Wages, Experience, and Training of Women over the Life Cycle

Richard Blundell, *University College London and Institute for Fiscal Studies*

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

David Goll, *University College London and Institute for Fiscal Studies*

Costas Meghir, *Yale University, National Bureau of Economic Research, Institute of Labor Economics, Center for Economic and Policy Research, and Institute for Fiscal Studies*

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

We are very grateful to Michael Webb and William Elming for the many discussions and inputs at different stages of this project. We are also grateful to conference and seminar participants at German Institute for Economic Research (DIW)/Berlin Social Science Center (WZB Berlin), Chicago Harris School, Center for Research in Economics and Statistics (CREST; École Nationale de la Statistique et de l'Administration Économique [ENSAE Paris]), Luxembourg Institute of Socio-Economic Research (LISER), University of Bristol, University of Essex, University

[*Journal of Labor Economics*, 2021, vol. 39, no. S1]
 0734-306X/2021/39S1-0016$10.00
Submitted April 17, 2018; Accepted August 27, 2020

I. Introduction

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

In this paper we specify a model of female labor supply over the life cycle, 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 life cycle 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 recognize 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 on the basis of wage data.

Our focus is on the two human-capital-enhancing activities, working and training. Each of these activities responds to incentives in a different way, which poses interesting policy questions. For example, passive learning in work is encouraged by any factor increasing the incentives to work, such as in-work benefits (the Earned Income Tax Credit [EITC] in the United States, and the Working Families Tax Credit [WFTC] in the United Kingdom). By making working more desirable, these work-conditioned policies

of Sussex, University of York, Bank of Spain, Birkbeck, and University College London for their many useful comments. Financial support from the Economic and Social Research Council (ESRC) through the ESRC Centre for the Microeconomic Analysis of Public Policy at the Institute for Fiscal Studies (IFS; reference ES/M010147/1), from project grant ES/N015304/1, and from European Union Horizon 2020 program Dynamics of Inequality Across the Life-Course (DIAL; project 462-16-090) is gratefully acknowledged. Costas Meghir is grateful to the Cowles Foundation and the Institution for Social and Policy Studies (ISPS) at Yale for financial support. The usual disclaimer applies. Contact the corresponding author, Costas Meghir, at c.meghir@yale.edu. Information concerning access to the data used in this paper is available as supplemental material online.

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 subsidize the cost of training associated with forgone earnings (see Heckman, Lochner, and Cossa 2002). The design of the subsidy may also interact with the return to training in ways that may increase or reduce its return. Understanding the importance of work-related training for human capital and wages is thus central to designing policy that could help reduce the earnings costs of children on women. In turn, this discussion also reveals that policy reforms that change incentives to work—and to work more—may also affect training rates. In such case, they can be used to identify the effects of training on future wages. We will exploit such variation together with our model to quantify these effects.

Our basic data source is the UK British Household Panel Survey (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 United Kingdom 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 include multiple cohorts entering the labor market at different times. Each is facing a different welfare and tax system, implying changes in incentives. During their lifetimes they face reforms that affect a number of cohorts but at different ages. This generates exogenous variation in the incentives that people face at different parts of the distribution. Thus, individuals of different cohorts and education groups face both different work and training incentives. This is the key idea that underlies our identification strategy and provides the variation we need to estimate the model.

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

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

II. Data

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

The BHPS collects detailed demographic information that we use to characterize the dynamics of family formation as well as socioeconomic information mapping the education attainment, labor supply, earnings, training events, childcare expenditures, and assets of all household members 16 years old 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 their parents when 16 years old, books at home during childhood, and so on. We synthesize this information into two indices of socioeconomic background that will be used to qualify individual earnings capacity and choices.

Our observation unit is women who have completed education, who are 19–60 years old, 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 onward. Finally, we exclude women who are not UK citizens or who are ever observed claiming disability benefits. The records of women in the cleaned sample are then linked to information on a present partner and children as relevant.

Our final sample is an unbalanced panel of 7,359 women and 55,591 observations. We arrange them into three groups by highest level of completed education, corresponding to less than high school, high school qualifications

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

Table 1
Sample Size and Distribution of Family Types, by Education

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

Source.—British Household Panel Survey data for the years 1991–2008.

and equivalent, and 3-year college degree and above.[2] Table 1 shows the sample composition by family type and education of the woman.

We consider both the extensive margin and the intensive margin of labor supply and discretize the distribution of labor supply to three points: not working for pay, which we take to be zero hours in paid work per week and corresponds empirically to the cases of workers doing less than 5 weekly hours of work; working part time, which we take to be 20 hours of work per week and combines all those doing 5–20 hours; and full-time work, which we take to be 40 weekly hours and combines workers doing 21 or more hours per week. The underlying measure of weekly hours we use is for usual hours in the main job, including paid and unpaid overtime. We also consider only employees and delete the paths of workers becoming self-employed from that moment onward. More details on data selection can be found in the appendix (available online).

Wages are measured on a per-hour rate by dividing weekly earnings in the main job, including paid overtime, by weekly hours also in the main job (including any overtime, as detailed above). Since our model does not deal with macroeconomic fluctuations, we net out aggregate wage growth from the wage rates and from all monetary values of the tax and benefit system, described below in section III. 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

[2] In the United Kingdom, these levels correspond, respectively, to General Certificate of Secondary Education (GCSE) qualifications (which are acquired at the end of secondary school, at age 16) and below, A-level qualifications (obtained at the end of high school, at age 18) and equivalent, and 3-year university degree and higher.

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 more than 80% of all recorded training episodes, of which 96% happen among those in paid work at the time of the interview. The data document 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 led to any qualification.

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

Table 2 briefly describes training spells among women, by education. We show figures for our measure of training, labeled "selected training," and for a similar measure constructed on all work-related training, labeled "any training." Panel A of the table shows that training is a common event, with between 17% and 37% of employed women receiving some form of training in each year. It is also much more common among those in the middle and top education groups. Our more demanding measure of training accounts for just over 40% of all training spells. These are nonnegligible investments, with a median length of between 80 and 96 hours per year, or between two and three full-time weeks (panel B). In a working year of 48 weeks, the

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

Table 2
Training Descriptives for Women, by Education (British Household Panel Survey)

	Education			
	Less than High School	High School	University	Total
	A. Training Rates for Employed (%)			
Any training	17.1	33.4	37.0	27.4
Selected training	5.4	14.3	16.2	11.1
	B. Median Hours of Training for Trainees (Hours per Year)			
Any training	24	40	40	32
Selected training	80	96	88	88
	C. Where Did Training Take Place (Selected Training, %)			
At work	50.3	36.4	28.6	36.3
College/university	22.8	27.6	26.2	26.5
Other	26.9	35.9	45.2	37.2
	D. Who Paid Explicit Fees, If Charged (Selected Training, %)			
Fees paid by employer	69.3	71.0	71.5	70.9
No fees paid by employer	30.7	29.0	28.5	29.1

SOURCE.—British Household Panel Survey data for the years 1991–2008.
NOTE.—All figures exclude instances of education or training spells that are not work related. Training is measured only for those in work at the time of the interview. Selected training further excludes induction training and instances of training that add up to 40 or fewer hours of training in the course of 1 year.

median training duration amounts to an average of about 2 hours of job-related training per week.

Panels C and D in table 2 narrow the sample to include only trainees under our preferred definition. Women who have not completed a high school education are more likely to receive training at work (50%) than either high-school-educated women (36%) or university-educated women (28%). University-educated women are often trained at work, at private training centers. Around one-quarter of training occurs at a university or further education college across all three education groups. When explicit fees are charged for training, these fees are paid by the employer in between 69% and 72% 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.

III. Institutional Background

The personal tax and welfare benefit systems operating in the United Kingdom during the 1990s and 2000s 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 occurred between the autumn of 1999 and April 2002, which introduced the 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 modeling 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 1990s, 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 United Kingdom can be found in Adam, Browne, and Heady (2010) and Blundell et al. (2016).

