High School educated women, but is much less effective in dealing with the loss in human capital from part-time work among University graduates. This reflects the lower human capital returns to training among University graduates.

8.2 Utility parameters and the cost of training

Tables 7 and 8 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 17 and 18. In both Tables, a positive parameter reflects higher costs of working or training.

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 7). 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. (2016a)). 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.

The bottom of Table 7 presents estimates of the distribution of unobserved heterogeneity in preferences. This is a two-point distribution with Type I representing women with comparatively high preference for working. Our estimates show that about one third of the population

Table 7: Parameters determining utility cost of working

	Coeff. (1)	St.Error (2)	Coeff. (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, Sec	0.34	(0.01)	-0.27	(0.01)
Singles, no children, HS	0.47	(0.11)	-0.33	(0.01)
Singles, no children, Univ	0.48	(0.02)	-0.32	(0.01)
Single mothers, Sec	0.39	(0.01)	-0.17	(0.01)
Single mothers, HS	0.32	(0.11)	-0.12	(0.02)
Single mothers, Univ	0.37	(0.04)	-0.05	(0.03)
Married, no children, Sec	0.25	(0.01)	-0.17	(0.01)
Married, no children, HS	0.24	(0.11)	-0.17	(0.01)
Married, no children, Univ	0.26	(0.03)	-0.15	(0.01)
Married mothers, Sec	0.27	(0.01)	-0.20	(0.01)
Married mothers, HS	0.27	(0.11)	-0.20	(0.02)
Married mothers, Univ	0.26	(0.04)	-0.18	(0.02)
Child aged 0-2, Sec	0.22	(0.02)	-0.12	(0.01)
Child aged 0-2, HS	0.22	(0.02)	-0.13	(0.01)
Child aged 0-2, Univ	0.16	(0.03)	-0.07	(0.02)
Child aged 3-5, Sec	0.23	(0.02)	-0.14	(0.01)
Child aged 3-5, HS	0.21	(0.03)	-0.13	(0.02)
Child aged 3-5, Univ	0.17	(0.03)	-0.10	(0.02)
Child aged 6-10, Sec	0.14	(0.02)	-0.10	(0.02)
Child aged 6-10, HS	0.16	(0.03)	-0.12	(0.02)
Child aged 6-10, Univ	0.11	(0.03)	-0.07	(0.02)
Spouse working	-0.10	(0.01)	0.08	(0.01)
High background factor 1	-0.02	(0.01)	-0.00	(0.01)
High background factor 2	0.00	(0.01)	-0.00	(0.01)
	Unobserved Heterogeneity in Cost of Work			
	Full-Time Employment (θ_{38})		Part-Time Employment (θ_{18})	
Type I	-0.21	(0.01)	-0.09	(0.01)
Type I: probability		0.36	(0.00)	

of women is in this group, and that the cost of working is significantly lower for them, especially in full-time hours.

Estimates for the parameters governing the utility cost of training are shown in Table 8. We

Table 8: Parameters determining utility cost of training

		Coeff.	St.Error	Coeff.	St.Error	Coeff.	St.Error
		Panel A: Utility Parameters in U_T, $(\alpha_T, \alpha_{Th,s})$					
		Secondary		High School		University	
(1)	Intercept	0.046	(0.012)	0.038	(0.011)	0.033	(0.014)
(2)	Mothers	0.011	(0.003)	0.011	(0.012)	0.004	(0.013)
(3)	Child aged 0 to 2	-0.010	(0.009)	0.005	(0.020)	0.017	(0.018)
(4)	Child aged 3 to 5	-0.007	(0.009)	0.006	(0.022)	0.006	(0.020)
(5)	Child aged 6 to 10	-0.001	(0.006)	0.005	(0.019)	0.018	(0.026)
(6)	High Bartik instrument	0.000	(0.002)	-0.005	(0.010)	-0.002	(0.008)
(7)	Part-time interaction	-0.004	(0.002)	-0.006	(0.012)	-0.011	(0.010)
		Panel B: Utility Parameters in Restricted Model U_T, $(\alpha_T, \alpha_{Th,s})$					
		Secondary		High School		University	
(8)	Intercept	0.044	(0.012)	0.037	(0.008)	0.030	(0.013)
(9)	Mothers	0.006	(0.002)	0.013	(0.007)	0.013	(0.005)
(10)	High Bartik instrument	0.000	(0.002)	-0.006	(0.009)	-0.010	(0.006)
(11)	Part-time interaction	-0.004	(0.002)	-0.002	(0.005)	-0.001	(0.006)
		Panel C: Unobserved Heterogeneity in Cost of Training					
		Correlated with Cost of Work θ_T			Correlated with Ability ω_T		
(12)	Type I	-0.030	(0.01)		-0.028	(0.01)	

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. The utility cost of training is identified from the discrepancy between the predicted take up of training (if costs were zero) and the actual take up. Our estimates show that training is also a costly activity. In the presence of frictions or other market failures justifying that firms and workers share the cost of training, these parameters would be pushed downwards in order to partly offset the assumed financial cost. Our estimates imply that, at estimated returns to training, the low training rates can only be rationalised if the cost of training to the worker is higher than the pure monetary cost of foregone earnings that we have assumed.

Panel A of Table 8 shows our main estimates for the utility cost of training. The cost is higher for mothers and varies little with age of the youngest child (rows 2 to 5). This is interesting because the training rates among working mothers with more than Secondary education drop to their lowest level during the pre-school period and recover as the child grows up. This pattern is well captured by the model, as shown in Figure 10. It also underlies the fluctuation in training over the course of life displayed in Figure 11, which the model also fits well.[20] However, the underlying parameters are not precisely estimated.

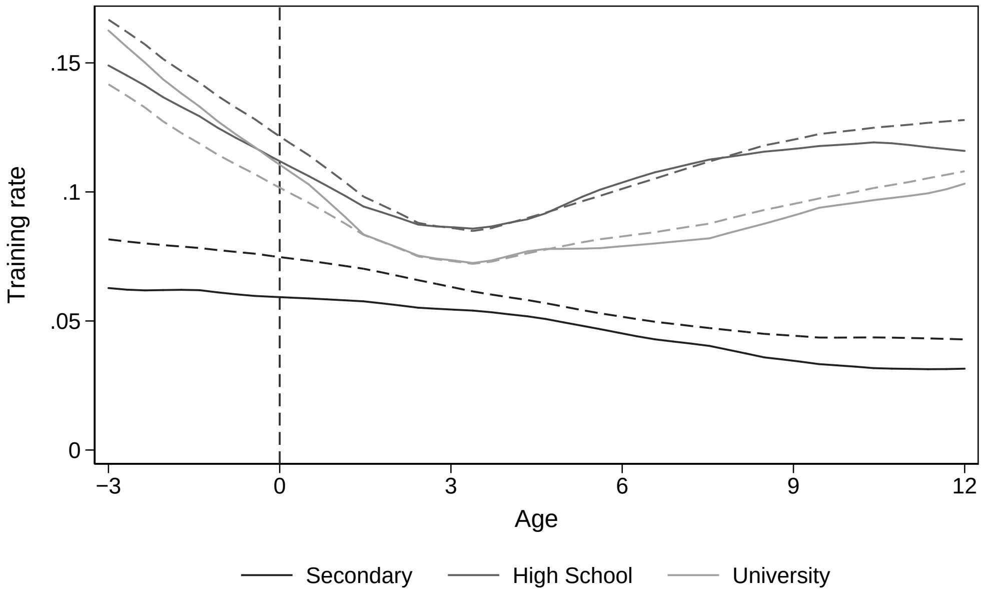
Panel B therefore contrasts our main specification with an alternative, restricted specification. In the restricted specification we impose a constant utility cost of training across all mothers, regardless of the age of their child. While this specification is less successful in replicating the lifecycle patterns of training, particularly relative to the birth of a first child, we estimate the utility cost associated with having a child more precisely. Row 9 shows the restricted estimates and standard errors. We reject the hypothesis that training costs are unaffected by motherhood at 99% confidence for Secondary and University educated mothers, and at 90% confidence for High School educated mothers.

