IFS

Institute for Fiscal Studies

David Goll
Robert Joyce
Tom Waters

Working paper 23/03

Intensive margin labour supply and the dynamic effects of in-work transfers

UKRI Economic and Social Research Council

Intensive margin labour supply and the dynamic effects of in-work transfers*

David Goll† Robert Joyce‡ Tom Waters†

January 13, 2023

Abstract

Policy-makers have increasingly turned to 'in-work transfers' to boost incomes among poorer workers and strengthen work incentives. One attraction of these is that labour supply elasticities are typically greatest at the extensive margin. Because in-work transfers are normally subject to earnings-related phase-outs, they tend to most strongly incentivise part-time work, weakening intensive margin incentives for many. But part-time work may generate relatively little in the way of human capital and career progression. How should these dynamic considerations affect the design of in-work transfers? To assess this we use a dynamic model of female labour supply with endogenous human capital accumulation. Among reforms that would cost the same amount on a no-behavioural-response basis, those that incentivise full-time work can end up costing considerably less than those that incentivise part-time work, once the dynamic responses - including human capital accumulation - are accounted for. They also do more to increase incomes, including among poorer households, and to raise welfare. Our results suggest that in-work transfers could be refined by paying greater attention to the intensive margin effects through the design of their phase-outs.

Keywords: Life-cycle model, human capital, transfers, labour supply, intensive margin, extensive margin, welfare.
JEL Codes: H2, H31, J22, J24.

*This work was funded by the Economic and Social Research Council grant 'Welfare integration, labour supply and take-up' (ES/V003704/1). Co-funding from the Economic and Social Research Council through the Centre for the Microeconomic Analysis of Public Policy at IFS (ES/H021221/1) is also gratefully acknowledged. We thank Richard Blundell, Monica Costa Dias, and Hilary Hoynes for helpful comments and advice. This work was produced using the UK Household Longitudinal Study, an initiative funded by the Economic and Social Research Council and various government departments, with scientific leadership by the Institute for Social and Economic Research, University of Essex, and survey delivery by NatCen Social Research and Kantar Public (University of Essex, 2021). The research data are distributed by the UK Data Service.

†Institute for Fiscal Studies and University College London
‡Institute for Fiscal Studies

1. Introduction

The growth of "in-work" transfers, such as the US' Earned Income Tax Credit (EITC) and the UK's Working Families Tax Credit (WFTC), has been a major development in transfer policy in developed economies in recent decades. These have been viewed as a tool for raising living standards among low-earning households but also, crucially, for changing incentives. The focus has been on encouraging participation in paid work or, viewed more holistically, to help offset the *dis*incentives to participation that income-related transfer systems as a whole tend to create.

This approach to structuring incentives has fitted well with insights from public economics. Saez (2002) set out the key theoretical and empirical arguments: Mirrleesian logic suggests that, all else equal, one should focus the strongest work incentives where labour supply effects are largest (Mirrlees, 1971) which, according to most empirical estimates, is the extensive margin. These considerations point towards work-contingent support as an element of the transfer system, along the lines of the EITC and WFTC. But in-work transfers are typically targeted at low earners, and hence are subject to phase-outs, which creates a natural tendency for the boosts to participation incentives to go alongside labour supply *dis*incentives along the intensive margin for some workers. The apparent costs of those disincentives in the traditional single-period framework may understate the true costs, if working hours are linked to wage dynamics. Phasing out support at higher earnings levels could discourage precisely the kind of work that brings the most wage progression, undermining the central goal of raising living standards among low earners. A critical question, unaddressed thus far in the large literature on in-work transfers, is therefore whether they could be reformed so as to better support that goal over the long term once we properly account for the dynamic effects of policy.

Our contribution in this paper is to examine that question. We use a dynamic model of female labour supply and consumption with endogenous human capital accumulation, estimated on rich longitudinal UK data, to show how different policy reforms can have different long term effects on incomes, consumption and welfare even if their short term effects on employment levels are the same. The key mechanism generating these results is that different reforms create different work incentives along the intensive margin, and the wage returns to full-time and part-time experience are different. In fact, the difference between full-time work and part-time work in their impact on future wages is considerably larger than the difference between part-time work and non-employment. As a result, once we take a dynamic view then labour supply incentives at the intensive margin become a key variable in shaping the impacts of welfare reform.

Given that higher wages and incomes bring a positive fiscal externality, this insight may offer the potential for welfare improvements in transfer design. A welfare comparison of full-time and part-time incentivising policies also needs to factor in both the value of leisure, which will tend to be reduced most when intensive margin incentives are strongest, as well as the insurance properties of the respective approaches. Weak intensive margin incentives could have an upside via better insurance: with positive labour supply elasticities part-time work is more likely to be favoured after a negative wage shock, and so along this margin a higher transfer to part-time workers improves welfare for risk averse individuals. Our empirical framework captures all these effects and the resulting welfare calculations suggest that full-time incentivising policies are strongly favoured overall for a given expansion in the government budget.

We build on a large literature that examines the impacts on labour supply behaviour of tax and transfer systems in general, and in-work transfers in particular (see e.g. Blundell and Macurdy, 1999 and Keane, 2011; for a recent review focusing on transfers specifically, see Chan and Moffitt, 2018). Many important empirical regularities have emerged from this literature, including wide heterogeneity in labour supply elasticities across demographic groups and over the lifecycle, and generally higher responsiveness of labour supply at the extensive margin than the intensive margin (though the size of the extensive margin response to the EITC has recently been the subject of some debate; see Kleven, 2019). For the most part, and especially until recently, this literature has run in parallel with a large body of research on the determinants of wages and wage growth over the lifecycle, in which the role of human capital, including through on-the-job learning, is central (e.g. Huggett, Ventura and Yaron, 2011). Our paper fits most closely with work which attempts to tie these ideas together, examining the interaction between endogenous human capital accumulation and the incentives created by tax and/or welfare policy (Krueger and Ludwig, 2013; Keane and Wolpin, 2007; 2010; Blundell et al., 2016; Stantcheva, 2017).

The study that most directly motivates the questions addressed in this paper is that of Blundell et al. (2016). That paper exploited changes to the tax and transfer system to separately identify the impacts of part- and full-time work on labour market experience capital accumulation and the impact of that experience capital on wages. Interpreting the relationship between working hours and wage dynamics as a human capital effect on wages is natural given the empirical persistence of the effect, and we follow this interpretation for ease of exposition (though it is possible that other mechanisms are at work too). They used the identified model to analyse the impacts of the UK's system of in-work transfers on women - the group who, especially during motherhood, are most responsive to financial incentives and who are a key focus of transfer policies. The paper finds that the dynamic effects of expansions of in-work transfers in the late 1990s and early 2000s were of limited magnitude: though they changed labour supply, the impact on careers beyond the period of eligibility was modest. This is because the reforms tended to encourage part-time work among relatively low skilled workers, which generates little human capital.

A critical unanswered question is whether, and how, in-work transfers could be designed so as to have more profound long-term effects on careers. This is the focus of our contribution. To make headway we use an empirically-estimated model that shares the same basic structure as that of Blundell et al. (2016) to conduct a number of policy experiments. We use these to illuminate the key dependence between the long-term effects of welfare reform and their impacts on intensive margin labour supply incentives, and show the relevance of these empirical findings for live real-world policy choices. While our focus here is on the design of income-related in-work transfers, and in particular their phase-outs, we note that these insights also likely have relevance for a much wider set of tax or transfer policies. For example, recent evidence suggests that universal transfer payments may also result in shifts to part-time work (Jones and Marinescu, 2022).

The paper is structured as follows. Section 2 describes the data we use and the empirical lifecycle patterns that motivate our analysis. Section 3 outlines the core features of our model and our approach to estimation and calibration of its parameters. Section 4 uses the model parameters to illustrate how current labour supply responds to financial incentives and how it in turn affects human capital and future wages. Section 5 shows the results from simulations of a number of policy experiments, which have different effects on part-time and full-time work incentives. Section 6 concludes.

2. Data and descriptives

The data we use are drawn from the UK Household Longitudinal Study (UKHLS). This annually surveys a panel of individuals and contemporaneous members of their household, collecting a rich set of information including hours worked, earnings, educational qualifications, and demographics. We use data from 2012 to 2019 (UKHLS began in 2009, but we wish to avoid the contaminating effects of reduced labour demand during the Great Recession). Our sample for analysis is a panel of women aged between 19 and 60 who have entered the labour market, where labour market entry is defined as two consecutive periods of employment when younger than 25 or any employment when older than 25. We drop all observations covering years in which women are in full-time education. We are left with 16,941 women in our sample, whom we observe for 4.25 periods on average. We allow parameters of our model to vary by, and present results disaggregated by, the highest educational qualification that women have, which we categorise as those with GCSEs only (who typically left school at the compulsory schooling age of 16); A-Levels (age 18), and degrees (age 21 or 22). We refer to these groups as the low-, mid- and high-educated.