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

The UK benefit system is more complex. We model a range of benefits, including the following: the Jobseekers Allowance (JSA), which is the UK unemployment benefit; IS, a minimum income floor that carries no work or job-search requirement; tax credits, an umbrella for various benefits including the WFTC, which provide additional income for families with individuals in work; the Child Benefit, a universal benefit for families with children; the Housing Benefit, which subsidizes housing costs for families who live in rented accommodation; and the Council Tax Benefit, which subsidizes 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 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 labor market attachment. The reform implemented a significant increase in the generosity and coverage of IS and the tax credits. For lone mothers, the IS award increased by more than 10% relative to wage levels over the period and remained taxed at a 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-hour threshold.

Figure 1 summarizes 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 (*A*) and disposable income (*B*) 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 shows 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

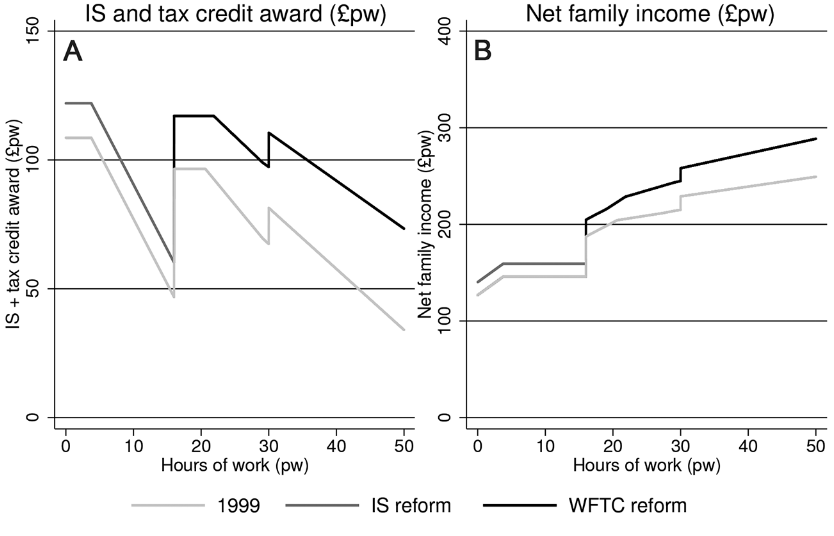

FIG. 1.—Income Support (IS) and tax credit for minimum-wage lone parent with 1 child. Simulations are from Fortax for a lone mother of one child aged 4 who is earning the 2004 minimum wage and not paying housing rents or for childcare. The *X*-axis represents hours of paid work per week. *A* shows the IS plus tax credit award and *B* shows the disposable income of the family, both in 2008 prices by working hours of the mother. From Blundell et al. (2016). WFTC = Working Families Tax Credit.

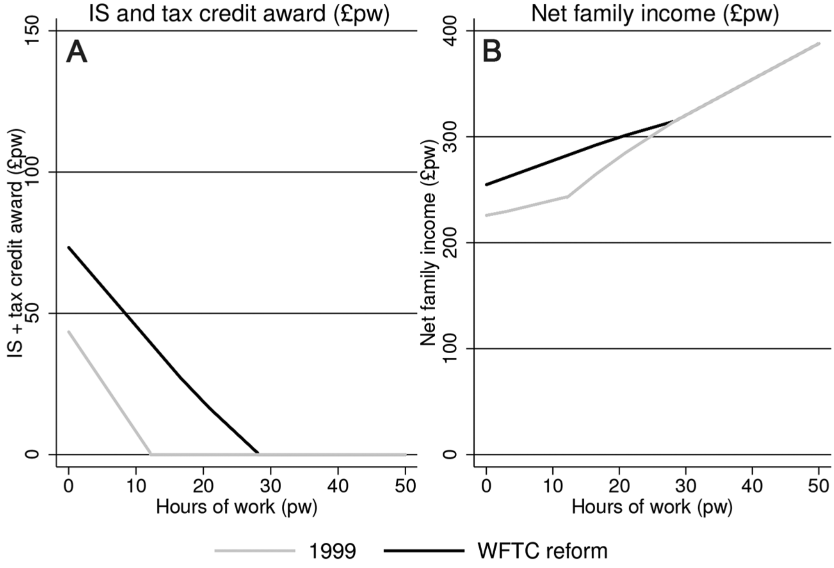

FIG. 2.—Income Support (IS) and tax credit for low-paid couple with 1 child. Simulations are from Fortax for a couple of one child aged 4 who are not paying housing rents or for childcare; both spouses are earning the 2004 minimum wage, and one spouse is working 40 hours per week. The *X*-axis represents hours of paid work per week. *A* shows the IS plus tax credit award and *B* shows the disposable income of the family, both in 2008 prices by working hours of the second earner. From Blundell et al. (2016). WFTC = Working Families Tax Credit.

if anything it reduced the incentives to work of the second earner in the family by taxing additional earned income more heavily.

IV. Life-Cycle Profiles of Employment and Training

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

The solid lines for men show wages increasing with education and growing rapidly in the early years of working life. However, the wages of men continue to grow far later into working life than the wages of similarly educated women, independent of education. The continued growth of men's wages compared with a flattening of women's wage profiles opens up a

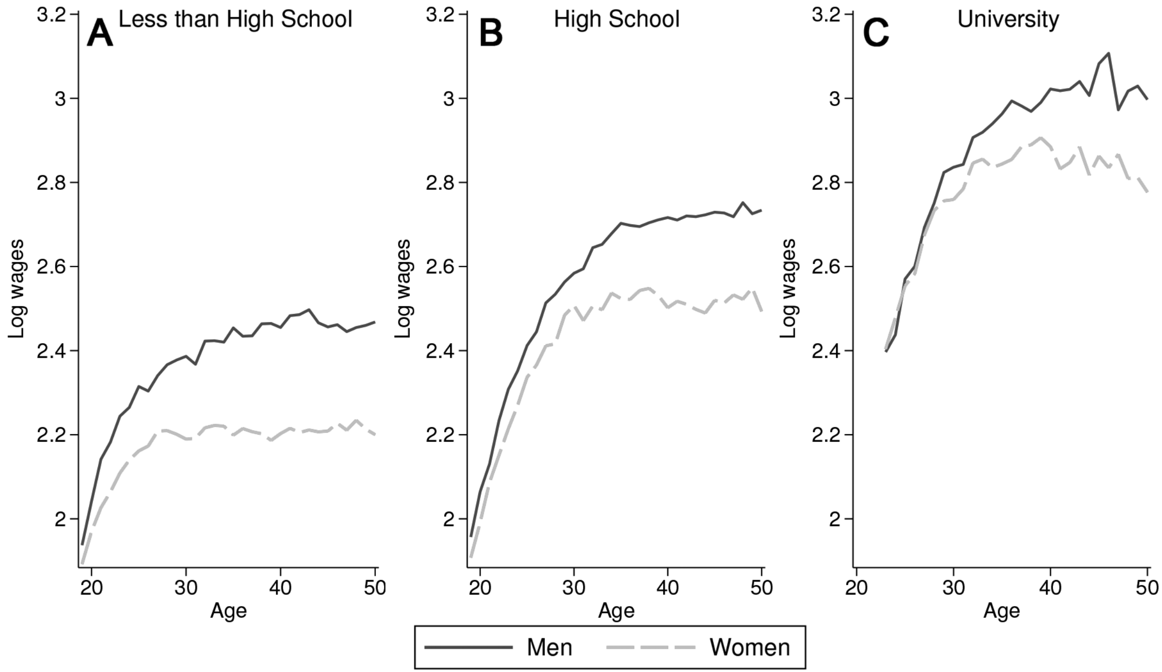

FIG. 3.—Average log wages of employed women and men over the life cycle, by education. Real wages are measured on a per-hour rate in logs. Source: British Household Panel Survey data for the years 1991–2008. A color version of this figure is available online.

gender wage gap. For low-educated women, this gap is already apparent by their early 20s. For higher-educated women, the gap opens in their late 20s. These patterns coincide with differences across women by education in the timing of childbirth. For instance, 51% of women with less than high school qualifications in our sample have at least one child by age 23. This compares to 4% of university-educated women. University-educated women reach comparable levels only at age 32, where 50% of our sample have at least one child.