Our model implicitly points to two additional mechanisms in addition to the utility costs 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 12 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 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

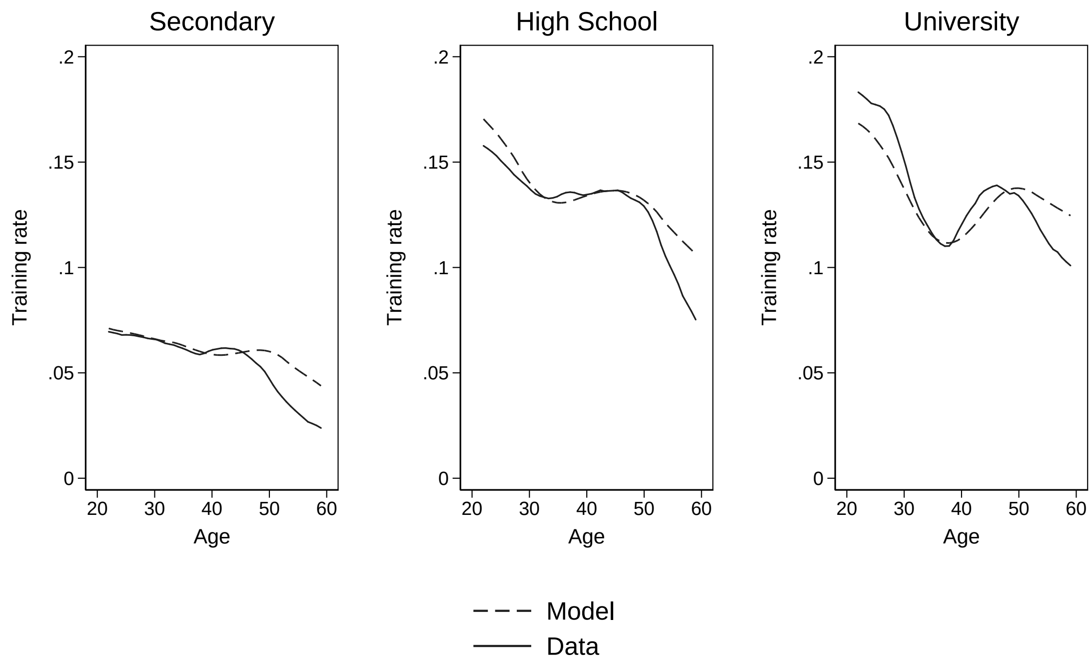
[20] The Online Appendix details evidence that the model also fits well the life-cycle profiles of wages, employment and hours, and compares the full set of data and corresponding model moments used in estimation.

Figure 10: Model versus data – Training incidence among working mothers, by time since/to birth of oldes child and maternal education



average return by a factor of 4 to 5 for most age groups. Figure 13 plots similar figures but by time to/from the birth of the first child. Clearly, the life-cycle variation is strongly associated with the dynamics of family demographics. In particular, High-School and University educated mothers of young children find training substantially more costly than they used to before giving birth, a pattern that drives the peak in the cost of training at around age 40 for them. The increase in cost around childbirth has a clear education gradient. In particular, it is noticeably larger among University graduates, who are also less likely to be entitled to work-conditional benefits. In turn, higher entitlement rates among lower educated women may be partly responsible for the milder increase in the cost of training around childbirth given its association with higher marginal tax rates, and hence lower costs of training in terms of foregone earnings. The returns to training also change around childbirth but by a much more modest amount, and then slowly recover as the child grows up.

Figure 11: Model versus data – Training incidence over the life-cycle among working women, by maternal age and education

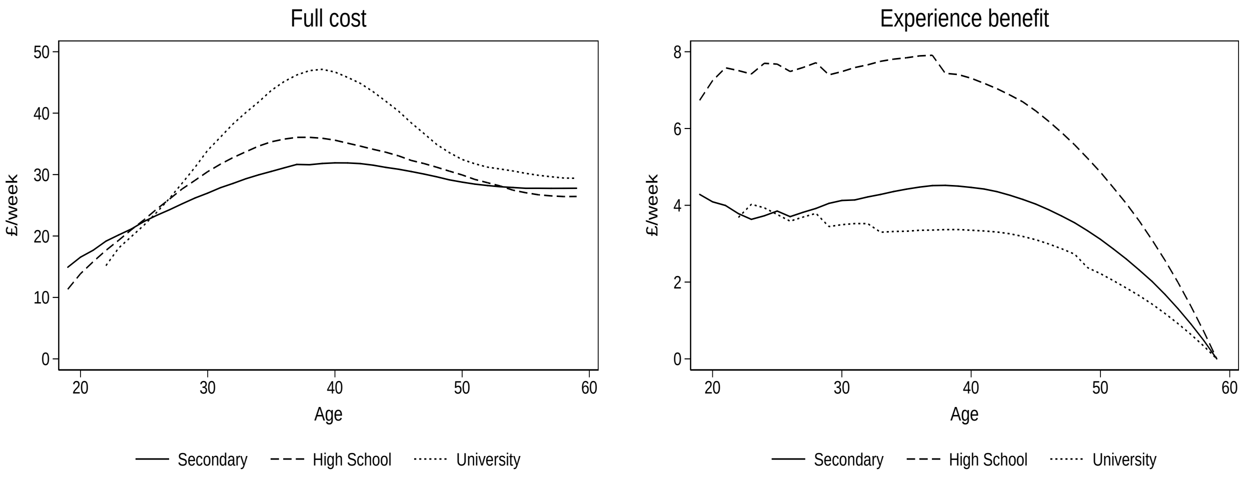


8.3 Responses of employment and training to changes in prices

We use the model to quantify responses to changes in the monetary incentives to work and train. Table 9 shows responses in employment rates (Panel A) and training rates among employed women (Panel B) to changes in the wage rates (columns 2 and 3) and in the earnings foregone while training (column 4). Column 1 provides a sense of scale by displaying the simulated levels of employment and training by education and family demographics. All simulations are run under the 2002 tax system.

Column 2 reports average responses to an anticipated 1% increase in the post tax wage rate at each age in the 23 to 50 interval, and for one year only; the small perturbation is compensated by a one-off transfer at the start of working life that keeps the expected lifetime utility constant, thereby netting out wealth effects. Column 3 shows similar responses to an unanticipated and

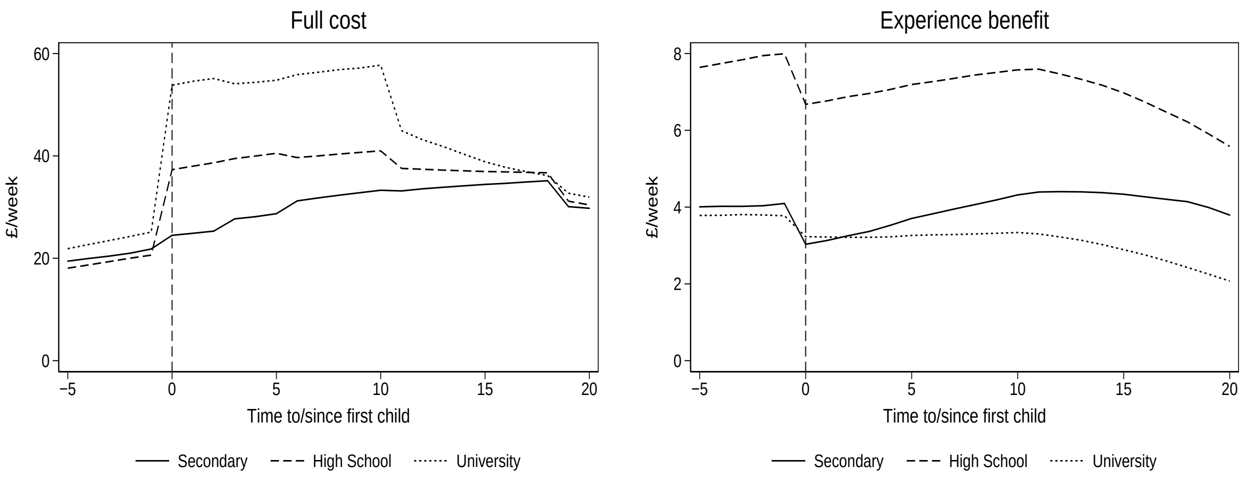
Figure 12: Monetised total cost of and experience return to training across whole population, by age and education