Some basic descriptive patterns from these data are worth outlining at the outset, based on figures included in the appendix. First, lower-educated women - often the recipients of means-tested transfers - are much less likely to be in full-time work than other women. Second, for women of all levels of education we see a drop in full-time work and a rise in part-time work around the typical age at which families are formed (see Appendix Figure B.1), and plotting employment profiles before and after the birth of the oldest child confirms that this switch to part-time work occurs immediately once that child is born (Appendix Figure B.2). The presence of dependent children is also associated with a higher rate of entitlement to means-tested transfers, meaning that this period - where transfer receipts are high and the prevalence of full-time work is relatively low - is one in which the potential for policy-related incentives to have effects on labour supply may be substantial.

Finally, and key to the rest of this paper, there are important relationships between education level, hours of work and wage progression. The more highly educated see a steeper age-profile in wages. As we will show with the benefit of our model in Section 4, this is driven both by more full-time work among those with higher education levels, and by larger complementarities between education and (full-time) working experience in the determination of wages.

3. Model and estimation

We apply a dynamic lifecycle model in which women choose labour supply and consumption each year. The structure of the model is based on Blundell et al. (2016). As described in more detail below, we re-estimate a subset of parameters and all exogenous processes to ensure that the model provides a good fit to survey data collected between 2012 and 2019; the model estimated in Blundell et al. (2016) uses data collected between 1991 and 2008 (from the British Household Panel Survey, to which UKHLS is a successor).

Agents enter the model at age 22 if they completed a degree or at age 19 otherwise. Each year, they can choose either to provide no market labour, work part-time (at 18 hours per week) or work full-time (at 38 hours per week). Women retire at the age of 60 and live for another 10 years from their accumulated savings. Labour supply generates income this period but also has long-term human capital benefits: women who have supplied more labour in the past have, on average, higher productivity and can therefore command a higher hourly wage. Human capital depreciates over time, so labour supply last year has a larger impact on wages this period than labour supply ten years ago. Agents account for human capital accumulation effects when making their labour supply decisions.

The wage, human capital and unobserved productivity processes are presented below (Equations 1 to 3). Human capital (e_t) this period is equal to human capital last period, minus depreciation (δ_s, where s indexes education level) and plus any additional human capital acquired through labour supply. Women gain a single unit of human capital if they work full-time and $\gamma_s < 1$ units of human capital if they work part-time. Log wages are a linear function of log human capital scaled by $\beta_{s,1} < 1$; as a result, the marginal effect of an additional unit of human capital on wage levels is declining in the human capital stock. Women's wages also depend on unobserved productivity (v_t). Productivity is idiosyncratic and persistent, following an AR(1) process.

$$\ln w_t = \beta_{s,0} + \beta_{s,1} \ln(e_t + 1) + v_t \quad (1)$$

$$e_t = (1 - \delta_s)e_{t-1} + 1(l_{t-1} = 38) + \gamma_s 1(l_{t-1} = 18) \quad (2)$$

$$v_t = \rho_s v_{t-1} + \xi_t \quad (3)$$

The parameters of the wage, human capital and unobserved productivity processes depend on the level of formal education acquired prior to entering the labour market (s) - allowing for the possibility that there are complementarities in education and the human capital accumulation effects of work. We also allow the disutility of labour supply to vary based on education level, family demographics and their interactions. There are stochastic, exogenous changes in family demographics: a woman may have a child in each period (up to the age of 43), and form or dissolve a partnership. When in a couple, total household income incorporates the income of their partner, which is determined by an exogenous stochastic process. All exogenous processes, including fertility, the formation of partnerships and partner earnings, are estimated outside the model using UKHLS data.

Both the persistent productivity process shown in Equation 3 and the stochastic changes in family demographics and partner earnings expose women to idiosyncratic, uninsurable risk. In any period they choose how much of their accumulated assets and current income to spend and how much to set aside for the future. They are unable to borrow against future income.

We use FORTAX, a detailed tax and benefit microsimulation model, to generate accurate budget constraints reflecting contemporaneous UK welfare policy (a description of FORTAX is available in Shaw, 2011).

Most of the parameters of the model are estimated using method of simulated moments, with the moments estimated using the UKHLS data described in Section 2. The parameters we estimate, all of which can vary by education, determine:

- The impact of human capital on wages
- Idiosyncratic wage risk
- Average disutility of full-time and part-time labour supply
- Probability distributions for unobserved heterogeneity "type"

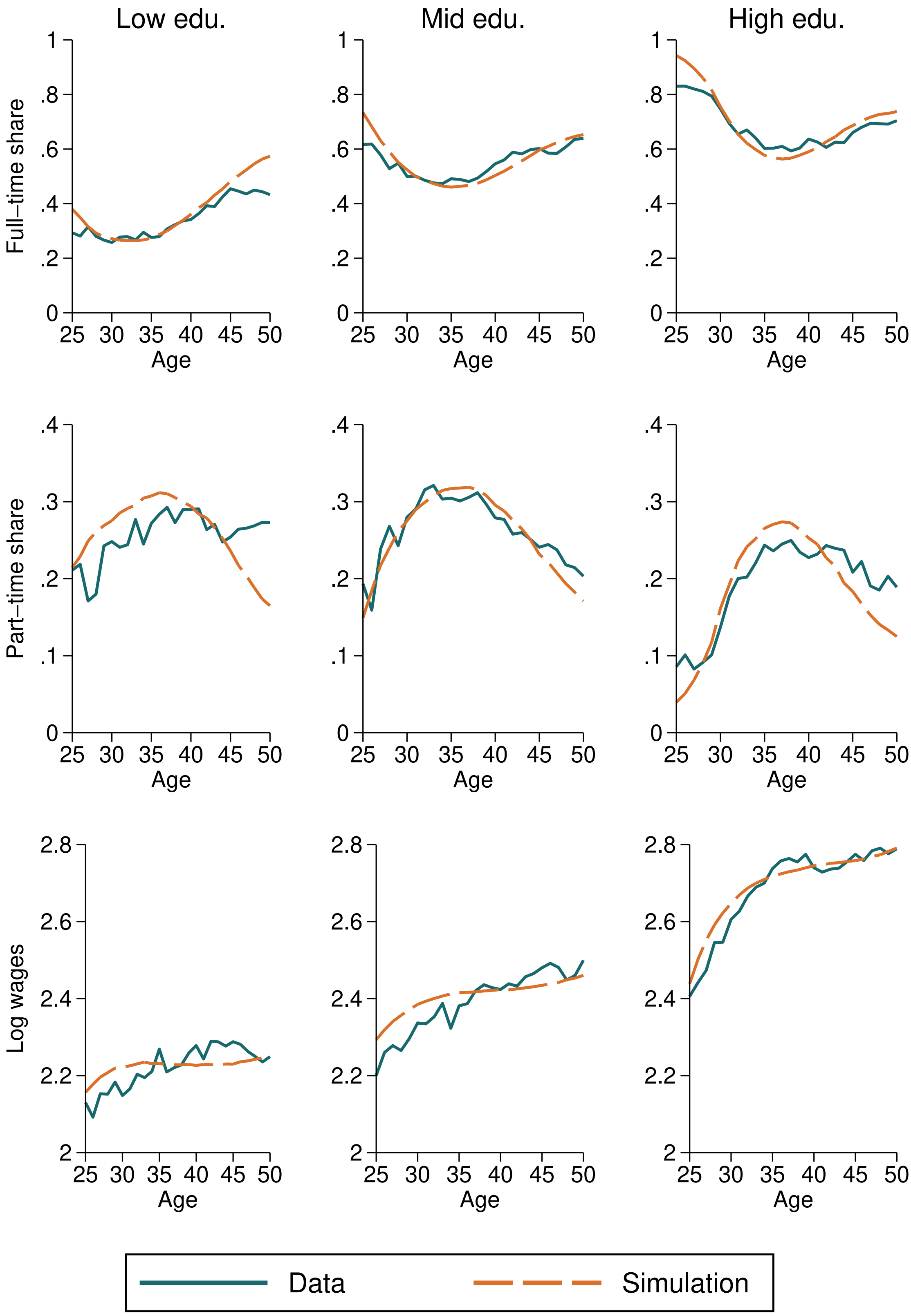
Details of the data moments used in estimation, and their simulated counterparts, are included in the appendix. Finally, some parameters are calibrated, drawn from the results in Blundell et al. (2016). These include parameters determining:

- Human capital accumulation and depreciation
- Impact of demographics on the disutility of full-time and part-time labour supply
- Impact of unobserved heterogeneity on the disutility of full-time and part-time labour supply

These parameters are identified in Blundell et al. (2016) by exploiting a series of large reforms to the welfare system. Specifically, they construct moments which vary by an individual's family background and the year in which they turned 16, which determines the tax-transfer systems they faced at different stages of their life. The impact of policy reforms on the incentives that people with different backgrounds face therefore help identify the parameters which determine how part- and full-time work contribute to human capital accumulation and how that human capital translates to wages. The variation in benefits used to identify the key human capital parameters closely mirrors the type of reforms that we consider in this paper; in particular, the Working Family Tax Credit introduced over their sample period included a minimum working hours eligibility rule that assists in the separate identification of the human capital returns to part-time and full-time work.