This wage profile is accompanied by strong changes in labor supply. Figure 4*A* shows 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. Figure 4*B* shows the proportion working part time among women in work. The same period witnesses a strong growth in part-time hours that persists into late working life, particularly for those with high school qualifications and less. Overall, employment and full-time working hours seem strongly complementary with education.

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

FIG. 4.—Employment and working hours over the life cycle, by education. *A* shows employment rates by age and education, and *B* shows the proportion of working women in part-time hours conditional on being in work, also by age and education. Source: British Household Panel Survey data for years 1991–2008. A color version of this figure is available online.

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 nonworking periods, then it may be more prevalent among women returning to the labor market after a long fertility-related interruption than among men of similar age.

We start investigating this by contrasting the training patterns of women and men over the course of life in figure 5. Figure 5*A* of this figure shows training rates by gender and education for all individuals, independent of work status (with training for those out of work always set to zero). Several features are noteworthy. First, on-the-job training is very common among high school and university graduates. There is a clear education gradient in training, with workers with less than high school qualifications being much less likely to invest. This suggests that, like work experience, the type of training that we measure is complementary with education instead of being used to compensate for the lack of academic skills.[4] Second, despite women being much more likely to interrupt their careers during the main child-rearing period, the training rates of women and men are surprisingly similar. This holds even at the start of working life, at which point women may foresee a long career interruption linked to fertility in the near future. Third, the overall pattern of training is downward sloping, as predicted by the classical Mincer/Ben-Porath human capital framework. Noticeably, however, the slope is not monotonic for women, particularly so for the more educated. Instead, training rates peak for a second time when women in these education groups are in their 40s or early 50s, a period that coincides with many of them returning to full-time work.

Conceivably, these patterns can be mechanically driven by the life cycle of employment among women. Specifically, since female employment rates drop markedly during the main child-rearing periods and recover once children are older, lower training rates at that stage and their subsequent pickup may just reflect that movement out and back into work. Figure 5*B* refutes that hypothesis by showing similar life-cycle variation in training rates among those in work.

Figure 6 provides further insight into the timing of training by plotting its frequency around the birth of the first child. It shows that the training rates are flat around the time of first birth for women with less than high school qualifications, seemingly unaffected by childbirth. In contrast, the training rates of women with high school or university qualifications vary significantly around childbirth, first declining to reach a minimum while the child

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

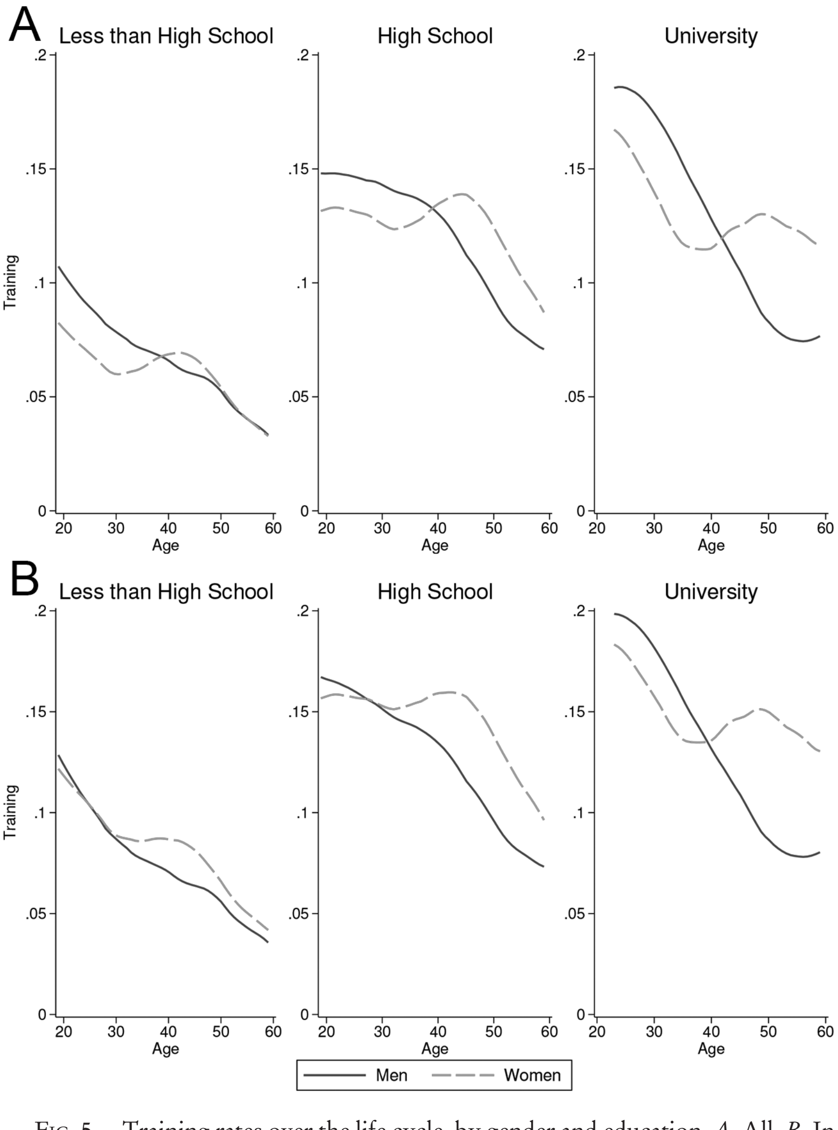

FIG. 5.—Training rates over the life cycle, by gender and education. *A*, All. *B*, In work. The training variable is an indicator for having had 40 or more hours of work-related training over the last 12 months. *A* shows training rates for the entire population, by age, gender, and education. *B* additionally conditions on working at least 5 hours per week in an usual week, which is the measure of employment used in this paper. Lines are smoothed using an Epanechnikov kernel. Source: British Household Panel Survey data for the years 1991–2008. A color version of this figure is available online.

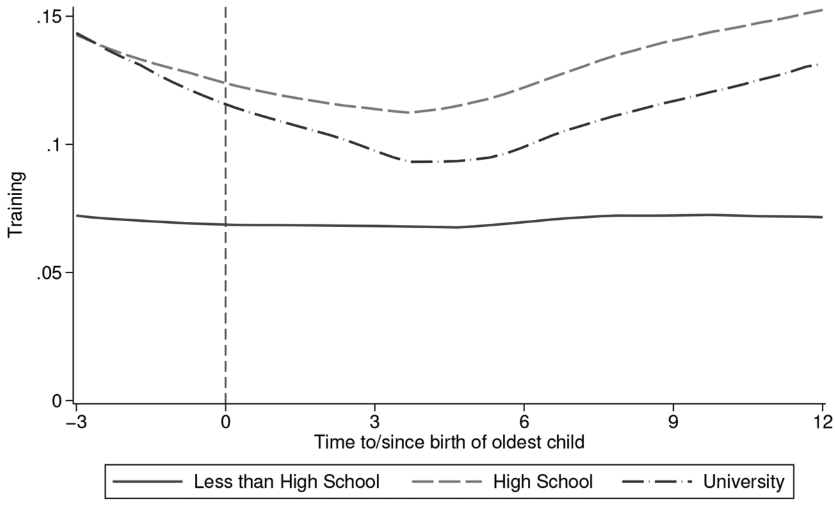

FIG. 6.—Training rates among mothers and mothers-to-be in paid work, by time to/since birth of first child and education. The training variable is an indicator for having had more than 40 hours of work-related training over the last 12 months. Source: British Household Panel Survey data for the years 1991–2008. A color version of this figure is available online.

is very young and later partly recovering as the child moves to primary and secondary schools.