(a) Monetized cost of training

(b) Consumption value of extra human capital

Figure 13: 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

permanent 1% increase in the post tax wage rate starting at each age in the 23 to 50 interval. No compensation is issued in this case, and hence the figures in column 3 conflate substitution and

Table 9: Model simulations – Employment and training responses to changes in wages and the monetary cost of training

	Baseline level (%)	Responses to 1% increase in wages (pp)	
		transitory	permanent
	(1)	(2)	(3)
		Panel A: Employment	
All women	80.5	0.564	0.369
By education			
Secondary	72.2	0.698	0.449
High school	83.3	0.530	0.354
University	88.1	0.417	0.267
By family demographics			
Singles, no kids	93.0	0.489	0.322
Single mothers	59.2	1.416	0.681
Couples, no kids	93.6	0.178	0.144
Mothers in couples	72.4	0.641	0.457
		Panel B: Training conditional on employment	
All women	11.6	-0.148	-0.014
By education			
Secondary	6.0	-0.214	-0.011
High school	14.4	-0.114	-0.006
University	13.1	-0.135	-0.034
By family demographics			
Singles, no kids	11.5	-0.097	0.021
Single mothers	8.1	-0.111	-0.018
Couples, no kids	15.1	-0.133	-0.027
Mothers in couples	9.3	-0.193	-0.019

Notes: Calculations based on model simulations. Column 2 shows contemporaneous effects on employment (panel A) and training (Panel B) of an anticipated and compensated 1% transitory (single period) increase in the after tax wage rate of women. Column 3 shows effects of an unanticipated and uncompensated 1% permanent increase in wages, on the employment and training rates on the period the change in wages is first realised. Column 3 shows contemporaneous effects of an anticipated and compensated 1% transitory (single period) drop in the foregone earnings associated with training. In all case, responses are averages of effects for women aged 23 to 50.

wealth effects. Both columns report immediate responses at the time of the wage perturbation.

Overall, that transitory change in the wage rate leads to an increase in the overall employment rate of 0.56 percentage points (pp, top row in Panel A, column 2). If the change is permanent, however, the immediate impact on employment is smaller, at 0.37pp (column 3). This difference is expected if wealth effects are important, as they partially offset the substitution effect of changes in wages.[21]

While training responses to changes in the wage rates are much smaller than those of employment, they are nevertheless important given current training rates (columns 1 to 3 of Panel B). Overall, an expected transitory change in the wage rate reduces training rates contemporaneously for all groups as training becomes temporarily more costly. Its overall impact is to reduce the training rate by about 0.15pp, or by 1.3% of the existing rate (-0.148/11.6). The training responses to transitory changes in wages are larger for women with Secondary education and for mothers, who are more likely to be liquidity constrained.

The immediate training responses to permanent changes in wages, however, are much smaller but still negative. Overall, training rates drop by 0.01pp among women in work in the period when wages are first perturbed, or by about 0.1%. The permanent increase in wages lifts future returns to training, hence partly offsetting the negative impact that higher contemporaneous wages have on the propensity to train. What these figures show, however, is that the higher future returns are not sufficient to turn the training responses positive except for single women without children.

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

[21]Experience capital may reverse this relationship (as Keane and Rogerson (2012), and Keane (2015) demonstrate). In particular, young workers may respond more to a permanent than to a transitory increase in the wage rate as they foresee future additional gains from working.

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. The way one achieves revenue neutrality is relevant, since, for example, changing the tax rate to fund subsidies has its own incentive effects.

Table 10 shows model predictions of the effects of subsidized training on training rates, employment, hours, wages, savings, income and welfare. The first three rows display the effects of a £500 lump-sum subsidy for mothers of children aged 0 to 7 in training. Under our assumption of standard training units of 104 hours per year, the subsidy amounts to approximately £5 per hour. This is not a trivial subsidy, making up about 60%, 50% and 30% of the average hourly wage rate of eligible mothers with Secondary, High-School and University education respectively. However, it is more modest than other work related subsidies such as Tax Credits because it only supports a limited amount of training.

Panel A shows that training rates respond strongly to the subsidy, particularly among eligible mothers with Secondary and High-School education. As compared to the baseline, these effects

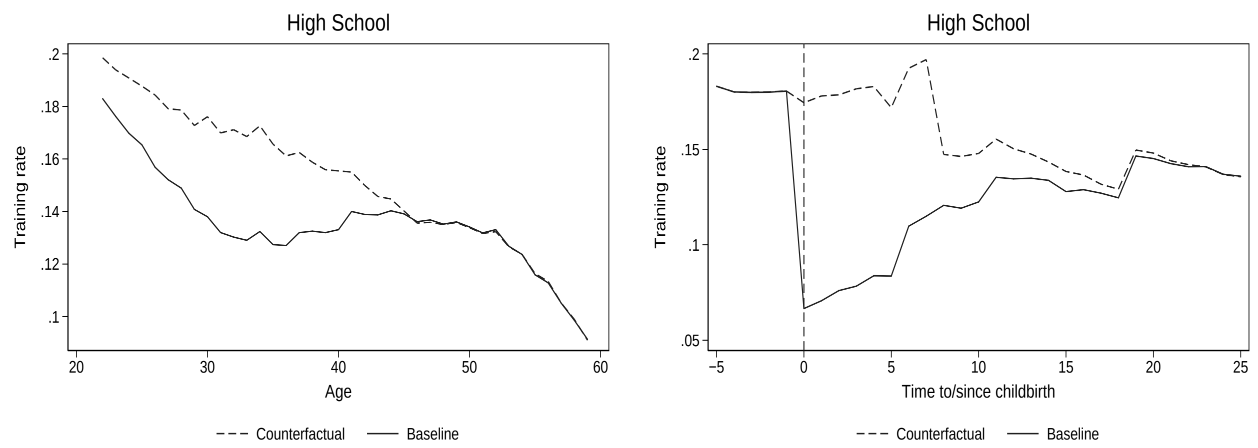
Table 10: Long term effects of subsidised training for mothers of children aged 7 and below

	Training Subsidy £500			£1,500			foregone earnings		
	Sec (1)	HS (2)	Uni (3)	Sec (4)	HS (5)	Uni (6)	Sec (7)	HS (8)	Uni (9)
	Panel A: Labor supply and training, mothers of child aged 0-7 (pp)								
Training	5.18	6.21	3.67	32.6	33.9	23.9	4.95	6.95	7.20
Employment	1.90	1.54	0.82	11.1	9.7	4.3	1.92	1.68	1.20
Full-time	-0.05	-0.16	-0.45	-1.5	-0.9	-1.6	-0.04	-0.14	-0.64
Part-time	1.94	1.70	1.28	12.6	10.6	5.9	1.96	1.83	1.84
	Panel B: Wages and assets when child aged 8 (%)								
Maternal wage rate	0.27	0.25	-0.03	1.29	1.34	0.10	0.26	0.29	-0.01
Family assets	0.49	0.12	0.05	3.87	1.64	0.26	0.74	0.45	0.27
	Panel C: Labor supply and training, mothers of child aged 8-18 (pp)								
Training	-0.31	-0.12	0.01	-0.50	-0.12	0.06	-0.32	-0.14	0.02
Employment	0.06	0.10	0.00	0.15	0.42	0.02	0.02	0.07	-0.00
Full-time	-0.12	-0.04	-0.03	-0.70	-0.59	-0.41	-0.15	-0.10	-0.08
Part-time	0.19	0.14	0.04	0.85	1.01	0.43	0.18	0.18	0.08
	Panel D: Wages and assets when child aged 19 (%)								
Maternal wage rate	0.04	0.08	-0.02	0.21	0.37	-0.05	0.03	0.08	-0.03
Family assets	0.17	0.07	-0.02	1.05	0.38	-0.29	0.21	0.15	0.02
	Panel E: Lifetime disposable income and welfare (%)								
Disposable income	0.13	0.08	-0.04	0.83	0.48	-0.33	0.10	0.07	-0.05
Consumption equiv	0.02	0.02	-0.02	0.06	-0.05	-0.40	-0.00	0.01	-0.03
	Panel F: Basic rate of income tax to fund reform:								
Change	+0.05pp			+0.6pp			+0.1pp		