Figure 1 shows the fit of the model in terms of part and full-time employment, and wages. The model is able to reproduce key features of the data which motivate this paper and will be central to its results. In particular, lower-educated women - disproportionately targeted by means-tested transfers - have significantly lower full-time employment rates over the lifecycle, but particularly during the years that approximately correspond to the period of child-rearing. For those in paid work, the age profile of wages is much flatter for the low-educated, and this lack of progression results in an education gradient in wages which increases steeply with age. Appendix A shows that the model is also successful in fitting wage distributions and the relationship between wages, work experience, and age.

Figure 1. DATA AND MODEL FIT FOR KEY LABOUR MARKET OUTCOMES



Notes: Log wage figures show the average simulated log wage for all women, regardless of whether or not they are simulated as working.

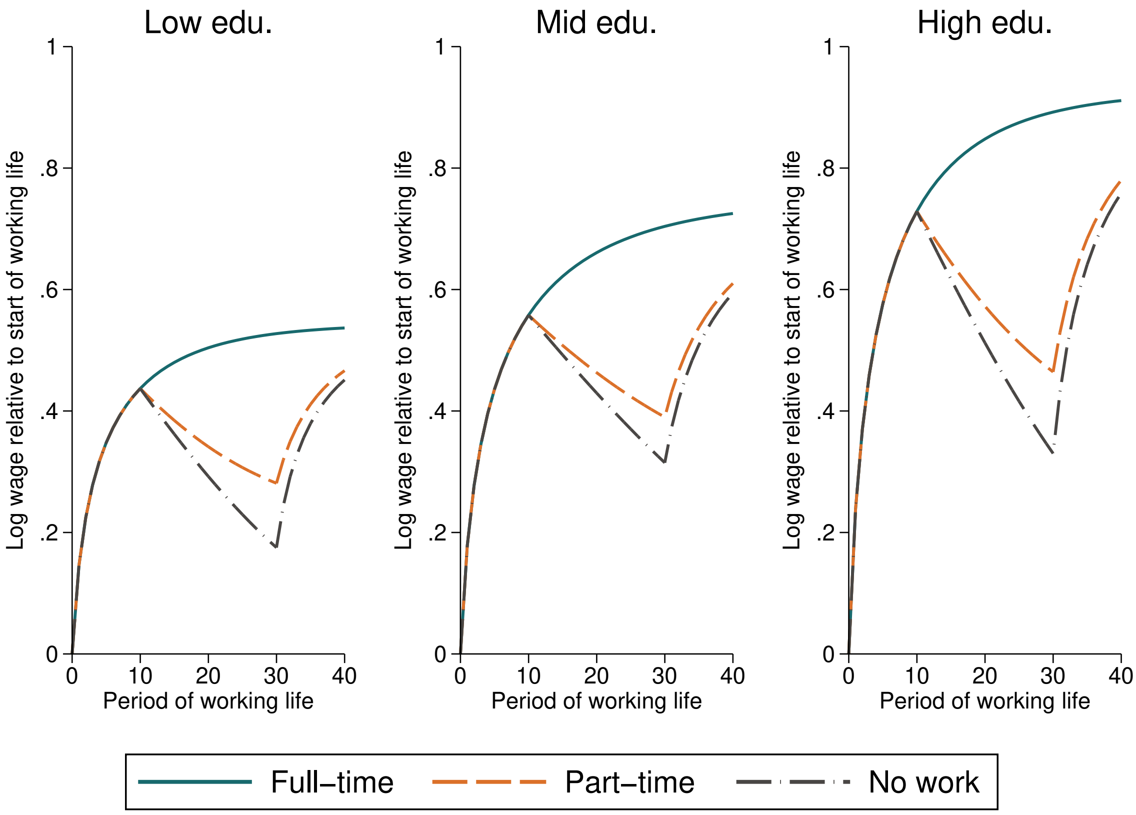
4. Hours of work, human capital accumulation, and labour supply elasticities

As background to the policy simulations that we present in the next section, we summarise the model's predictions of the human capital returns from work, and the responsiveness of employment to financial incentives.

We begin by illustrating how full-time and part-time work choices affect future wages, using the estimates of the parameters governing human capital accumulation and the wage function. We perform a simple simulation, where women work full-time for the first 10 years of adulthood, and from year 30 onwards. For the intervening 20 years we simulate three scenarios, where women work full-time, part-time, or not at all, every year. This is intended to crudely mimic labour supply options over the child-rearing period.

Figure 2 plots the impact on log wages. This shows four things. First, part-time work during these 20 years does relatively little to build up human capital, and therefore wages. In fact its impacts on future wages are more similar to not working at all than to working full-time. Second, the returns to full-time work experience (and hence the wages foregone by not working full-time) are larger for higher education groups (observe the gaps between full-time and the other lines in the 30th period of working life), reflecting a complementarity between education and experience. Third, the effects of reductions in working hours on wages are persistent - even after 10 years back in full time work the legacy of work choices during the childrearing years is substantial. Fourth, there are decreasing wage returns to human capital, resulting in the concavity of the wage profiles for those that always work full-time.

Figure 2. WAGE RETURNS TO FULL- AND PART-TIME WORK BY EDUCATION



Notes: Figure shows results of a a women work full-time for the first ten years of working life, and from year 30 on. In the intervening 20 years they always supply a given amount of labour as indicated by the legend. Graph shows simulated average log hourly wage that women could command if they were in work.

The impact that welfare policy can have on these dynamics is constrained by the responsiveness of women to financial incentives. We now briefly summarise the model's estimates of labour supply elasticities. A full description of these is in Appendix Figure B.3 and Table B.1.

Looking across the lifecycle, elasticities tend to show an inverse U shape - low early in life, then higher around the time when childrearing begins, then lower as children grow older and, eventually, leave home. Because lower educated groups tend to have children earlier, their peak in elasticities is earlier too. Looking between groups, we see that those with lower levels of education are much more responsive to financial incentives. For example, both the Frisch and Marshall elasticities of those with low levels of education are approximately double the elasticities of those with high levels. Mothers, and particularly single mothers, are also considerably more responsive than women without children. Single mothers have Marshall elasticities seven times as large as single women without children, and over 15% larger than mothers in couples.

While previous work taking a dynamic approach has suggested that actual transfer programmes have achieved relatively little in the way of human capital, these estimates highlight that the target populations for a typical transfer program are relatively sensitive to financial incentives when it comes to their contemporaneous choices. This raises the question of whether that sensitivity can be leveraged in a way that is better suited to encouraging the kinds of work that would have longer-term payoffs.

5. Policy experiments

We now ask whether in-work transfer policy can be designed differently so that it has larger (positive) longer-term effects on career dynamics, and consider the wider welfare consequences of such a change. We focus on the insight, demonstrated above, that part-time work tends to have negligible effects on human capital accumulation, while income-related in-work transfers or tax credits have a tendency to discourage progression to full-time work for many recipients due to the high marginal tax creates that they create when phased out.

When setting out our results we first focus on 'positive' analysis, describing effects on incomes and the government budget. Arguably it is these which largely shape policy-making in practice, as memorably argued by Besley and Coate (1992; 1995) who noted "There is little evidence that the poor's leisure is valued by policy makers". We also provide welfare calculations. These incorporate the impacts on welfare of changes in leisure as well as income, and they factor in the insurance properties of different systems which allow individuals to smooth consumption to different degrees. A particular relevant consideration in this context is that, with positive labour supply elasticities, part-time work is more likely to be chosen in 'bad' states of the world (after a negative wage shock), and so more relative generosity in that state of the world can have insurance benefits. Overall though the welfare effects across reforms turn out to show similar patterns to the changes in income, with stronger intensive margin incentives clearly favoured at a given cost.