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

The life-cycle patterns of training also suggest a role for public policies subsidizing working mothers that has received little attention so far (one notable exception being Heckman, Lochner, and Cossa 2002). Specifically, working incentives targeting mothers—such as the UK tax credits that we described before or the US EITC—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 increasing overall the total return to the investment. Third, by topping up low pay, the benefits may indirectly subsidize the cost of training associated with forgone earnings. And finally, the design of the subsidy may interact with the return to training among subsidized women in ways that may increase or reduce its return.

V. Training Responses to Work Incentives

One observation from the discussion in the previous section is that reforms in incentives to work may provide useful exogenous variation to identify the impact of training on the earnings of women. Existing studies have mostly focused on the impact of tax reforms on employment and hours. For instance, it has been shown that the WFTC reform affected the labor supply of lone mothers (e.g., Brewer et al. 2006; Blundell et al. 2016). Here we show that the various reforms to the tax and benefit system that happened in the United Kingdom over the 1990s and 2000s, of which the WFTC reform is a prominent example, also affected the probability that women take up training.[5] This implies that tax and benefit variation can be used to help identify the returns to training in the context of a life-cycle model.

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

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

In the above, the dependent variable T_{it} is an indicator for having had more than 40 hours of training over the last 12 months for woman i at time t, and $(\hat{Y}^{\mathrm{O}}_{it}, \hat{Y}^{\mathrm{P}}_{it}, \hat{Y}^{\mathrm{F}}_{it})$ are the respective simulated income variables. They measure family disposable income for three scenarios of female labor supply: not working (superscript O), working part-time hours (superscript P), and working full-time hours (superscript F). We use the tax system in place in period t to simulate these incomes on the basis of average female wages (by age and education) and details of the demographics of the family.[6] The term $\hat{Y}$ singles out how policy reforms differentially affect the resources of families of different types depending on the labor supply of women. We also control for a set of other covariates X, which includes time dummies, a quadratic polynomial in age, indicators for family composition, and two indices that summarize parsimoniously a set of observed variables characterizing the socioeconomic background of the woman.[7] These variables are

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

[6] We use the Institute for Fiscal Studies microsimulation program Fortax, which provides a detailed description of the taxes and benefits operating at each time period.

[7] The indices are the first and second principle components of a set of observed retrospective variables on parental background from when the woman was 16 years

Table 3
Regression of Training on Simulated Income

	Less than High School (1)	High School (2)	Degree (3)
Simulated income: no work ($\hat{Y}^{O}$)	−.000254**	−.000280	−.000178
	(.0000881)	(.000145)	(.000209)
Simulated income: part time ($\hat{Y}^{P}$)	.000606***	.000524*	.000751*
	(.000146)	(.000238)	(.000376)
Simulated income: full time ($\hat{Y}^{F}$)	−.000705***	−.000878***	−.000960***
	(.000105)	(.000150)	(.000223)
Observations	30,383	17,260	7,328
Demographic controls	Yes	Yes	Yes
Family background controls	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes
Age polynomial (second order)	Yes	Yes	Yes
F-test on instruments	20.93	15.53	8.312
F-statistic *p*-value	.00	.00	.00

SOURCE.—British Household Panel Survey data for the years 1991–2008.
NOTE.—The outcome variable is an indicator for whether the woman has taken more than 40 hours of work-related training during the year that precedes the interview. Estimates show effects of simulated family disposable income for different levels of female labor supply on the probability of taking up training. The simulations are constructed using a detailed microsimulation model for the United Kingdom. We use the tax system in place in period t to simulate these incomes on the basis of average female wages (by age and education) and details of the demographics of the family. The regressions also control for year dummies, demographic characteristics (including a quadratic in age and dummies indicating family composition), and family background (including the first two principal components drawn from a collection of variables that describe the childhood household of each individual and an indicator for whether this information is missing). The *F*-statistics at the bottom of the table test the joint significance of the three simulated income variables. Standard errors, shown in parentheses under the estimates, are clustered at the individual level.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

meant to control for variation in the disposable income variables not induced by policy reforms.

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

Estimates in table 3 are for all women, regardless of their employment status. Since the type of training that we are considering only happens among those in work, it could be thought that our estimates are effectively capturing the effects of monetary incentives to work on employment and through

of age. They summarize information on the education of both parents (five levels each), number of siblings and sibling order (dummies for no siblings, for three or more siblings, and for whether respondent is the first child), books in childhood home (three levels), and whether lived with both parents when 16 years old.

employment on training. To check this possibility, we estimated the same regression model for the restricted sample of women in paid work. Results are shown in table A1 (tables A1–A18 are available online). They demonstrate that this is not the case, particularly for women in the middle education group. For them, the *F*-statistic that we estimate is still strong (at 8.8). The effect of the simulated income variation is weakest for college graduates (*F*-statistic of 6.4) and is in between the two for the group of women with less than high school qualifications (7.3). Given the strength of the policy variation in affecting the training rates of high school graduates and the fact that training is very prevalent among women this group as well, our focus will be on this group for the remainder of the paper.

VI. The Model

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

A. Overview of the Model and Its Key Components

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

All women initiate their adult life as a single woman with no children. They are characterized 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 socioeconomic conditions of their parental home when they were 16 years old. These affect their productivity in and preferences for work. The other component of observed heterogeneity is the cohort to which women belong. Different cohorts are affected by different sequences of work incentives shaped by the policy reforms, which may affect their working and training choices.

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

During their working life, women decide in each period whether to work and for how many hours, whether to invest in training if they are working, and how much to consume today and save for the future. In the model, labor supply is discrete and can assume three values that indicate not working, working part time, and working full time, corresponding to 0, 18, and 38 weekly hours of paid work, respectively. Training is fixed at 2 hours per week, the median value of the distribution of training conditional on it exceeding 40 hours over the previous year, or one full-time working week worth of training.

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

In a competitive labor market framework with general training, workers bear the full cost of training and capture its entire return. However, firm-specific training and labor 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 labor 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 forgone earnings due to time taken away from work and that it bears a return through human capital that is reflected in future wages. However, we also allow training to carry a utility cost that may partly capture, in a reduced-form sense, incidence in the cost of training. It also captures other drivers of training, such as actual preferences, effort, or congestion in training places. In the same vein, the contribution of training to the human capital 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 as 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 of 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 human capital, then the monetary cost of training is lower and returns are larger when human capital is low. This creates stronger incentives to invest at the start of the working life—when there is also a longer period ahead to bear returns, as in a Ben-Porath model—and when returning to work after long separations, to compensate for the depreciation of skills.

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

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

B. Female Wages and Human Capital

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

$$\ln w_t = b_0 + b_1 x_1 + b_2 x_2 + (\gamma_0 + \gamma_1 x_1 + \gamma_2 x_2)\ln(\kappa_t + 1) + \omega + \upsilon_t, \tag{2}$$

where $\upsilon_t = \rho \upsilon_{t-1} + \zeta_t$.

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

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

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

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

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

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

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

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

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

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

C. 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 labor supply choices and human capital process are not endogenously modeled, we adopt a stochastic specification that captures the main features of the richer female model.