Notes: Calculations based on model simulations. Columns 1-3 show the effects of a £500 yearly subsidy, while columns 4-6 and 7-9 show similar calculations for, respectively, yearly amounts of £1,500 or matching foregone earnings. Age of the child in Panels A to D refer to youngest child in the family. The consumption equivalent in Panel E is calculated at the start of working life to keep expected lifetime utility constant.

amount to an increase in their training rates of 88% and 44% percent, respectively. The subsidy is timed to coincide with the fall in training we observe around the birth of first child. Figure 14 shows, as an example, the impact of the subsidy on the lifecycle timing of training for High School educated mothers. 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 6).

Figure 14: 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 subsidy also has a sizeable effect on employment rates, which again is larger for those with less formal education. However, all additional employment is in part-time hours, and indeed full-time employment drops for all education groups and more so for the more educated. This is likely to counterbalance the positive effect that additional training has on wages through accumulated human capital. The top line of panel B shows that the cumulative effect of the additional training on the wage rates of women at the end of the eligibility period is positive but modest for women with Secondary and High-School qualifications, with both benefiting from a 0.25% increase in wages. The wages of University graduates are unaffected by the policy.

This demonstrates that the policy has only a small impact of the accumulated human capital of the mothers at the end of the eligibility period. The second line of this Panel shows that the subsidy also affects savings during the eligibility period for the families of Secondary educated women, increasing assets by 0.5% by the end of the period; here again, the effects are much milder for families of university educated women.

Since neither human capital nor assets at the end of the eligibility period are much affected, we expect the policy to have negligible dynamic effects. In fact, the figures in Panels C and D on the Table confirm that the choices and outcomes of mothers past the eligibility period are barely affected. Overall, the policy helps increase lifetime disposable income of families of Secondary and High-School educated women by a small amount, increasing their welfare by just 0.02% in consumption equivalent terms (see Panel E).[22] This is a relatively inexpensive policy that can be funded by a small increase in in the basic tax rate of 0.05pp (Panel F). Since the bulk of the cost is supported by the families of university graduates, it carries a small reduction in welfare for them.

Columns 4 to 6 of the Table show similar results for a more generous lump-sum subsidy of £1,500 per year. The additional generosity comes with a high price, requiring an increase of 0.6pp in the basic tax rate to balance the public budget. For comparison, Blundell et al. (2016b) calculations suggest that funding for the 2002 Tax Credit scheme in the UK adds 0.9pp to the basic tax rate. Our simulations show that this more generous policy succeeds in bringing up very significantly the training and employment rates of eligible mothers, although again full-time hours are negatively affected. Nevertheless, Secondary and High-school graduates see their wages increase by about 1.3% at the end of the eligibility period, but their savings

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

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

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: $\iota = 1 - \left(\frac{EV_0}{EV_1}\right)^{\frac{1}{\mu}}$.

are also higher by almost 4% among those with less education. This combination has mixed consequences for future choices as higher wages make working more attractive but higher savings push in the opposite direction. Overall, we see ineligible mothers more likely to be working but at fewer hours, and this gradually eliminates the gains in wages and savings accumulated over the eligibility period.

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 foregone earnings of trainee mothers of children aged 0 to 7. Results for this policy are displayed in columns 7 to 9 of Table 10. The cost of such policy is similar to that of the least generous training-conditional lump-sum transfer, requiring an increase of 0.1pp in the basic tax rate to balance the public accounts. Its effects are also of a very similar order of magnitude to that less generous policy, particularly for women with Secondary and High-school education. For example, those with High School education there is a positive impact on training, wages, and welfare. University graduates respond more robustly to this policy, which is especially generous towards them, by increasing employment and training rates during the eligibility period. However, they increase disproportionately part-time hours, which washes away any human capital gains from additional training and employment and results in no wage gains when their youngest child reaches 8. Overall, University graduates do not prefer this policy to the least generous lump-sum subsidy, which is preferred by the other groups.

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. We find the largest gains accrue to those who leave education after completing high school. This contrasts with the fact that most training takes place among University graduates.

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 subsidy of 1,500 pounds increases the take up of employment and training substantially, particularly for the lower education groups. It also increases lifetime disposable income by about 0.83% for the lowest education group and about half of that for the middle group. It decreases disposable income for the highest group due to increased taxation funding the program. Note, however, that welfare hardly increases for the lowest education group and is negative for the other two. We also consider other less effective 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) [23]

[23] see also Flinn et al. (2017); Lentz and Roys (2015)

References

Acemoglu, Daron and Jorn-Steffen Pischke, "Beyond Becker: Training in Imperfect Labour Markets," *The Economic Journal*, feb 1999, *109* (453), 112–142.

Adam, Stuart, James Browne, and Christopher Heady, "Taxation in the UK," in James Mirrlees, Stuart Adam, Timothy Besley, Richard Blundell, Stephen Bond, Robert Chote, Malcolm Gammie, Paul Johnson, Gareth Myles, and James Poterba, eds., *Dimensions of Tax Design: The Mirrlees Review*, Oxford University Press, 2010.

Altonji, Joseph G. and Charles R. Pierret, "Employer Learning and Statistical Discrimination," *Quarterly Journal of Economics*, February 2001, pp. 313–350.

Altonji, Joseph G and James R Spletzer, "Worker Characteristics, Job Characteristics, and the Receipt of On-the-Job Training," *Industrial and Labor Relations Review*, October 1991, pp. 58–79.

Altonji, Joseph G. and Lewis M. Segal, "Small-Sample Bias in GMM Estimation of Covariance Structures," *Journal of Business & Economic Statistics*, jul 1996, *14* (3), 353.

Andrews, Isaiah, Matthew Gentzkow, and Jesse M. Shapiro, "Measuring the Sensitivity of Parameter Estimates to Estimation Moments," *The Quarterly Journal of Economics*, jun 2017, *132* (4), 1553–1592.

Attanasio, Orazio, Hamish Low, and Virginia Sánchez-Marcos, "Explaining Changes in Female Labor Supply in a Life-Cycle Model," *American Economic Review*, aug 2008, *98* (4), 1517–1552.

Attanasio, Orazio P. and Guglielmo Weber, "Is Consumption Growth Consistent with Intertemporal Optimization? Evidence from the Consumer Expenditure Survey," *Journal of Political Economy*, dec 1995, *103* (6), 1121–1157.

Blundell, R., M. Browning, and C. Meghir, "Consumer Demand and the Life-Cycle Allocation of Household Expenditures," *The Review of Economic Studies*, jan 1994, *61* (1), 57–80.

Blundell, Richard, Lorraine Dearden, and Barbara Sianesi, "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)*, jul 2005, *168* (3), 473–512.

_ **, Luigi Pistaferri, and Itay Saporta-Eksten**, "Consumption Inequality and Family Labor Supply," *American Economic Review*, feb 2016, *106* (2), 387–435.