5.1 Incentivising part- and full-time work

We compare reforms which would have different impacts on intensive margin labour supply choices, and then trace out their long run effects. We begin by examining in detail two policies which incentivise part- or full-time work particularly sharply and relatively directly. This illustrates starkly the difference in effects of policies which create different intensive margin incentives. We then go on to show that even when comparing commonplace changes to the basic parameters of transfer programmes, the relative intensive margin responses are key to a proper comparison of the long-term effects of the policies.

The baseline tax and transfer system for our policy simulations is the UK's 2022-23 system. This is primarily comprised of an income tax (levied at the individual level), payroll tax, and a means-tested household benefit called 'universal credit' (UC) - the result of a major recent reform which has integrated most of the UK in-work and out-of-work means-tested transfers for those of working age into one benefit. UC entitlement for a household with no other source of income is determined by whether they are a single or a couple, the number of children they have, and the level of rent they pay (if any). This is sometimes known as the 'basic entitlement'. As the household increases its earnings, that basic entitlement is steadily withdrawn at a 55% taper rate with respect to net-of-tax earnings above a disregard (though the disregard is zero for those without children). We simulate two reforms which are calibrated such that, in the absence of behavioural responses, the two reforms would cost the government the same amount (£4 billion per year).

The first reform we consider is the addition of a separate disregard for the earnings of the lower earner in a household. Under current policy, no distinction is made between the earnings of different household members when applying the means-test – all earnings are simply added together and compared to the household-level disregard. Hence, second earners in a household receiving UC will typically find themselves subject to the taper from the first pound that they earn. A "lower earner disregard" is often proposed in policy debate as a way of strengthening the work incentives of second earners, and women in particular (e.g. Brewer, Finch and Tomlinson, 2017; Butler and Rutter, 2016). However, as we will show, an important detail is that it is *part-time* work incentives that would typically be strengthened.[1]

The second reform we examine is the introduction of a 'lowest earner tax credit' in the UC system. Under this reform, a family's UC entitlement is increased if the lowest earner in the household earns at least a set amount. This tax credit is then tapered away with the rest of their UC as earnings rise. This has some commonalities with the US's Earned Income Tax Credit (EITC), though it is based on the earnings of the lowest earner, rather than total household earnings, and it does not slowly 'phase in' as EITC does. The earnings minimum is set such that the lower earner would need to work at least 30 hours per week at the national minimum wage to qualify for the tax credit (around £13,700 per year post-tax), and the credit is worth £6,800 per year. For a minimum wage worker it functions like a full-time 'bonus' equal to that amount. [2]

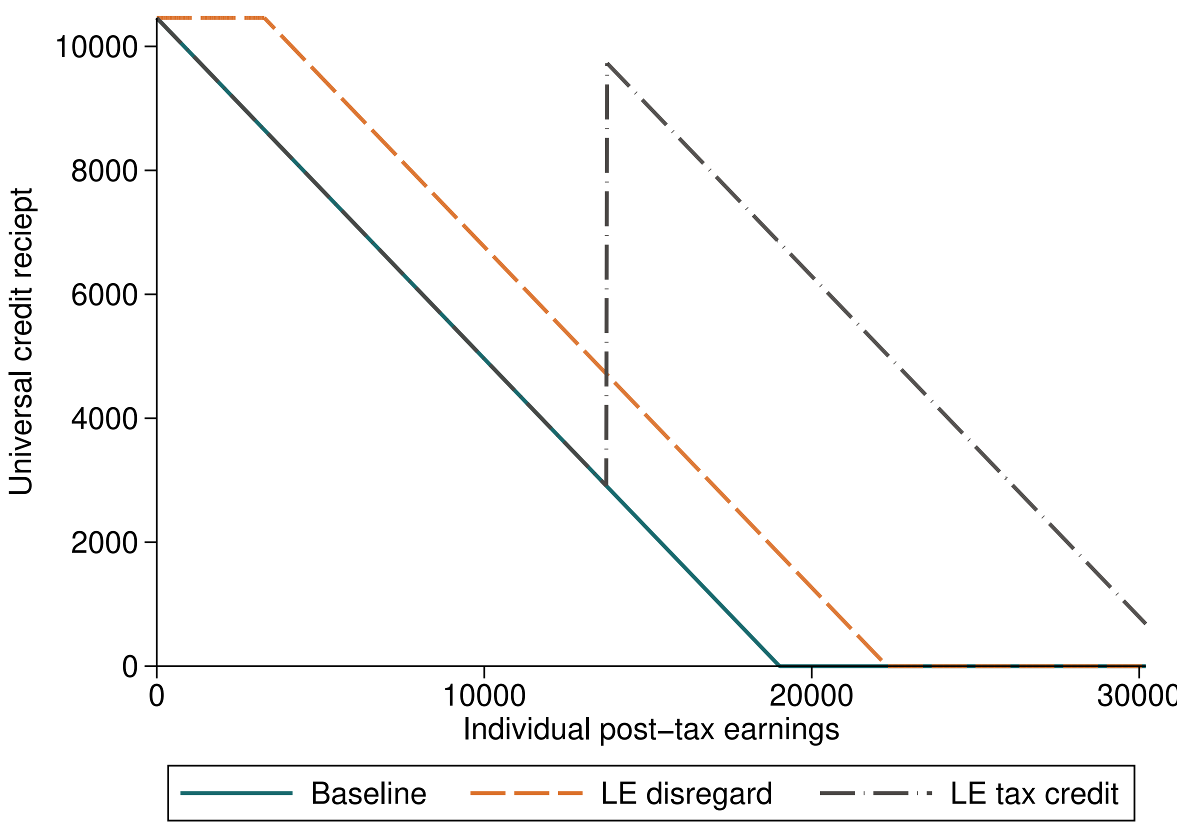
To illustrate how these reforms affect work incentives, Figure 3 shows UC entitlements for an example second earner under the baseline system and these two reforms. The first earner in the household is assumed to be earning above the household's disregard, meaning that under the baseline system UC begins to be withdrawn from the first pound earned by the second earner.

[1] We simulate the impact of an additional lower earner disregard with a value of around 80% of the current disregard. For example, the total annual disregard for a renting couple with a child is currently £4,128. Under our simulated reform, the lower earning member of the household can earn up to £3,509 before their earnings start to reduce the household's universal credit receipt, meaning a maximum disregard for the household as a whole of £7,637. For single mothers we simply increase their disregard by 80%; in other words we treat them as the 'lowest earner' in the household.

[2] An alternative reform would be to explicitly tie receipt of the EITC element to working a set number of hours per week. This would have precedent in the UK, which has had an hours-based full-time 'premium' in its tax credit system since 1995, and it is only the replacement of tax credits and out-of-work benefits by the integrated UC system which is removing it. But it would involve the government collecting information on hours worked - something which they currently do not do in the context of UC and is potentially open to fraud.

But with a lower earner disregard, she can earn up to £3,509 without affecting the household's UC entitlement. As the figure shows, the effect of this is to incentivise low earning work - which, for many, could be achieved by working part-time. Conversely, because the household is only eligible for the lower earner tax credit if the second earner earns at least £13,700 per year, the effect of its introduction is to incentivise earning at least that much. Working at the minimum wage in the UK this would require 30 hours per week or more as an employee. In Appendix Figure B.5 we confirm using micro-simulation that across a representative sample of women the effect of the lower earner disregard is to incentivise part-time work relative to full-time, on average, while the opposite is true of the lowest earner tax credit.