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

$$\ln \tilde{w}_t = \tilde{b}_{\tilde{s}} + \tilde{\gamma}_{\tilde{s}} \ln(t-18) + \tilde{\upsilon}_t, \tag{5}$$

$$\text{where } \tilde{\upsilon}_t = \tilde{\rho}_{\tilde{s}} \tilde{\upsilon}_{t-1} + \tilde{\zeta}_t. \tag{6}$$

The term $\tilde{v}$ is the productivity shock, initially drawn from a $\tilde{s}_t$-specific normal distribution when the couple is formed and later modeled as a $\tilde{s}$-specific autoregressive process with normal independent and identically distributed (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 labor 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 as follows:

$$\text{in new couples: Prob}\left[\tilde{h}_t = 40 \mid t, \tilde{s}_t, m_{t-1} = 0\right] = \psi_0(t, \tilde{s}_t), \tag{7}$$

$$\text{in existing couples: Prob}\left[\tilde{h}_t = 40 \mid t, \tilde{s}_t, \tilde{h}_{t-1}, m_{t-1} = 1\right] = \psi_1\left(t, \tilde{s}_t, \tilde{h}_{t-1}\right). \tag{8}$$

D. The Budget Constraint

Family resources include 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 formalized in terms of the evolution of assets:

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

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

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

[8] Fortax describes most of the UK personal taxes and benefits and how they changed over the period we model, including income tax, social security contributions, and

E. 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 labor supply, we keep track only 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 resetting t^k to zero and happens at a rate that depends on the woman's age, whether she has other children (denoted by the indicator n^k) and the age of the youngest, and whether she is married (m):

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

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

The probability that a woman marries or remains married to a man of education $\tilde{s}$ depends on her past marital circumstances, her age, whether she has children, and the education of her spouse if he is present in the previous period:

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

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

Otherwise, she will be single at time t.

F. Utility and Value Functions

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

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

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

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

---

the main subsidies for working-age families (namely, IS, JSA, tax credits, the Housing Benefit, the Council Tax Benefit, and the Child Benefit).

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

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

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

with (U_h, U_T) defined as

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

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

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

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

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

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

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

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

VII. Estimation

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

Estimation relies on a set of 139 moments capturing various aspects of life-cycle behavior and wages.[10] We construct the simulated moments to reproduce their data counterparts based on the simulation of five lifetime profiles for each of the 1,443 high-school-educated women who are observed in the BHPS with observed socioeconomic background and life histories of employment. From the resulting 7,215 profiles we select a window that exactly matches the observation window of the corresponding woman in the survey data. This way, we exactly reproduce the time, age, and socioeconomic structure of the data.

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

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

Our moments include pre- and post-2002 measures of employment, working hours, and training, which explicitly capture the variation induced by the reform. Responses to the reform are likely to vary by cohort, as they are differently exposed to the reform, and individual permanent characteristics. We exploit these interactions to identify the value of working and

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

training for future wages by explicitly modeling the differential exposure to the reforms of different cohorts and by allowing responses to depend on socioeconomic background.

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

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

where K is the total number of moments used in estimation, $M^d_{\kappa,N}$ is the estimate of moment κ from N observations of observed data, and $M^s_{\kappa,S}$ is the corresponding moment calculated on S model simulations for parameter values Θ.[11] We calculate asymptotic standard errors following Gourieroux, Monfort, and Renault (1993).

VIII. Parameter Estimates and Implications for Behavior

A. Wages, Human Capital, and the Return to Training

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

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

Training affects wages through its impact on human capital. Our estimates show the incremental effect of training over work experience for the duration of training; that is, they show how much more human capital workers gain if they choose to take time away from working and use it to train instead. The top row of table 5 shows the estimate of this effect for women

[11] It is implicit in the maximization 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 4
Wage Parameters

	Parameter Value	SE
	A. Wage Coefficients	
Intercept, exp(b_0)	6.86	.065
Increment: high factor 1, exp(b_1)	.64	.093
Increment: high factor 2, exp(b_2)	−.31	.028
Return to human capital, γ_0	.27	.004
Increment: high factor 1, γ_1	−.04	.005
Increment: high factor 2, γ_2	.03	.004
	B. Unobserved Heterogeneity in Ability, ω	
ω type I: wage effect	.24	.012
ω type I: probability	.79	.002
	C. Distribution of Persistent Productivity Shock, ν	
Persistence of productivity, ρ	.95	.002
SD of productivity innovation, ζ_t	.12	.003
SD of initial productivity, υ_0	.27	.007

at the start of working life, when they have not yet accumulated human capital from work (τ_1). Our estimate suggests that, at that stage of the working life, training increases human capital by 16% of the return to 1 year of full-time work (which is normalized to 1). We allow for more flexibility in how training affects human capital—and hence wages—by adding an interaction term with the stock of human capital (τ_2 in the second row of the table). Our estimates, however, suggest that this term is not needed.

The magnitude of the effect of training is slightly larger than the human capital return from working part-time hours, which are estimated to be 13% of the full-time return at the start of working life ($g_1(18)$ in second row of the table). We also allow for an interaction term with the stock of human capital ($g_2(18)$) and again find no evidence of the need to allow for more flexibility in how part-time hours affect human capital and wages. The only interaction of the stock of human capital that is statistically significant at conventional levels is that with full-time hours ($g_2(38)$), but even

Table 5
Parameters in the Human Capital Accumulation Process

	Parameter Value	SE
Training, τ_1	.16	.008
Training × human capital, τ_2	.00	.004
Part time, $g_1(18)$	.13	.009
Part time × human capital, $g_2(18)$	.00	.003
Full time × human capital, $g_2(38)$	−.02	.005
Depreciation rate (δ)	.08	.002

there the effect is small. Our estimate shows that one additional unit of human capital reduces the human capital return to full-time hours by 2%. Since human capital never increases beyond 12 in simulations, at the maximum this parameter is responsible for a 24% drop in the human capital return to 1 year of full-time work.

The size of the impact of training on wages depends on the interactions between its impact on human capital (determined by τ_1) and its wage return (determined by a combination of ($\gamma_0, \gamma_1, \gamma_2$) for different groups), the depreciation rate (δ), and the stock of human capital at the time of training. Figure 7 illustrates the overall short- and long-term wage effects of one episode of training taking place at different stages of the working life. Figure 7*A* shows the impulse response to one training episode in year 1 of working life for women in full-time hours; figure 7*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 in human capital. For instance, training increases the wage rate by 2% if taken in the first period of work but only by 0.5% if taken after 10 years

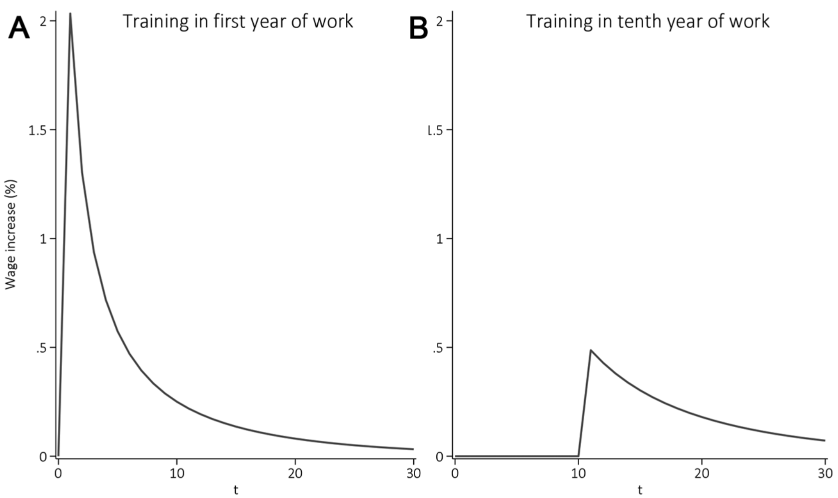

FIG. 7.—Wage return to one episode of training while working full time, by education. Shown is the percentage change in wage rates due to single episode of training in years 1 (*A*) and 10 (*B*) of full-time work. The agent is assumed to have no human capital at $t = 0$ except for that acquired through formal education and is working full time over the entire period.

of working full time. The falling returns to training with accumulated human capital is an important determinant of the timing of training in our model.