_ **, Monica Costa Dias, Costas Meghir, and Jonathan Shaw**, "Female Labor Supply, Human Capital, and Welfare Reform," *Econometrica*, 2016, *84* (5), 1705–1753.

Brewer, Mike, Alan Duncan, Andrew Shephard, and María José Suárez, "Did working families tax credit work? The impact of in-work support on labour supply in Great Britain," *Labour Economics*, dec 2006, *13* (6), 699–720.

Card, David, "The Causal Effect of Education on Earnings," in "Handbook of Labor Economics," Elsevier, 1999, pp. 1801–1863.

Chiappori, Pierre-André, Monica Costa Dias, and Costas Meghir, "The Marriage Market, Labor Supply, and Education Choice," *Journal of Political Economy*, oct 2018, *126* (S1), S26–S72.

_ , Murat Iyigun, and Yoram Weiss, "Investment in Schooling and the Marriage Market," *American Economic Review*, dec 2009, *99* (5), 1689–1713.

Fan, Xiaodong, Ananth Seshadri, and Christopher Taber, "Understanding Earnings, Labor Supply, and Retirement Decisions," Working Papers wp367, University of Michigan, Michigan Retirement Research Center September 2017.

Flinn, Christopher, Ahu Gemici, and Steven Laufer, "Search, Matching, and Training," *Review of Economic Dynamics*, April 2017, *25*, 260–297.

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

Heckman, James, Lance Lochner, and Ricardo Cossa, "Learning-By-Doing Vs. On-the-Job Training: Using Variation Induced by the EITC to Distinguish Between Models of Skill Formation," Technical Report jul 2002.

Keane, Michael and Richard Rogerson, "Micro and Macro Labor Supply Elasticities: A Reassessment of Conventional Wisdom," *Journal of Economic Literature*, jun 2012, *50* (2), 464–476.

Keane, Michael P., "Effects of Permanent and Transitory Tax Changes in a Life-Cycle Labor Supply Model with Human Capital," *International Economic Review*, apr 2015, *56* (2), 485–503.

Lentz, Rasmus and Nicolas Roys, "Training and Search On the Job," Technical Report nov 2015.

Shaw, Jonathan, "FORTAX: UK tax and benefit system documentation," may 2011.

Web Appendix A: Data and model moments

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. Our full data set is an unbalanced panel of 4,298 women observed for some varying period during the years 1991 to 2008. Almost 75% of these women were observed for at least 5 years and just over 40% were observed for at least 10 years. 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. Moreover, historical data provide information 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.

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. 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 (for both men and women of all education levels), 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. Education is classified in three categories: secondary or compulsory (completed by the age of 16), high school or equivalent (corresponding to A-levels or equivalent qualifications), and university (3-year degrees and above)

Tables 11 to 27 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 348 moments, which fall into the following categories:

- Mean employment, part-time hours and training conditional on demographics (Table 11, 12 and 13)

- Mean employment and training conditional on age band (Table 14 and 15)
- Transition rates from unemployment to employment conditional on demographics (Table 16)
- Transition rates from employment to unemployment conditional on demographics and wage decile (Table 17)
- Mean, variance and quantiles of log wage at entrance to working life (Table 18)
- Log wage regression in first differences on training dummy and change in log experience (Table 19)
- Log wage regression on lagged wage, family background, log years of work experience and lagged log years of work experience (Table 20)
- Log wage regression on training, experience and working status last period (Table 21)
- Log wage regression on age and family background (Table 22)
- Mean yearly change in wages conditioning on working status last period (Table 23)
- Mean wages and proportion of population with wages below pre-defined empirical wage deciles, conditional on working hours (Table 24 and 25)
- Mean training conditional on wages below pre-defined empirical wage deciles (Table 26)
- Mean log wages conditional on family background (Table 27)

All moments are constructed from the BHPS and are education-specific. Among the 348 simulated moments, 85 fall outside the 95% confidence interval for the respective data moment, but many amongst these are very similar to their BHPS counterparts.

Figures 15 to 17 show that the model fits well moments outside the estimation sample, including the wage, employment and hours profiles by education.