Figure 3. UC ENTITLEMENT FOR AN EXAMPLE SECOND EARNER UNDER DIFFERENT REFORMS



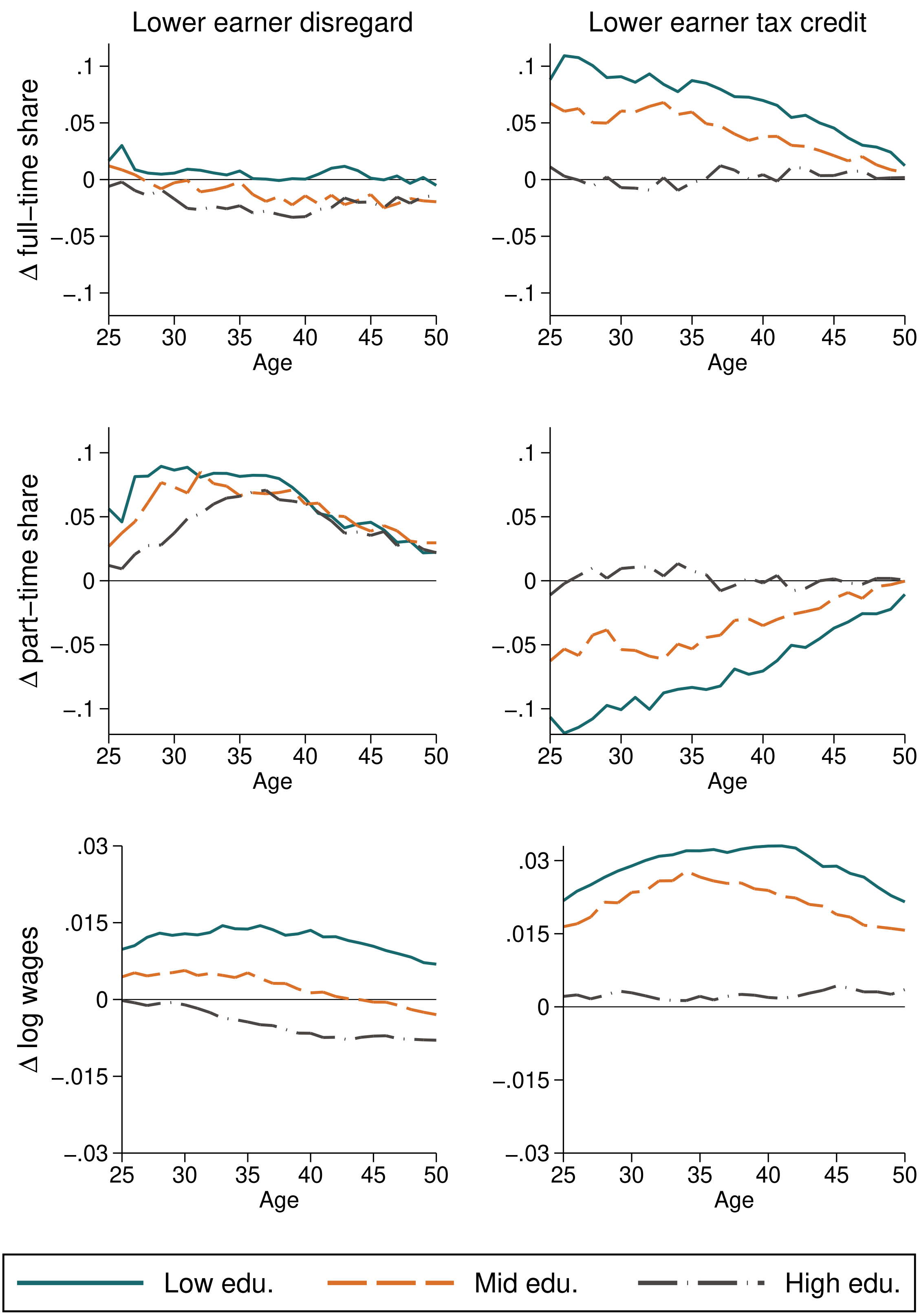
Notes: All values shown are annualised (UC is in fact assessed at a monthly frequency). Example household is a couple with weekly rent of £150 and one child. The horizontal axis shows the after tax (but pre-transfer) earnings of the woman. The woman's partner has annual earnings of £19,760 (pre-tax) - equivalent to 40 hours per week at the minimum wage. 'Baseline' is the actual 2022-23 tax and benefit system. 'LE disregard' and 'LE tax credit' are the lower earner disregard and lower earner tax credit discussed in the text.

Figure 4 shows the simulated impact of these policies (relative to the baseline system). The left hand subfigures show the impact from the lower earner disregard, and the right hand the lowest earner tax credit. The rows of subfigures show, in order, the impact of the policy on full-time work, part-time work, and the resulting effect on (log) hourly wages, in each case across the lifecycle and split by education. Consistent with the financial incentives, the introduction of the lower earner disregard results in a shift towards part-time work, to a small extent from full-time work but mainly from non-employment, meaning that the policy raises overall employment. Conversely, the introduction of the lowest earner tax credit has almost zero impact on overall employment, but causes workers to switch from part- to full-time work, with effects larger for lower-educated groups and earlier on in life.

The bottom panel of the figure traces out how these labour supply responses ultimately feed through to hourly wages via accumulated working experience.[3] Strikingly, the lower earner disregard results in higher female employment and yet lower wages than the lower earner tax credit. The key mediating mechanism is the intensive margin labour supply effect. The lower earner disregard strongly incentivises part-time work which generates little human capital, so its impacts on hourly wages are small and in fact slightly negative for the higher educated group. Conversely, the lower earner tax credit has the larger impacts on the level of full-time employment, which ultimately increases wages. These impacts are bigger for those with lower educational qualifications, consistent with the magnitude of the full-time labour supply response.

[3]Here we report the average simulated wage across all workers, regardless of whether or not they are simulated as actually being in work

Figure 4. Impact of introducing a lower earner disregard or tax credit



Notes: See text for description of simulated policies. The bottom panel shows the impact on simulated log wages, regardless of whether the woman is in fact in work or not. Lines show changes relative to the baseline tax-transfer system.

As noted above, we calibrated both policies such that they would have the same fiscal cost if wages and labour supply were unaffected, but when we allow for behaviour to adjust we find that the lower earner disregard costs 56% more than the lower earner tax credit. This in part reflects the fact that the human capital channel is associated with a fiscal externality - when a woman chooses to work full-time and raise her human capital, her future wage goes up which in turn means that she receives fewer transfers and/or pays more in taxes in future.

We now turn to the distributional consequences. There are three mechanisms by which the reform changes incomes. First, there is the direct consequence of raising benefit entitlements, before behavioural change. Second, women change their labour supply, which exerts its own impact on their income. Third, in the long run these changes in labour supply affect their wages. To help illustrate these mechanisms, we show how the reform changes incomes under three simulations. In the first, which we term the 'static' simulation, labour supply is fixed at optimal choices under the baseline system - and therefore so are human capital and wages - so we only include the direct impact of changing benefit entitlements on income. In the second ('Δ labour supply, fixed human capital'), we simulate incomes under the assumption that labour supply choices are as women choose under the reform system, but wages are those they would have under the baseline system. In the third, ('Δ labour supply, Δ human capital') labour supply, human capital, and wages follow what women optimally choose under the reform system. These simulations allow us to provide an accounting decomposition of the total change in income. The difference between the first and second simulation is the "labour supply effect" on income; the difference between the second and third is the "wage effect".[4]

Figure 5 shows simulated impacts of these policies on the household income distribution, under the three simulations described above. In all three subfigures we put individuals into deciles based upon their average lifetime equivalised household income under the baseline system and show effects on average lifetime household income by decile. Two results are worth drawing out. First, while increases in labour supply boost incomes under both reforms (compare the first and second subfigure), changes to wages only increase overall income under the lower earner tax credit reform. In fact, under that reform, wages contribute as much to overall income growth as the increase in labour supply. Because of this, while the income gains in the static scenario are by definition on average identical overall (see the 'all' bars), once behavioural responses are taken into account the income gain under the lower earner tax credit reform is about 40% larger than under the disregard. Second, the behavioural responses change the distributional profile of the reform. In the static analysis the lower earner disregard is clearly more progressive than the lower earner tax credit. Incorporating labour supply changes closes part of that difference as the lower earner tax credit draws more low income women into full-time work. And after accounting for changes to wages the lower earner tax credit looks more progressive still, doing