Our estimates of the wage impact of training can be compared with estimates of the impact of one additional year of education found in the broader literature once adjusted for the relatively small number of hours spent in training. Assuming that school requires 30 hours of study per week and takes place over 40 weeks, a year of schooling requires 1,200 hours of time investment. This is approximately 12 times longer than the 100 hours corresponding to a training episode within our model. Card (1999) surveys the vast literature on returns to education and finds estimates implying increases in wages of between 5% and 15% associated with an additional year of high school, or approximately 0.4%–1.3% per 100 hours invested. Blundell, Dearden, and Sianesi (2005) estimate a wage return of 24% for the 2 years of education differentiating high school graduates from those who leave school at 16 (with less than high school qualifications) in the UK context, or approximately 1% per 100 hours invested. Our estimates of the initial return from training at the start of working life fall on very similar values.

In figure 8 we document the extent by which training can offset the part-time penalty in wages. The diagram compares the loss in wages that results from a shift from full-time work to (*a*) part-time work (solid line) or (*b*) part-time work plus training (dashed line). It represents how the impact of training

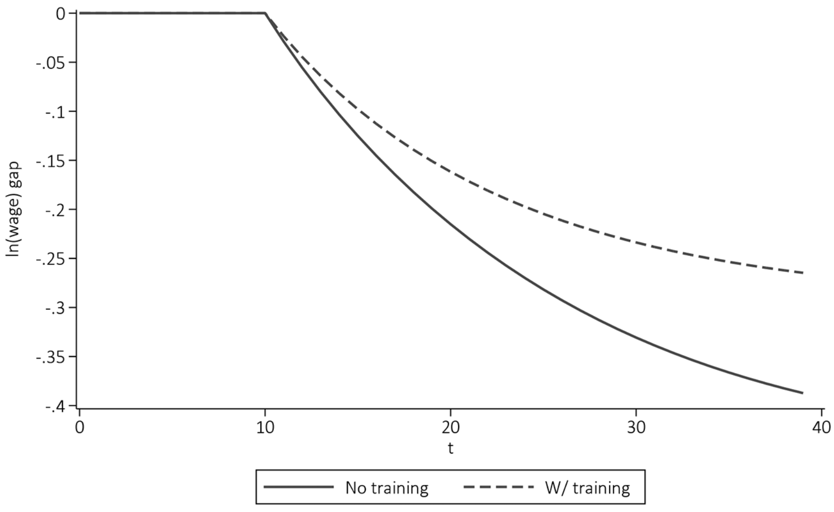

FIG. 8.—Training and the wage penalty from working part-time hours, by education. Solid lines represent the wage penalty, in log points, from moving to continuous part-time work after 10 years of continuous full-time work. Dotted lines factor in continuous training starting in year 10, together with part-time working hours.

compares with that of part-time hours. The solid lines in the figure show that part-time work is associated with a large wage penalty. The dashed lines show that taking training together with part-time hours offsets almost one-third of the part-time penalty.

B. Utility Parameters and the Cost of Training

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

To reproduce the observed employment rates at the given monetary incentives to work, the model requires working to carry a utility cost for all groups (see the estimates in cols. 1 and 2 of table 6). The costs are lower for married women than for single women, partly offsetting differences in incentives to work between the two groups due to a 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, Pistaferri, and Saporta-Eksten 2016). Mothers of young children, particularly of preschool 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 hours are less onerous in utility terms than full-time hours.

Estimates for the parameters governing the utility cost of training are shown in table 7. We have fixed the monetary cost of training to equal the forgone wage for 2 hours of training per week, or 104 hours per calendar year, which corresponds in the data to the median level of training among

Table 6
Parameters Determining Utility Cost of Working

	Utility Parameters in U_h			
	Full-Time Employment (α_{38})		Part-Time Employment (Increment: $\alpha_{18} - \alpha_{38}$)	
	Parameter Value (1)	SE (2)	Parameter Value (3)	SE (4)
Single, no children	.56	.006	−.37	.004
Single mothers	.47	.011	−.22	.009
Married, no children	.33	.014	−.23	.015
Married mothers	.34	.013	−.24	.012
Child aged 0–2	.16	.009	−.07	.008
Child aged 3–5	.11	.010	−.05	.009
Child aged 6–10	.06	.010	−.04	.006
Spouse working	−.07	.013	.08	.012
High background factor 1	.02	.008	.00	.005
High background factor 2	.03	.008	−.02	.005
$l_h(\theta)$ type I	−.38	.178	.00	.005

Table 7
Parameters Determining Utility Cost and Benefits of Training and the Terminal Value of Human Capital

	Parameter Value	SE
	Utility Parameters in U_T, (α_T, α_{Tb})	
1. Single, no children	.002	.010
2. Single mothers	.007	.008
3. Married, no children	−.002	.015
4. Married mothers	−.002	.016
5. Child aged 0–2	.010	.025
6. Child aged 3–5	.004	.014
7. Child aged 6–10	.003	.008
8. Spouse working	.009	.014
9. High background factor 1	.004	.007
10. High background factor 2	.002	.005
11. Part-time interaction	.016	.006
12. $l_T(\theta)$ type I	−.028	.003
	Terminal Value of Human Capital	
13. Scale parameter, ϕ_1	.05	.009
14. Curvature parameter, ϕ_2	.21	.137

trainees undergoing more than 1 week of training over the 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.

Most parameters in the utility of training are small and mostly not statistically significant at conventional levels: the utility cost of training does not seem to depend on the demographic structure of the household or even on the family background factors. Perhaps this is not surprising, since most of the cost associated with the household structure relates to the decision to work or not, and once that has been paid it is no longer relevant for the training decision itself. However, the interaction with part-time hours (row 11) shows that training is more costly when women are doing short working hours, and the unobserved heterogeneity term (row 12) shows that the group with higher preferences for work also has a positive preferences for training (which mirrors a higher training cost for those with lower preferences for work). Our model captures well the training profiles of women around the birth of their first child and over the life cycle, as illustrated in figures 9 and 10. Given the estimated returns to training discussed before, the model rationalizes observed training levels with a preference for training among higher-ability women (who constitute 80% of the population according to estimates in panel B of table 4). A model that admits search frictions or other imperfections could provide a structural interpretation of this, since in that case the firm and the worker share the costs of training.

Our model implicitly points to two additional mechanisms explaining the life-cycle patterns of training. First, families with children have higher needs and may be more likely to face liquidity constraints. In those circumstances,

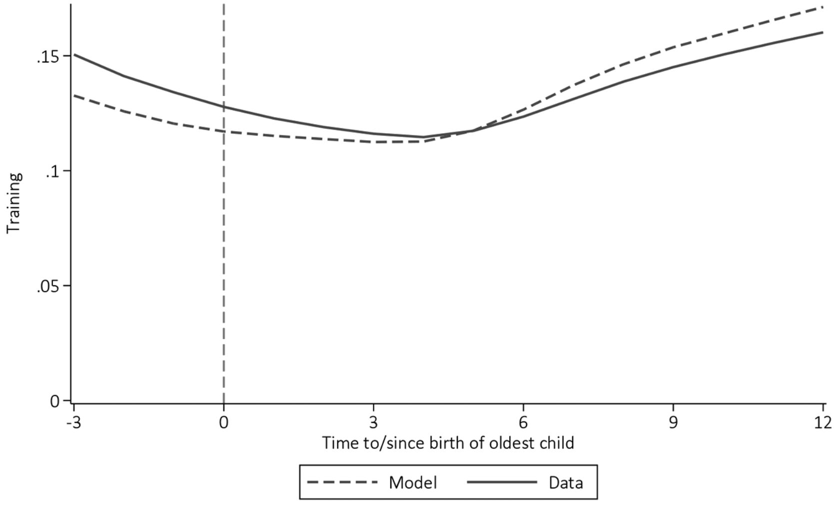

FIG. 9.—Model versus data: training incidence among working mothers, by time since/to birth of oldest child and maternal education. A color version of this figure is available online.