Figure 15: Employment over life-cycle

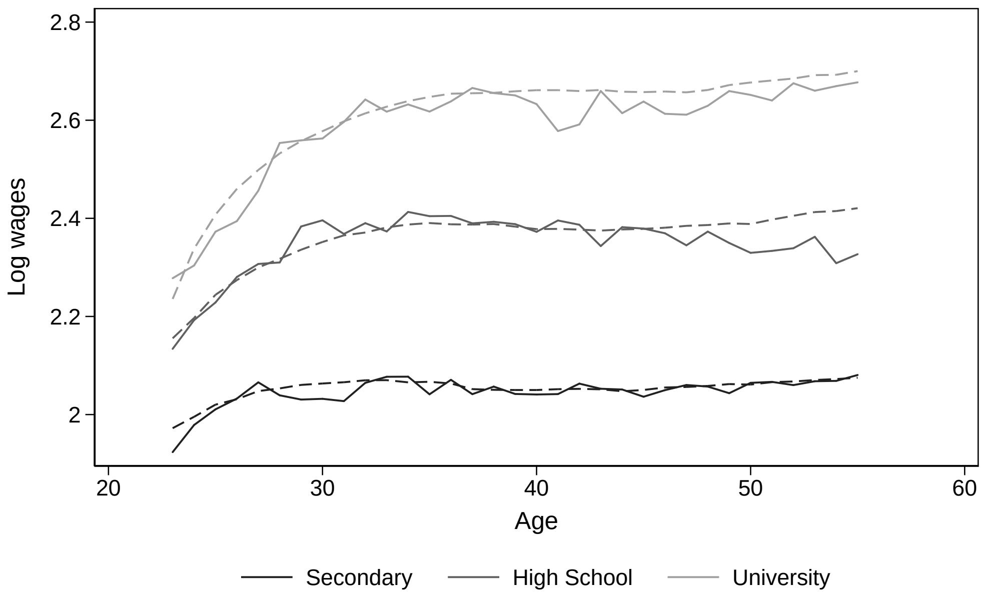


Figure 16: Employment over life-cycle

Figure 17: Employment of mothers

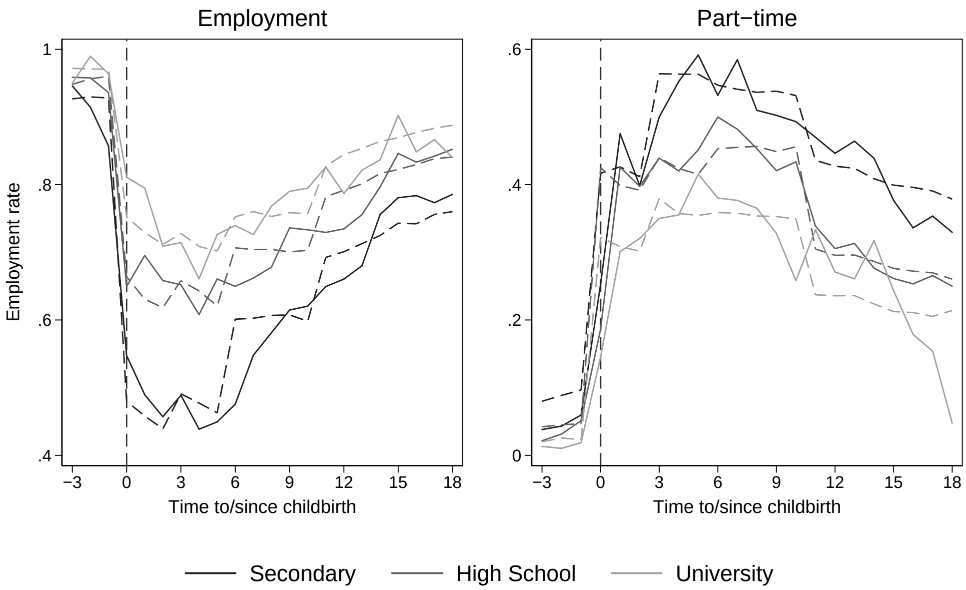


Table 11: Mean employment during working life

Moment	Data	Simulated	SE data	No. SE diff
Secondary				
All	0.733	0.717	0.011	1.552
Single women, no child	0.898	0.854	0.013	3.338
Married women, no child	0.834	0.852	0.014	1.220
Lone mothers	0.459	0.417	0.031	1.378
Married mothers	0.641	0.626	0.016	0.955
Partner working	0.767	0.763	0.012	0.332
Youngest child 0-2	0.414	0.450	0.021	1.689
Youngest child 3-5	0.525	0.465	0.022	2.764
Youngest child 6-10	0.706	0.631	0.021	3.594
Youngest child 11+	0.799	0.774	0.019	1.286
Family bkg: factor 1	0.756	0.787	0.017	1.832
Family bkg: factor 2	0.732	0.729	0.016	0.191
High Bartik instrument	0.749	0.713	0.012	2.956
Before-after (1999) difference	0.037	0.023	0.012	1.173
High School				
All	0.833	0.825	0.011	0.674
Single women, no child	0.916	0.887	0.011	2.568
Married women, no child	0.937	0.931	0.010	0.638
Lone mothers	0.674	0.647	0.045	0.595
Married mothers	0.723	0.742	0.019	0.953
Partner working	0.833	0.852	0.012	1.582
Youngest child 0-2	0.595	0.642	0.026	1.838
Youngest child 3-5	0.709	0.659	0.026	1.942
Youngest child 6-10	0.770	0.748	0.026	0.845
Youngest child 11+	0.846	0.884	0.025	1.517
Family bkg: factor 1	0.821	0.860	0.014	2.867
Family bkg: factor 2	0.837	0.833	0.013	0.257
High Bartik instrument	0.837	0.825	0.014	0.844
Before-after (1999) difference	0.020	0.018	0.012	0.186
University				
All	0.870	0.865	0.014	0.333
Single women, no child	0.938	0.963	0.011	2.297
Married women, no child	0.933	0.930	0.014	0.266
Lone mothers	0.870	0.597	0.050	5.422
Married mothers	0.771	0.779	0.028	0.264
Partner working	0.848	0.868	0.019	1.094
Youngest child 0-2	0.702	0.720	0.034	0.523
Youngest child 3-5	0.733	0.676	0.042	1.353
Youngest child 6-10	0.852	0.776	0.034	2.223
Youngest child 11+	0.898	0.885	0.039	0.323
Family bkg: factor 1	0.870	0.885	0.015	1.025
Family bkg: factor 2	0.868	0.877	0.019	0.477
High Bartik instrument	0.855	0.863	0.018	0.467
Before-after (1999) difference	0.020	0.004	0.014	1.166

Table 12: Mean part-time employment during working life

Moment	Data	Simulated	SE data	No. SE diff
Secondary				
All	0.210	0.191	0.009	2.219
Single women, no child	0.082	0.077	0.015	0.326
Married women, no child	0.171	0.161	0.013	0.789
Lone mothers	0.177	0.169	0.022	0.354
Married mothers	0.299	0.264	0.013	2.652
Partner working	0.246	0.221	0.010	2.327
Youngest child 0-2	0.219	0.161	0.015	3.719
Youngest child 3-5	0.301	0.261	0.019	2.061
Youngest child 6-10	0.331	0.321	0.020	0.464
Youngest child 11+	0.273	0.251	0.022	1.034
Family bkg: factor 1	0.176	0.177	0.013	0.103
Family bkg: factor 2	0.213	0.201	0.013	0.902
High Bartik instrument	0.201	0.191	0.011	0.928
Before-after (1999) difference	-0.019	-0.031	0.011	1.022
High School				
All	0.166	0.176	0.009	1.151
Single women, no child	0.064	0.064	0.011	0.017
Married women, no child	0.096	0.121	0.013	1.957
Lone mothers	0.164	0.108	0.033	1.694
Married mothers	0.280	0.292	0.017	0.709
Partner working	0.194	0.220	0.012	2.235
Youngest child 0-2	0.256	0.256	0.020	0.003
Youngest child 3-5	0.318	0.284	0.026	1.297
Youngest child 6-10	0.282	0.331	0.025	1.963
Youngest child 11+	0.202	0.175	0.027	1.018
Family bkg: factor 1	0.158	0.177	0.012	1.527
Family bkg: factor 2	0.181	0.169	0.013	0.913
High Bartik instrument	0.156	0.180	0.013	1.908
Before-after (1999) difference	-0.010	-0.005	0.012	0.423
University				
All	0.123	0.124	0.012	0.086
Single women, no child	0.047	0.044	0.013	0.200
Married women, no child	0.051	0.049	0.012	0.206
Lone mothers	0.089	0.000	0.029	3.115
Married mothers	0.237	0.255	0.025	0.736
Partner working	0.145	0.161	0.016	0.994
Youngest child 0-2	0.241	0.233	0.031	0.229
Youngest child 3-5	0.257	0.279	0.036	0.597
Youngest child 6-10	0.230	0.263	0.040	0.847
Youngest child 11+	0.153	0.148	0.036	0.142
Family bkg: factor 1	0.127	0.126	0.014	0.058
Family bkg: factor 2	0.113	0.124	0.015	0.754
High Bartik instrument	0.122	0.132	0.015	0.650
Before-after (1999) difference	0.007	0.006	0.015	0.027

Table 13: Mean training during working life

Moment	Data	Simulated	SE data	No. SE diff
Secondary				
All	0.042	0.047	0.002	2.000
Youngest child 0-2	0.022	0.038	0.005	3.487
Youngest child 3-5	0.018	0.032	0.004	3.142
Youngest child 6-10	0.034	0.034	0.005	0.113
Youngest child 11+	0.036	0.028	0.005	1.454
High Bartik instrument	0.045	0.047	0.004	0.459
Before-after (1999) difference	-0.020	-0.020	0.004	0.000
High School				
All	0.112	0.121	0.006	1.549
Youngest child 0-2	0.047	0.070	0.008	2.992
Youngest child 3-5	0.078	0.062	0.011	1.404
Youngest child 6-10	0.094	0.095	0.013	0.069
Youngest child 11+	0.133	0.135	0.016	0.097
High Bartik instrument	0.123	0.127	0.008	0.559
Before-after (1999) difference	-0.052	-0.052	0.007	0.000
University				
All	0.124	0.117	0.008	0.937
Youngest child 0-2	0.035	0.074	0.007	5.504
Youngest child 3-5	0.077	0.058	0.018	1.128
Youngest child 6-10	0.107	0.078	0.020	1.450
Youngest child 11+	0.123	0.147	0.023	1.014
High Bartik instrument	0.136	0.128	0.010	0.776
Before-after (1999) difference	-0.072	-0.072	0.012	0.000

Table 14: Employment by age

Moment	Data	Simulated	SE data	No. SE diff
Secondary				
25 - 29 yrs	0.633	0.628	0.021	0.205
30 - 34 yrs	0.648	0.611	0.020	1.869
35 - 39 yrs	0.739	0.645	0.018	5.071
40 - 44 yrs	0.783	0.719	0.019	3.413
45 - 49 yrs	0.818	0.772	0.017	2.642
50 - 54 yrs	0.805	0.793	0.019	0.590
55 - 59 yrs	0.756	0.788	0.022	1.429
High School				
25 - 29 yrs	0.820	0.829	0.017	0.521
30 - 34 yrs	0.791	0.769	0.019	1.122
35 - 39 yrs	0.784	0.776	0.021	0.401
40 - 44 yrs	0.848	0.816	0.020	1.589
45 - 49 yrs	0.886	0.857	0.020	1.423
50 - 54 yrs	0.888	0.860	0.024	1.158
55 - 59 yrs	0.885	0.835	0.028	1.764
University				
25 - 29 yrs	0.901	0.945	0.017	2.588
30 - 34 yrs	0.867	0.843	0.021	1.158
35 - 39 yrs	0.834	0.778	0.028	2.023
40 - 44 yrs	0.853	0.787	0.029	2.285
45 - 49 yrs	0.858	0.848	0.032	0.313
50 - 54 yrs	0.860	0.847	0.042	0.309
55 - 59 yrs	0.901	0.842	0.050	1.202