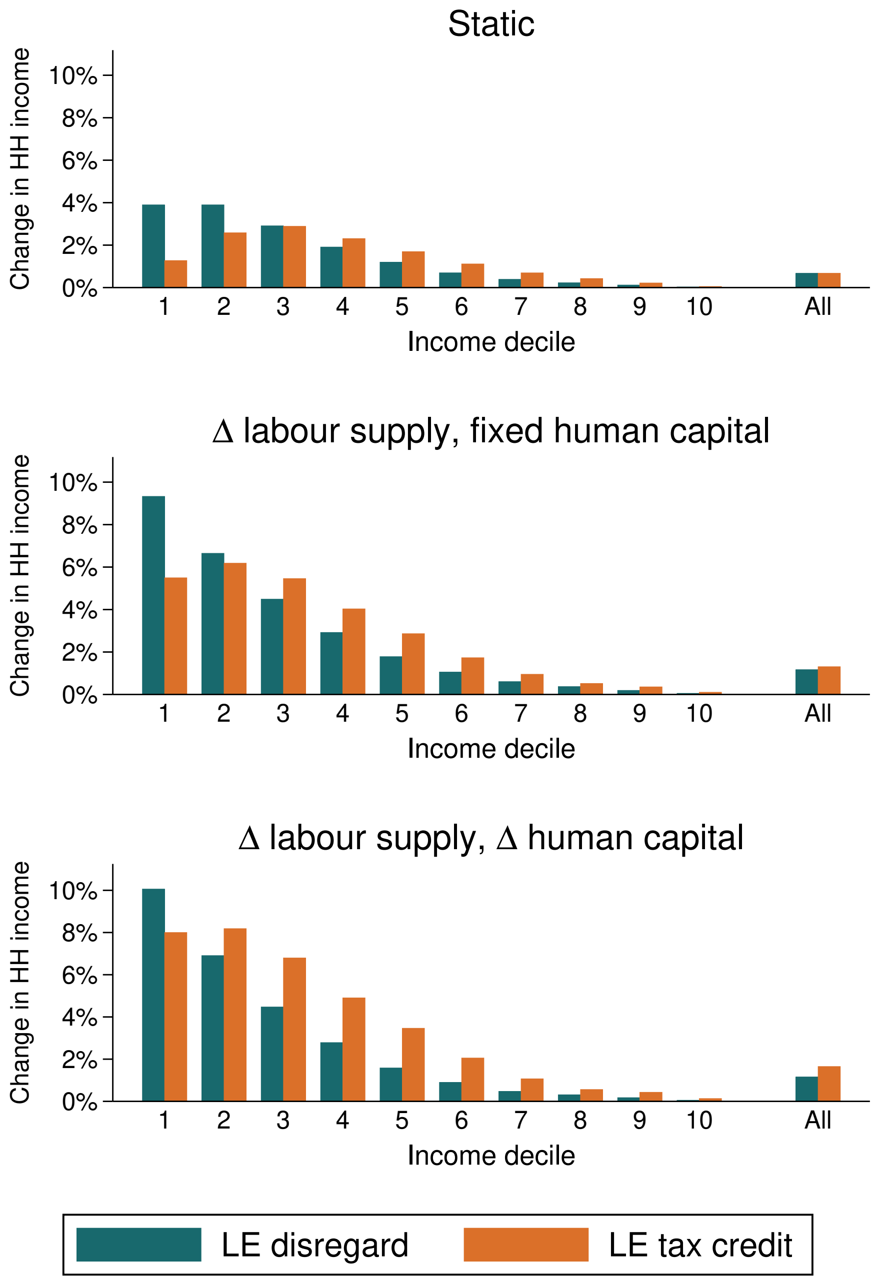
[4]The labour supply choices that the second simulation incorporates are predicated on the wage changes that women know they will experience as a result. Hence, this should be thought of as an accounting decomposition rather than a meaningful counterfactual from an economic point of view - the difference between the first and second simulations show the effect of *any* change in labour supply resulting from the reform, rather than the response that would be observed if there were no human capital accumulation.

more to increase incomes than the disregard policy in each of the deciles other than the very bottom.

We have so far discussed changes in incomes, not welfare. These can differ for two key reasons. First, increases in incomes induced by higher labour supply are also associated with lower leisure, meaning that all else equal the impacts on income may look more positive than the welfare effects. Second, a key motivation for income-contingent transfers is insurance - increasing resources specifically in states of the world where consumption is low - and alternative transfer designs can achieve that aim to different degrees.

We calculate the change in welfare that the reforms induce by estimating the permanent change in consumption that would be necessary to equalise expected utility between the baseline and each policy reform (for further details, see Footnote 51, Blundell et al., 2016). Using this metric the disregard policy increases aggregate welfare by 1.9%; the tax credit policy, 2.5%. This is despite the fact that, after accounting for behavioural effects and consequent human capital accumulation, the tax credit policy costs only two thirds as much as the disregard policy.

Figure 5. Distributional impacts of lower earner disregard and tax credit



Notes: All three figures are constructed based on sample of households that contain a woman aged between 25 and 50. Individuals are assigned to income deciles based on lifecycle average equivalised household income under the baseline system, using the OECD equivalence scale. The bars show change in the average income in each decile. 'LE disregard' and 'LE tax credit' are the lower earner disregard and lower earner tax credit discussed in the text. The 'static' subfigure shows the changes in income that would result if labour supply and wages remained unchanged at the patterns seen in the baseline scenario. The 'Δ labour supply, Δ human capital' subfigure shows incomes accounting for changes to labour supply and wages. The 'Δ labour supply, fixed human capital' subfigure shows incomes under the scenario where labour supply choices are those from the 'Δ labour supply, Δ human capital' simulation (i.e. allowing labour supply to change), while wages are those from the 'static' scenario.

5.2 Changes to existing benefit parameters

In the previous section we investigated policies that quite directly targeted part- and full-time work. One might ask how relevant the results are for the most typical policy reforms - those which simply change the values of existing parameters. We now show that even in these cases the dynamic considerations can be of first order importance when making policy comparisons.

The two reforms we consider are increasing the main disregard (the level of household earnings at which UC begins to be withdrawn) and reducing the taper rate (the speed at which UC is withdrawn as earnings increase). These are the two parameters that govern the phase-out of UC, and both parameters have already been changed multiple times since UC's introduction in 2013. Again we calibrate each policy such that, if behaviour and wages were unaffected, they would each cost approximately the same. We again choose the static cost of both policies to be £4bn per year, as with the policies analysed in Section 5.1.[5]

The policies are superficially similar - both incentivise households to have at least one person in work, and both increase the marginal returns to work for some households already in work and claiming UC. But the two reforms have quite different effects on part- vs. full-time incentives (see Figure B.4 in the Appendix for example budget constraints). We draw out a couple of key differences which drive responses. Raising the disregard represents a fixed cash increase in income for any UC receiving household with earnings above the new disregard. This means that, for example, a non-working household is incentivised to move to part-time work - but, at least so long as that part-time work puts them above the new disregard, it does not provide any additional incentive to work full-time. In contrast, the reduction in the taper rate increases the returns to any increase in earnings for a UC-receiving family (until UC entitlement is exhausted entirely) as the amount a household loses when earnings increase is reduced). Thus, this reform tends to have a more positive impact on full-time work.[6]

We summarise the simulated impacts of these reforms in Table 1, where we also include the policies investigated previously for comparison. In Appendix A, Table B.2 shows the equivalent statistics only for those with a low level of education. As with the effects on the income distribution (Figure 5), we simulate the cost of the policy under three simulations: static (where neither labour supply nor human capital change in response to the reform); 'Δ LS, fixed HC', where only labour supply changes; and 'Δ LS, Δ HC', where both labour supply and human capital change. Again, differences between these cost estimates amount to an accounting decomposition of the role played by labour supply and wage changes. We also provide a welfare calculation, in the same manner as described above.

[5]In the increased disregard experiment, the income disregard is increased by £2,800 per year for all households with children. In the reduced taper rate experiment, the taper rate is cut by 15ppts.

[6]Both reforms will, for some households already on UC in work, have income effects. By extending UC entitlement further up the earnings distribution, both also reduce the marginal returns to work for those who were previously earning a little too much to be entitled to the benefit, creating a substitution effect. Both of these effects tend to reduce labour supply.

Both policies increase female employment, but the higher disregard results in a net shift to part-time work and away from both no work and full-time work, while the taper rate cut results in a shift to both part-time and full-time employment. As a result, increasing the disregard has a small negative effect on wages while cutting the taper rate has a small positive one. The consequence of these behavioural changes is that the taper rate cut costs less than half as much as the higher disregard - despite the fact that their static costs are the same - and delivers a higher level of average income. The accounting decomposition shows that the majority of the difference between the static and full ('Δ LS, Δ HC') costs are accounted for by labour supply changes, but the impact on wages plays an important supporting role. Per pound of government budget, reducing the taper has a considerably larger impact on welfare, with average welfare effects virtually the same despite the taper cut costing far less.

These reforms represent more moderate, but directionally similar, versions of the lower earner disregard and tax credit policies investigated previously. Across both pairs of policies, the two that incentivise full-time work (the lower taper rate and the introduction of the lower earner tax credit) have a more positive effect on wages than the part-time incentivising policies (increase disregard, lower earner disregard) and end up costing the government less than one would think based on a static costing alone. As a result of these effects, welfare calculations suggest that these full-time incentivising policies are strongly favoured.