the forgone earnings associated with training may be an especially high cost to pay that could drive training rates down during that period of life. 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

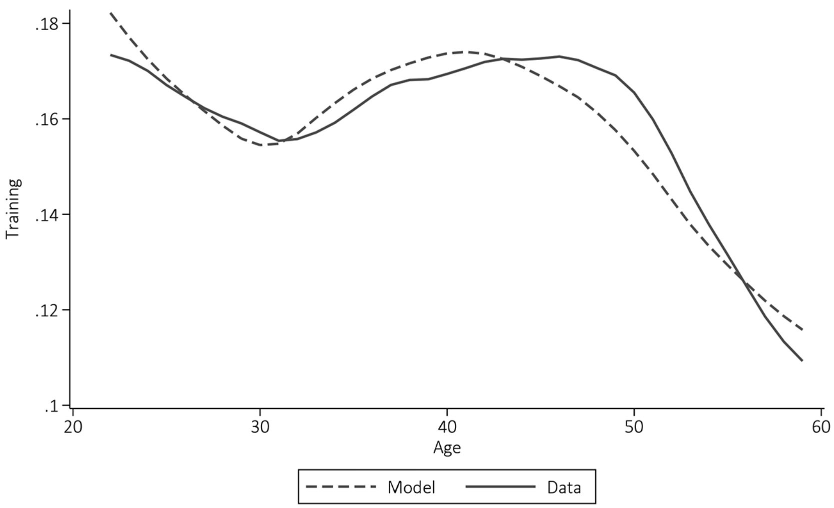

FIG. 10.—Model versus data: training incidence over the life cycle among working women, by maternal age and education.

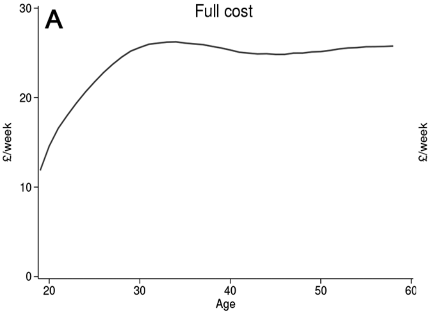

FIG. 11.—Monetized total cost of and experience return to training across the whole population, by age and education. *A*, Monetized cost of training (full cost). *B*, Consumption value of extra human capital (experience benefit).

before depreciation eventually washes out the human capital gains from training.

Figure 11*A* plots age profiles for the average total cost of training, including both the monetary cost associated with lost labor 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 in figure 11*B*. In line with the observed training rates, the average cost exceeds the average return by a factor of two for most age groups. Figure 12 plots similar graphs but by time to/from the birth of the first child. The life-cycle variation is strongly associated with the dynamics of family demographics

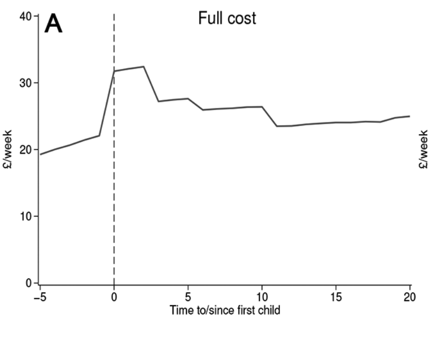

FIG. 12.—Monetized cost of and experience return to training across the whole population, by time to/since first birth and education. *A*, Monetized cost of training (full cost). *B*, Consumption value of extra human capital (experience benefit). *A* shows the average monetary compensation required to equalize period utility between (1) working full time and not training and (2) working full time and training. *B* shows the average monetary deduction required for an individual to be indifferent to receiving additional human capital equivalent to 1 unit of training. A color version of this figure is available online.

through employment behavior rather than through the utility cost of training. The returns to training also change around childbirth but by a much more modest amount, and they then slowly recover as the child grows up.

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

C. Responses of Employment and Training to Changes in Prices

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

Column 2 reports the average immediate response to an unanticipated and permanent 5% decline in the posttax wage rate starting at each age in the 23–50 interval. Overall, this change leads to a 2% decline in employment on a base of 85.7%, displaying the dominance of the wealth effect. The response

Table 8
Model Simulations: Employment and Training Responses to Changes in Wages and the Monetary Cost of Training

	Level (%) (1)	5% Permanent Decrease in Net Earnings (2)	5% Permanent Decrease in Training Cost (3)
	A. Employment		
All women	85.7	−2.0	.0
By family demographics:			
Single, no kids	93.3	−1.4	.0
Single mothers	69.7	−4.2	.0
Couples, no kids	94.3	−.8	−.1
Mothers in couples	79.0	−2.6	.0
	B. Training Conditional on Employment		
All women	16.7	.3	2.1
By family demographics:			
Single, no kids	16.3	.0	3.0
Single mothers	10.3	.7	2.9
Couples, no kids	18.4	.3	1.6
Mothers in couples	16.6	.2	1.9

NOTE.—Calculations are based on model simulations. Column 2 shows the effects of an unanticipated and uncompensated 5% permanent decrease in net earnings on the employment and training rates in the period the change in earnings is first realized. Column 3 shows the effects of an unanticipated and uncompensated 5% permanent decrease in the forgone earnings cost of training on the employment and training rates in the period the change in costs is first realized. In all cases, responses are averages of effects for women aged 23–50.

is larger for mothers, particularly single mothers, than it is for women without children, reflecting their larger labor supply elasticities. While training responses to changes in the wage rates are smaller than those of employment, they are nevertheless important given current training rates. A permanent drop in wages reduces future returns to training, but that is offset by the negative impact it has on the current cost of training. We find that the latter dominates, leading to a small overall increase in training rates particularly for single mothers. Finally, column 3 in the table shows that the training responses to a drop in the cost of training are large, particularly for single women. In turn, employment does not respond to changes in training incentives.

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

IX. Counterfactual Simulations and Discussion: Subsidized Training for Mothers

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

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

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

Table 9 shows model predictions of the effects of subsidized training on training rates, employment, hours, wages, savings, income, and welfare. The first column displays the effects of a £500 lump-sum subsidy for mothers of children aged 0–7 in training. The second column increases this to £1,500, and in the final column the subsidy provides full compensation for the monetary cost of training, which includes forgone earnings. The subsidy policy is made revenue neutral with a change in the basic tax rate.

Table 9
Impact of Training Subsidies

	Annual Training Subsidy		
	£500 (1)	£1,500 (2)	Full Compensation (3)
	A. Child Aged 0–7		
Training	9.33	40.95	19.84
Employment	.02	3.14	.01
Full time	.64	1.37	1.35
Part time	−.62	1.77	−1.34
	B. Child Aged 8		
Assets (%)	−.12	.51	.69
Wages (%)	.31	2.60	1.36
	C. Child Aged 8–18		
Training	−.19	−.38	−.31
Employment	.10	.52	.15
Full time	.41	−.30	.20
Part time	−.31	.82	−.05
	D. Child Aged 19		
Assets (%)	−.05	.24	.29
Wages (%)	.14	.72	.44
	E. Child Aged ≥19		
Training	.00	−.03	.03
Employment	.21	.26	.23
Full time	.72	.26	.62
Part time	−.51	.00	−.39
	F. Lifetime Outcomes		
Disposable income (%)	.24	.35	.23
Consumption equivalent (%)	.83	.74	.77
	G. Revenue Neutrality Adjustment		
Basic income tax change	−.02	.5	.15

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

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

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

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

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

Panel B shows that the cumulative effect of the additional training and full-time work on the wage rates of women at the end of the eligibility period is positive, with mothers benefiting from a 0.31% increase in wages. This demonstrates that the policy has a small but not negligible impact on the human capital of mothers at the end of the eligibility period. The subsidy also reduces savings by a modest 0.12% when the child reaches 8 years