Table 15: Training by age

Moment	Data	Simulated	SE data	No. SE diff
Secondary				
25 - 29 yrs	0.044	0.039	0.006	0.749
30 - 34 yrs	0.043	0.036	0.006	1.196
35 - 39 yrs	0.043	0.037	0.006	0.951
40 - 44 yrs	0.047	0.040	0.006	1.131
45 - 49 yrs	0.048	0.056	0.006	1.256
50 - 54 yrs	0.035	0.056	0.005	3.845
55 - 59 yrs	0.018	0.038	0.004	4.587
High School				
25 - 29 yrs	0.121	0.125	0.011	0.383
30 - 34 yrs	0.104	0.099	0.011	0.463
35 - 39 yrs	0.100	0.107	0.012	0.603
40 - 44 yrs	0.117	0.109	0.013	0.594
45 - 49 yrs	0.120	0.136	0.015	1.072
50 - 54 yrs	0.112	0.147	0.015	2.295
55 - 59 yrs	0.075	0.105	0.014	2.129
University				
25 - 29 yrs	0.156	0.150	0.016	0.375
30 - 34 yrs	0.117	0.103	0.014	0.999
35 - 39 yrs	0.090	0.095	0.014	0.391
40 - 44 yrs	0.113	0.096	0.020	0.857
45 - 49 yrs	0.130	0.105	0.024	1.031
50 - 54 yrs	0.094	0.135	0.021	1.924
55 - 59 yrs	0.109	0.122	0.030	0.426

Table 16: Transition rates from unemployment to employment

Moment	Data	Simulated	SE data	No. SE diff
Secondary				
All	0.167	0.182	0.009	1.716
Single women, no child	0.173	0.228	0.020	2.774
Married women, no child	0.117	0.127	0.016	0.627
Lone mothers	0.180	0.179	0.012	0.073
High School				
All	0.243	0.241	0.016	0.096
Single women, no child	0.396	0.294	0.047	2.179
Married women, no child	0.183	0.198	0.035	0.408
Lone mothers	0.207	0.227	0.017	1.116
University				
All	0.255	0.215	0.030	1.304
Single women, no child	0.431	0.234	0.068	2.882
Married women, no child	0.294	0.130	0.082	2.006
Lone mothers	0.185	0.226	0.029	1.399

Table 17: Transition rates from employment to unemployment

Moment	Data	Simulated	SE data	No. SE diff
Secondary				
All	0.057	0.063	0.003	2.050
Single women, no child	0.033	0.041	0.003	2.845
Married women, no child	0.143	0.151	0.019	0.439
Lone mothers	0.083	0.082	0.006	0.259
w_{t-1} below 1st decile	0.099	0.109	0.009	1.054
w_{t-1} below median	0.065	0.077	0.004	3.108
w_{t-1} below 9th decile	0.056	0.064	0.003	2.593
High School				
All	0.050	0.050	0.003	0.200
Single women, no child	0.027	0.029	0.003	0.782
Married women, no child	0.090	0.106	0.017	0.954
Lone mothers	0.084	0.074	0.008	1.358
w_{t-1} below 1st decile	0.130	0.093	0.017	2.238
w_{t-1} below median	0.073	0.076	0.006	0.541
w_{t-1} below 9th decile	0.051	0.055	0.004	1.105
University				
All	0.037	0.037	0.004	0.056
Single women, no child	0.024	0.018	0.004	1.337
Married women, no child	0.033	0.137	0.019	5.526
Lone mothers	0.059	0.062	0.008	0.319
w_{t-1} below 1st decile	0.077	0.068	0.032	0.264
w_{t-1} below median	0.074	0.062	0.014	0.835
w_{t-1} below 9th decile	0.041	0.043	0.005	0.423

Table 18: Log wage at entrance to working life

Moment	Data	Simulated	SE data	No. SE diff
Secondary				
Mean	1.818	1.819	0.019	0.043
Variance	0.071	0.090	0.007	2.887
Mean: high background factor 1	1.852	1.856	0.024	0.145
Mean: high background factor 2	1.838	1.834	0.031	0.119
w_t below 1st quartile	0.249	0.238	0.031	0.352
w_t below median	0.497	0.529	0.036	0.873
w_t below 3rd quartile	0.751	0.724	0.029	0.914
High School				
Mean	1.824	1.873	0.018	2.721
Variance	0.091	0.129	0.007	5.413
Mean: high background factor 1	1.827	1.896	0.021	3.281
Mean: high background factor 2	1.816	1.884	0.029	2.350
w_t below 1st quartile	0.248	0.238	0.026	0.394
w_t below median	0.500	0.448	0.029	1.820
w_t below 3rd quartile	0.752	0.707	0.026	1.688
University				
Mean	2.112	2.110	0.025	0.083
Variance	0.118	0.134	0.011	1.387
Mean: high background factor 1	2.108	2.119	0.026	0.397
Mean: high background factor 2	2.139	2.121	0.033	0.562
w_t below 1st quartile	0.247	0.271	0.033	0.725
w_t below median	0.500	0.508	0.037	0.211
w_t below 3rd quartile	0.753	0.745	0.031	0.260

Table 19: 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	No. SE diff
Secondary				
Diff in log years of work exp: $\Delta \ln(\kappa_t + 1)$	0.097	0.134	0.022	1.750
Lagged training dummy: d_{t-1}	0.015	0.011	0.011	0.451
High School				
Diff in log years of work exp: $\Delta \ln(\kappa_t + 1)$	0.189	0.275	0.017	5.170
Lagged training dummy: d_{t-1}	-0.000	0.004	0.007	0.625
University				
Diff in log years of work exp: $\Delta \ln(\kappa_t + 1)$	0.227	0.242	0.022	0.695
Lagged training dummy: d_{t-1}	-0.006	0.009	0.009	1.697

Table 20: 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	No. SE diff
Secondary				
Constant	0.465	0.423	0.035	1.179
High background factor 1: x_1	0.029	0.035	0.006	0.875
High background factor 2: x_2	-0.005	-0.003	0.006	0.422
Lagged log wages: $\ln(w_{t-1})$	0.762	0.767	0.013	0.447
Log years of work exp: $\ln(1+\kappa_t)$	-0.032	0.095	0.067	1.904
Lagged log years of work exp: $\ln(1+\kappa_{t-1})$	0.043	-0.079	0.059	2.045
Variance of ϵ_t	0.047	0.054	0.002	3.510
First-order auto-corr of ϵ_t	-0.010	-0.012	0.001	3.433
High School				
Constant	0.364	0.355	0.030	0.282
High background factor 1: x_1	0.015	-0.000	0.006	2.397
High background factor 2: x_2	0.000	0.006	0.006	0.933
Lagged log wages: $\ln(w_{t-1})$	0.807	0.820	0.010	1.244
Log years of work exp: $\ln(1+\kappa_t)$	0.172	0.223	0.054	0.941
Lagged log years of work exp: $\ln(1+\kappa_{t-1})$	-0.139	-0.200	0.046	1.309
Variance of ϵ_t	0.050	0.053	0.002	1.844
First-order auto-corr of ϵ_t	-0.011	-0.011	0.001	0.090
University				
Constant	0.596	0.581	0.053	0.281
High background factor 1: x_1	-0.006	-0.011	0.010	0.513
High background factor 2: x_2	-0.001	-0.008	0.008	0.778
Lagged log wages: $\ln(w_{t-1})$	0.769	0.754	0.018	0.867
Log years of work exp: $\ln(1+\kappa_t)$	0.069	0.169	0.065	1.552
Lagged log years of work exp: $\ln(1+\kappa_{t-1})$	-0.056	-0.139	0.055	1.494
Variance of ϵ_t	0.045	0.046	0.002	0.688
First-order auto-corr of ϵ_t	-0.008	-0.008	0.001	0.317