Table 1. Simulated impact of policy reforms (average across individuals and lifecycle)

	Labour supply						Cost		
	NW	PT	FT	Wages	HH income	Welfare	Static	Δ LS, fixed HC	Δ LS, Δ HC
Baseline level:									
2022-23 UC	19.1%	26.4%	54.5%	£11.7	£51659	-	-	-	-
change vs. baseline:									
Increase disregard	-0.5ppt	1.3ppt	-0.8ppt	-0.3%	0.6%	1.6%	£4.0bn	£5.6bn	£5.9bn
Reduce taper	-1.4ppt	1.0ppt	0.4ppt	0.2%	0.8%	1.4%	£4.0bn	£2.8bn	£2.5bn
LE disregard	-4.7ppt	5.5ppt	-0.8ppt	0.0%	1.2%	1.9%	£4.0bn	£4.9bn	£5.0bn
LE tax credit	-0.3ppt	-3.6ppt	3.9ppt	1.4%	1.7%	2.5%	£4.0bn	£3.9bn	£3.2bn

Notes: Table is constructed based on sample of households that contain a woman aged between 25 and 50. Columns show the average impact across the lifecycle and across all individuals. The labour supply 'NW' column shows the fraction of the population not working; 'PT' the fraction in part-time work; and 'FT' the fraction in full time work. 'Wages' show the average simulated hourly wage all individuals could earn, including among those who are not actually simulated to be working. Income shows average household net income per year (in £). Welfare is calculated as the permanent proportional change in consumption that would be necessary to equalise expected value between the baseline and each policy regime; for further details, see Footnote 51, Blundell et al. (2016). The 'static' cost is the cost to the government to implement the policy if labour supply and wages remained unchanged at the patterns seen in the baseline scenario. The 'Δ LS, Δ HC' cost is the cost to the government accounting for changes to labour supply and wages. The 'Δ LS, fixed HC' cost is the cost to the government under the scenario where labour supply choices are those from the 'Δ LS, Δ HC' simulation (i.e. allowing labour supply to change), while wages are those from the 'static' scenario. Costings account for changes in direct tax receipts and transfers; consumption taxes are not included. 'LE disregard' and 'LE tax credit' are the lower earner disregard and lower earner tax credit policies discussed in the text.

6. Conclusion

The incentives that welfare systems create have attracted a huge amount of academic and policy attention. Increasingly this analysis has turned to longer run effects, including through human capital. Indeed, a key motivation behind the growth of in-work transfers in developed economies has been to encourage participation in paid work and, with it, the accumulation of human capital and wage progression. In this paper we have examined how the design of in-work transfers could be refined further by taking account of the fact that human capital accumulation can depend not only on binary participation decisions but on hours of work. Once we recognise this then the importance of the design of the phase-outs in these schemes - which shape intensive margin incentives - is accentuated.

To study the implications of this we employed a dynamic model of female labour supply and savings, with endogenous human capital accumulation, using the UK's extensive system of in-work transfers as our application. We have shown that reforms which have similar effects on overall employment can still differ markedly in their long-term consequences, if they differ in the intensive margin incentives they provide. A reform that distributes more to part-time workers may, in the short-run, be a more effective tool for boosting the incomes of poor working households than a reform that targets full-time workers. For similar reasons it may also have insurance advantages over an approach more focused on strengthening intensive margin work incentives. But the behavioural effects are crucial in shaping the ultimate comparison between these policy approaches. Not only do reductions in labour supply along the intensive margin mechanically affect contemporaneous earnings; they can also have substantial impacts on human capital accumulation and hence, in the long run, wage levels.

Because higher wages mean an improved government budget, the fiscal externality from policies that strengthen intensive margin work incentives offers the potential for welfare gains. Our results suggest that, indeed, reforms to the phase-out of in-work transfers which strengthen incentives (or reduce disincentives) to work full-time are, relative to part-time incentivising policies, welfare-improving at a given long run cost. Our conclusion is that, while strengthening incentives at the extensive margin was a central and justifiable aim behind the growth of in-work transfers in the first place, a greater focus on the intensive margin offers scope for considerable refinements to their design going forward.

References

Besley, Timothy, and Stephen Coate. 1992. "Workfare versus Welfare: Incentive Arguments for Work Requirements in Poverty-Alleviation Programs." *The American Economic Review*, 82(1): 249–261.

Besley, Timothy, and Stephen Coate. 1995. "The Design of Income Maintenance Programmes." *The Review of Economic Studies*, 62(2): 187–221.

Blundell, Richard, and Thomas Macurdy. 1999. "Chapter 27 - Labor Supply: A Review of Alternative Approaches." In *Handbook of Labor Economics*. Vol. 3, , ed. Orley C. Ashenfelter and David Card, 1559–1695. Elsevier.

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

Brewer, Mike, David Finch, and Daniel Tomlinson. 2017. "Universal Remedy: Ensuring Universal Credit Is Fit for Purpose." Resolution Foundation.

Butler, Adam, and Jill Rutter. 2016. "Creating an Anti-Poverty Childcare System."

Chan, Marc K., and Robert Moffitt. 2018. "Welfare Reform and the Labor Market." *Annual Review of Economics*, 10(1): 347–381.

Huggett, Mark, Gustavo Ventura, and Amir Yaron. 2011. "Sources of Lifetime Inequality." *American Economic Review*, 101(7): 2923–2954.

Jones, Damon, and Ioana Marinescu. 2022. "The Labor Market Impacts of Universal and Permanent Cash Transfers: Evidence from the Alaska Permanent Fund." *American Economic Journal: Economic Policy*, 14(2): 315–340.

Keane, Michael P. 2011. "Labor Supply and Taxes: A Survey." *Journal of Economic Literature*, 49(4): 961–1075.

Keane, Michael P., and Kenneth I. Wolpin. 2007. "Exploring the Usefulness of a Nonrandom Holdout Sample for Model Validation: Welfare Effects on Female Behavior." *International Economic Review*, 48(4): 1351–1378.

Keane, Michael P., and Kenneth I. Wolpin. 2010. "The Role of Labor and Marriage Markets, Preference Heterogeneity, and the Welfare System in the Life Cycle Decisions of Black, Hispanic, and White Women." *International Economic Review*, 51(3): 851–892.

Kleven, Henrik. 2019. "The EITC and the Extensive Margin: A Reappraisal."

Krueger, Dirk, and Alexander Ludwig. 2013. "Optimal Progressive Labor Income Taxation and Education Subsidies When Education Decisions and Intergenerational Transfers Are Endogenous." *American Economic Review*, 103(3): 496–501.

Mirrlees, J. A. 1971. "An Exploration in the Theory of Optimum Income Taxation." *The Review of Economic Studies*, 38(2): 175–208.

Saez, Emmanuel. 2002. "Optimal Income Transfer Programs: Intensive versus Extensive Labor Supply Responses." *The Quarterly Journal of Economics*, 117(3): 1039–1073.

Shaw, Jonathan. 2011. "FORTAX: UK Tax and Benefit System Documentation." W11/08.

Stantcheva, Stefanie. 2017. "Optimal Taxation and Human Capital Policies over the Life Cycle." *Journal of Political Economy*, 125(6): 1931–1990.

University of Essex, Institute for Social and Economic Research. 2021. "Understanding Society: Waves 1-11, 2009-2020 and Harmonised BHPS: Waves 1-18, 1991-2009." UK Data Service.

A. Description of moments used in estimation

Tables A.2 to A.7 display the full list of data moments used in estimation, together with their simulated counterparts and the normalized (by the data standard error) differences between the two. Estimation used 51 moments, which fall into the following categories:

- Mean employment and part-time hours (Table A.2)
- Log wage mean and variance at entrance to working life (Table A.3)
- Log wage regression on lagged wage, log years of work experience and lagged log years of work experience (Table A.4)
- Log wage regression on log age (Table A.5)
- First differences regression of log wages on log years of work experience (Table A.6)
- Log wage mean and proportion of population with wages below pre-defined wage percentiles during working life (Table A.7)

Table A.2. Mean employment and part-time hours

Moment	Data	Simulated	SE	Normed diff.
Low edu.				
Employment	0.625	0.655	0.008	3.656
Part-time hours	0.260	0.231	0.007	4.356
Mid edu.				
Employment	0.817	0.819	0.006	0.473
Part-time hours	0.262	0.256	0.006	0.860
High edu.				
Employment	0.873	0.880	0.004	1.715
Part-time hours	0.198	0.200	0.005	0.346

Table A.3. Log wage mean and variance at entrance to working life

Moment	Data	Simulated	SE	Normed diff.
Low edu.				
Mean	2.062	1.976	0.057	1.521
Variance	0.131	0.226	0.042	2.235
Mid edu.				
Mean	2.204	2.063	0.034	4.153
Variance	0.105	0.159	0.018	3.027
High edu.				
Mean	2.392	2.192	0.030	6.600
Variance	0.127	0.187	0.018	3.301