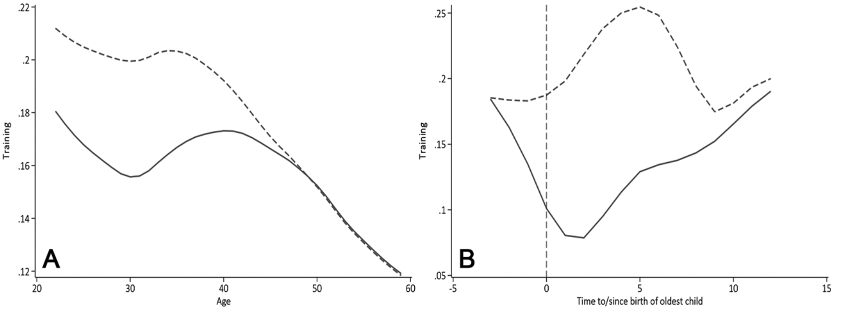

FIG. 13.—Training over the life cycle for high school educated under £500 subsidy. *A*, Training rate by age. *B*, Training rate by time since birth of oldest child. A color version of this figure is available online.

of age. This suggests that the additional human capital the woman accumulated over this period will make future work more likely and she will need to rely less on savings.

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

Column 2 of the table shows similar results for a more generous lump-sum subsidy of £1,500 per year, or about 120% of the pay of eligible mothers during training episodes. The additional generosity comes with a high price, requiring an increase of 0.5 percentage points in the basic tax rate to balance the public budget. For comparison, the calculations of Blundell et al. (2016) suggest that funding for the 2002 tax credit scheme in the United Kingdom adds 0.9 percentage points to the basic tax rate. Despite its cost, which is fully borne by this population of women and their partners, our simulations show that this policy is welfare increasing and drives up disposable income by more than the less generous policy. These effects result from the strong impact that the policy has on the training rates of eligible mothers, which increase by 41 percentage points, and their employment rates, which also increase by 3 percentage points. The combined effects of these responses result in higher wages at the end of the eligibility period (2.6%; panel B), which drives later employment gains and persistent increases in wages (by 0.72% when the child reaches 19 years of age).

The lump-sum subsidies provide a stronger incentive for those in low pay, who may also benefit less from training if they are from the low-ability group or on a flatter wage trajectory induced by low (persistent) productivity shocks. We therefore redesigned the subsidy to exactly cover the forgone earnings of trainee mothers of children aged 0–7. Results for this policy are displayed in column 3 of table 9. Because this type of design incentivizes

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

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

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

training among higher-paid mothers, it also ends up being more expensive for each trainee than the generous £1,500 lump-sum subsidy, costing £1,600 per trainee. However, it draws fewer women into training than the lump-sum benefit because it is less generous for lower-paid women. So in the end the cost of such a policy is smaller than that of the more generous lump-sum transfer, requiring an increase of 0.15 percentage points in the basic tax rate to balance the public accounts. Its effects lie between the figures in the first two columns of the table for the two lump-sum subsidies.

X. Conclusions

We have estimated a life-cycle 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 play 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 nonemployment following a return to the labor market after having children.

Finally, we evaluate a policy of subsidizing training for mothers with children younger than 8. All policies are revenue neutral and are funded by increasing taxes. A fixed modest subsidy of £500 increases the take-up of training substantially and leads to small but persistent gains in wages, lifetime disposable income, and welfare. It also pays for itself. We also consider other less effective and more expensive approaches.

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

References

Acemoglu, Daron, and Jorn-Steffen Pischke. 1999. Beyond Becker: Training in imperfect labour markets. *Economic Journal* 109, no. 453:112–42.

Adam, Stuart, James Browne, and Christopher Heady. 2010. Taxation in the UK. In *Dimensions of tax design: The Mirrlees review*, ed. James

Mirrlees, Stuart Adam, Timothy Besley, Richard Blundell, Stephen Bond, Robert Chote, Malcolm Gammie, Paul Johnson, Gareth Myles, and James Poterba. Oxford: Oxford University Press.

Altonji, Joseph G., and Charles R. Pierret. 2001. Employer learning and statistical discrimination. *Quarterly Journal of Economics* 116:313–50.

Altonji, Joseph G., and Lewis M. Segal. 1996. Small-sample bias in GMM estimation of covariance structures. *Journal of Business and Economic Statistics* 14, no. 3:353–66.

Altonji, Joseph G., and James R. Spletzer. 1991. Worker characteristics, job characteristics, and the receipt of on-the-job training. *Industrial and Labor Relations Review* 45:58–79.

Andrews, Isaiah, Matthew Gentzkow, and Jesse M. Shapiro. 2017. Measuring the sensitivity of parameter estimates to estimation moments. *Quarterly Journal of Economics* 132, no. 4:1553–92.

Attanasio, Orazio, Hamish Low, and Virginia Sánchez-Marcos. 2008. Explaining changes in female labor supply in a life-cycle model. *American Economic Review* 98, no. 4:1517–52.

Attanasio, Orazio, and Guglielmo Weber. 1995. Is consumption growth consistent with intertemporal optimization? Evidence from the Consumer Expenditure Survey. *Journal of Political Economy* 103, no. 6:1121–57.

Becker, Gary S. 1964. *Human capital: A theoretical and empirical analysis, with special reference to education*. Chicago: University of Chicago Press.

Blundell, Richard, Martin Browning, and Costas Meghir. 1994. Consumer demand and the life-cycle allocation of household expenditures. *Review of Economic Studies* 61, no. 1:57–80.

Blundell, Richard, Lorraine Dearden, and Barbara Sianesi. 2005. 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* 168, no. 3:473–512.

Blundell, Richard, Monica Costa Dias, Costas Meghir, and Jonathan Shaw. 2016. Female labor supply, human capital, and welfare reform. *Econometrica* 84, no. 5:1705–53.

Blundell, Richard, Luigi Pistaferri, and Itay Saporta-Eksten. 2016. Consumption inequality and family labor supply. *American Economic Review* 106, no. 2:387–435.

Brewer, Mike, Alan Duncan, Andrew Shephard, and Maria Jose Suarez. 2006. Did working families tax credit work? The impact of in-work support on labour supply in Great Britain. *Labour Economics* 13, no. 6:699–720.

Card, David. 1999. The causal effect of education on earnings. In *Handbook of labor economics*, 1801–63. Amsterdam: Elsevier.

Chiappori, Pierre-André, Monica Costa Dias, and Costas Meghir. 2018. The marriage market, labor supply, and education choice. *Journal of Political Economy* 126, no. S1:S26–S72.

Chiappori, Pierre-André, Murat Iyigun, and Yoram Weiss. 2009. Investment in schooling and the marriage market. *American Economic Review* 99, no. 5:1689–713.

Fan, Xiaodong, Ananth Seshadri, and Christopher Taber. 2017. Understanding earnings, labor supply, and retirement decisions. Working Paper no. wp367, University of Michigan, Michigan Retirement Research Center.

Flinn, Christopher, Ahu Gemici, and Steven Laufer. 2017. Search, matching, and training. *Review of Economic Dynamics* 25:260–97.

Gourieroux, Christian, Alain Monfort, and Eric Renault. 1993. Indirect inference. *Journal of Applied Econometrics* 8:S85–S118.

Heckman, James, Lance Lochner, and Ricardo Cossa. 2002. Learning-by-doing vs. on-the-job training: Using variation induced by the EITC to distinguish between models of skill formation. NBER Working Paper no. 9083, National Bureau of Economic Research, Cambridge, MA.

Lentz, Rasmus, and Nicolas Roys. 2015. Training and search on the job. NBER Working Paper no. 21702, National Bureau of Economic Research, Cambridge, MA.

Shaw, Jonathan. 2011. Fortax: UK tax and benefit system documentation. IFS Working Paper no. W11/08, Institute for Fiscal Studies, London.