Table 21: 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	No. SE diff
Secondary				
Constant	1.635	1.667	0.028	1.124
High background factor 1: x_1	0.094	0.119	0.019	1.272
High background factor 2: x_2	-0.015	-0.007	0.018	0.454
Log years of work exp: $\ln(1+\kappa_t)$	0.080	0.073	0.009	0.753
Lagged full-time dummy: $1(h_t = 38)$	0.206	0.188	0.017	1.058
Lagged part-time dummy: $1(h_t = 18)$	0.016	0.037	0.018	1.217
Lagged training dummy: d_{t-1}	0.140	0.094	0.021	2.187
High School				
Constant	1.672	1.644	0.037	0.749
High background factor 1: x_1	0.055	0.021	0.023	1.489
High background factor 2: x_2	0.021	0.010	0.023	0.477
Log years of work exp: $\ln(1+\kappa_t)$	0.162	0.165	0.011	0.277
Lagged full-time dummy: $1(h_t = 38)$	0.236	0.289	0.026	2.048
Lagged part-time dummy: $1(h_t = 18)$	-0.012	0.059	0.029	2.460
Lagged training dummy: d_{t-1}	0.080	0.079	0.014	0.009
University				
Constant	2.080	2.079	0.059	0.006
High background factor 1: x_1	-0.019	-0.022	0.035	0.101
High background factor 2: x_2	-0.011	-0.029	0.030	0.627
Log years of work exp: $\ln(1+\kappa_t)$	0.132	0.150	0.013	1.263
Lagged full-time dummy: $1(h_t = 38)$	0.277	0.221	0.046	1.212
Lagged part-time dummy: $1(h_t = 18)$	0.108	0.045	0.057	1.121
Lagged training dummy: d_{t-1}	-0.014	0.043	0.020	2.866

Table 22: 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	No. SE diff
Secondary				
Constant	1.934	1.885	0.031	1.588
High background factor 1: x_1	0.087	0.122	0.019	1.786
High background factor 2: x_2	-0.017	-0.006	0.018	0.612
Age: t	0.020	0.028	0.007	1.139
High School				
Constant	1.847	2.008	0.040	3.997
High background factor 1: x_1	0.050	0.009	0.023	1.823
High background factor 2: x_2	0.018	0.019	0.022	0.047
Age: t	0.111	0.080	0.009	3.446
University				
Constant	2.255	2.216	0.062	0.616
High background factor 1: x_1	-0.029	-0.032	0.035	0.077
High background factor 2: x_2	-0.017	-0.030	0.028	0.450
Age: t	0.095	0.110	0.013	1.142

Table 23: Mean yearly change in log wages given working hours at $t-1$

Moment	Data	Simulated	SE data	No. SE diff
Secondary				
Working full-time at $t-1$	0.018	0.010	0.002	4.466
Working part-time at $t-1$	-0.005	0.009	0.004	3.537
Not working at $t-1$	-0.005	-0.001	0.010	0.378
High School				
Working full time at $t-1$	0.030	0.023	0.002	3.680
Working part-time at $t-1$	-0.012	0.011	0.006	4.078
Not working at $t-1$	-0.006	-0.009	0.012	0.228
University				
Working full time at $t-1$	0.036	0.028	0.003	2.645
Working part-time at $t-1$	0.008	0.016	0.011	0.791
Not working at $t-1$	-0.007	0.001	0.020	0.383

Table 24: Other moments in log wages conditional on full-time work

Moment	Data	Simulated	SE data	No. SE diff
Secondary				
Mean log wages	2.090	2.075	0.010	1.635
w_t below 1st decile	0.100	0.138	0.005	7.442
w_t below 1st quartile	0.250	0.269	0.010	1.961
w_t below median	0.500	0.504	0.013	0.326
w_t below 3rd quartile	0.750	0.741	0.012	0.794
w_t below 9th decile	0.900	0.893	0.008	0.885
High School				
Mean log wages	2.326	2.357	0.012	2.679
w_t below 1st decile	0.100	0.095	0.006	0.798
w_t below 1st quartile	0.250	0.239	0.011	1.004
w_t below median	0.500	0.482	0.014	1.296
w_t below 3rd quartile	0.750	0.725	0.012	1.996
w_t below 9th decile	0.900	0.865	0.008	4.491
University				
Mean log wages	2.579	2.591	0.014	0.899
w_t below 1st decile	0.100	0.092	0.009	0.905
w_t below 1st quartile	0.250	0.252	0.014	0.127
w_t below median	0.500	0.508	0.018	0.481
w_t below 3rd quartile	0.750	0.749	0.016	0.039
w_t below 9th decile	0.900	0.880	0.010	2.042

Table 25: Other moments in log wages conditional on part-time work

Moment	Data	Simulated	SE data	No. SE diff
Secondary				
Mean log wages	1.908	1.914	0.010	0.584
w_t below 1st decile	0.100	0.138	0.007	5.585
w_t below 1st quartile	0.250	0.251	0.012	0.043
w_t below median	0.500	0.428	0.016	4.546
w_t below 3rd quartile	0.750	0.695	0.013	4.069
w_t below 9th decile	0.900	0.905	0.009	0.557
High School				
Mean log wages	2.120	2.184	0.019	3.310
w_t below 1st decile	0.100	0.065	0.009	3.870
w_t below 1st quartile	0.250	0.154	0.016	6.030
w_t below median	0.500	0.384	0.022	5.298
w_t below 3rd quartile	0.750	0.720	0.019	1.546
w_t below 9th decile	0.900	0.937	0.012	3.069
University				
Mean log wages	2.463	2.470	0.037	0.165
w_t below 1st decile	0.099	0.012	0.019	4.705
w_t below 1st quartile	0.250	0.140	0.031	3.571
w_t below median	0.501	0.546	0.037	1.231
w_t below 3rd quartile	0.751	0.864	0.033	3.421
w_t below 9th decile	0.901	0.975	0.025	3.021

Table 26: Mean training by wage decile

Moment	Data	Simulated	SE data	No. SE diff
Secondary				
w_t below 1st decile	0.052	0.015	0.015	2.342
w_t below 1st quartile	0.041	0.022	0.009	2.047
w_t below median	0.041	0.035	0.006	1.128
w_t below 3rd quartile	0.055	0.045	0.006	1.741
w_t below 9th decile	0.059	0.050	0.006	1.441
High School				
w_t below 1st decile	0.103	0.085	0.017	1.066
w_t below 1st quartile	0.113	0.102	0.012	0.981
w_t below median	0.108	0.119	0.009	1.207
w_t below 3rd quartile	0.108	0.130	0.007	3.003
w_t below 9th decile	0.122	0.135	0.007	1.787
University				
w_t below 1st decile	0.123	0.126	0.028	0.075
w_t below 1st quartile	0.167	0.135	0.021	1.565
w_t below median	0.173	0.139	0.015	2.227
w_t below 3rd quartile	0.171	0.141	0.013	2.236
w_t below 9th decile	0.164	0.144	0.012	1.629

Table 27: Mean log wages by family background

Moment	Data	Simulated	SE data	No. SE diff
Secondary				
High background factor 1	2.088	2.102	0.014	1.004
High background factor 2	2.028	2.020	0.012	0.680
High School				
High background factor 1	2.277	2.303	0.014	1.792
High background factor 2	2.305	2.338	0.016	2.053
University				
High background factor 1	2.543	2.556	0.015	0.869
High background factor 2	2.547	2.562	0.018	0.843