Table A.4. Log wage regression on lagged wage, log years of work experience and lagged log years of work experience

Moment	Data	Simulated	SE	Normed diff.
Low edu.				
Constant	0.731	0.732	0.039	0.034
Lagged log wages	0.631	0.651	0.019	1.095
Log years of work experience	0.183	0.062	0.047	2.548
Lagged log years of work experience	-0.144	-0.040	0.046	2.267
Mid edu.				
Constant	0.618	0.571	0.026	1.794
Lagged log wages	0.693	0.725	0.012	2.674
Log years of work experience	0.193	0.112	0.044	1.825
Lagged log years of work experience	-0.146	-0.077	0.043	1.619
High edu.				
Constant	0.736	0.767	0.032	0.964
Lagged log wages	0.668	0.665	0.012	0.261
Log years of work experience	0.264	0.183	0.044	1.854
Lagged log years of work experience	-0.199	-0.130	0.039	1.758

Table A.5. Log wage regression on log age

Moment	Data	Simulated	SE	Normed diff.
Low edu.				
Constant	1.931	2.136	0.050	4.110
Log age	0.101	0.035	0.016	4.010
Mid edu.				
Constant	1.876	2.151	0.040	6.865
Log age	0.179	0.089	0.014	6.409
High edu.				
Constant	2.010	2.156	0.040	3.627
Log age	0.235	0.190	0.014	3.182

Table A.6. First differences regression of log wages on log years of work experience

Moment	Data	Simulated	SE	Normed diff.
Low edu.				
Differenced log years of work experience	0.100	0.197	0.041	2.397
Mid edu.				
Differenced log years of work experience	0.114	0.185	0.032	2.250
High edu.				
Differenced log years of work experience	0.183	0.230	0.021	2.242

Table A.7. Log wage mean and proportion of population with wages below pre-defined wage percentiles during working life

Moment	Data	Simulated	SE	Normed diff.
Low edu.				
Mean	2.240	2.244	0.004	0.812
Wages below 10th percentile	0.100	0.111	0.003	3.477
Wages below 25th percentile	0.250	0.249	0.005	0.255
Wages below 50th percentile	0.500	0.448	0.005	9.686
Wages below 75th percentile	0.750	0.727	0.005	4.912
Wages below 90th percentile	0.900	0.928	0.003	8.649
Mid edu.				
Mean	2.404	2.414	0.004	2.698
Wages below 10th percentile	0.100	0.106	0.003	2.436
Wages below 25th percentile	0.250	0.240	0.004	2.519
Wages below 50th percentile	0.500	0.469	0.004	7.130
Wages below 75th percentile	0.750	0.750	0.004	0.000
Wages below 90th percentile	0.900	0.910	0.003	3.875
High edu.				
Mean	2.696	2.710	0.004	3.707
Wages below 10th percentile	0.100	0.084	0.002	6.511
Wages below 25th percentile	0.250	0.252	0.004	0.463
Wages below 50th percentile	0.500	0.519	0.004	4.608
Wages below 75th percentile	0.750	0.751	0.004	0.143
Wages below 90th percentile	0.900	0.890	0.002	4.063

B. Additional tables and figures

Table B.1. ESTIMATED LABOUR SUPPLY ELASTICITIES BY DEMOGRAPHIC GROUP

	Frisch elasticity	Marshall elasticity
All women	0.346	0.274
Low edu.	0.517	0.367
Mid edu.	0.337	0.269
High edu.	0.221	0.206
Singles, no kids	0.262	0.072
Single mothers	0.676	0.511
Couples, no kids	0.103	0.100
Mothers in couples	0.439	0.440

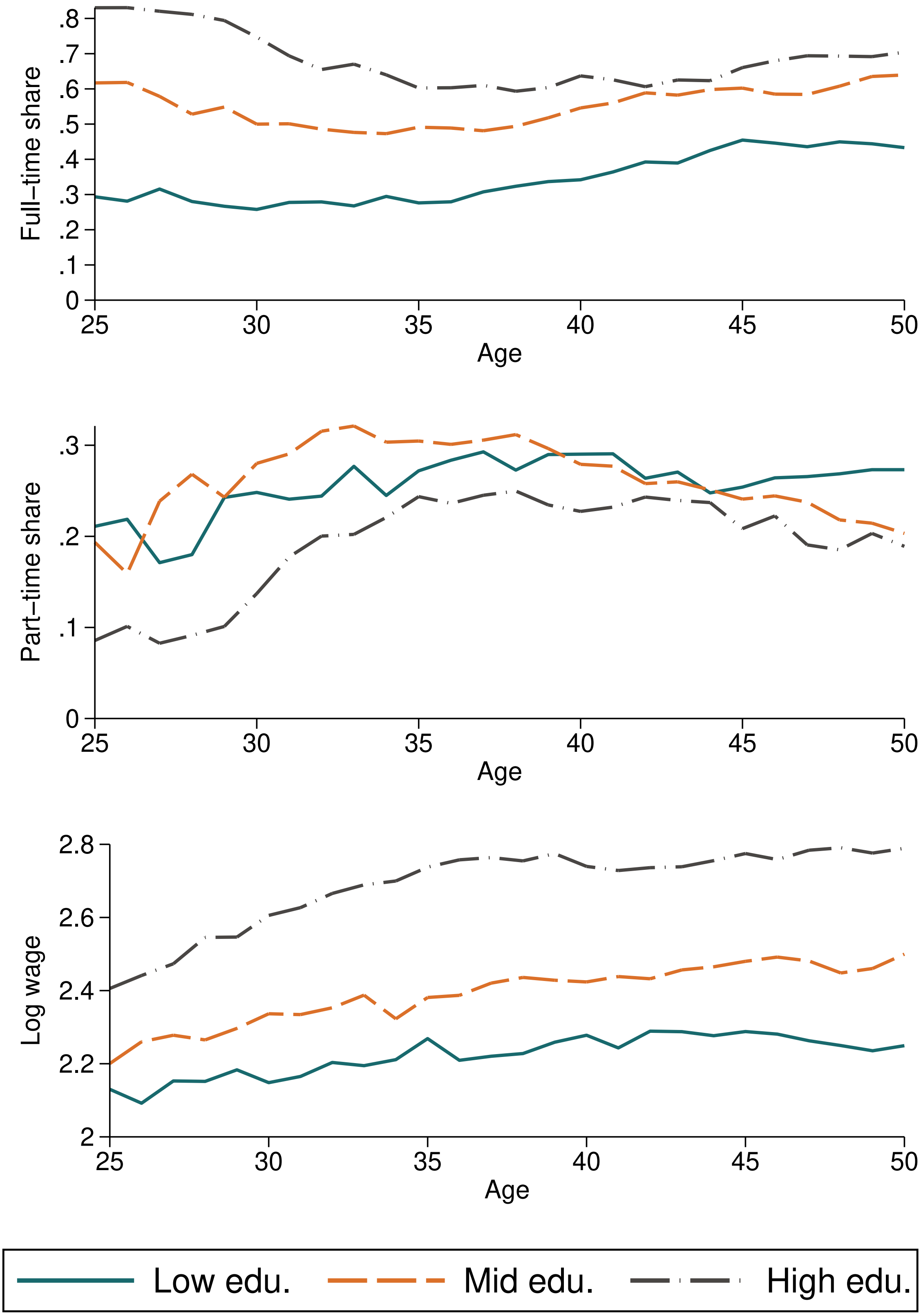
Notes: Elasticities shown are with respect to hours - i.e., the proportional change in total hours in response to a change in wages. Marshall elasticities are calculated by perturbing the entire profile of wages from the indicated age onwards; they therefore document the labour supply response this period to an expected, permanent increase in wages, and reflect both substitution and income effects. Frisch elasticities are calculated by perturbing wages in one period only. We provide transfers to compensate for wealth effects. Frisch elasticities document the labour supply response this period to an expected, temporary increase in wages.

Table B.2. SIMULATED IMPACT OF POLICY REFORMS FOR LOW EDUCATION

	Labour supply					
	NW	PT	FT	Wages	HH income	Welfare
Baseline level:						
2022-23 UC	29.6%	31.2%	39.1%	£8.9	£43891	-
% change vs. baseline:						
Increase disregard	-1.1ppt	1.1ppt	0.1ppt	0.1%	1.2%	2.2%
Reduce taper	-1.5ppt	0.7ppt	0.9ppt	0.3%	1.1%	1.7%
LE disregard	-7.2ppt	6.6ppt	0.5ppt	1.0%	2.1%	2.5%
LE tax credit	0.0ppt	-7.0ppt	7.0ppt	3.0%	2.6%	2.7%

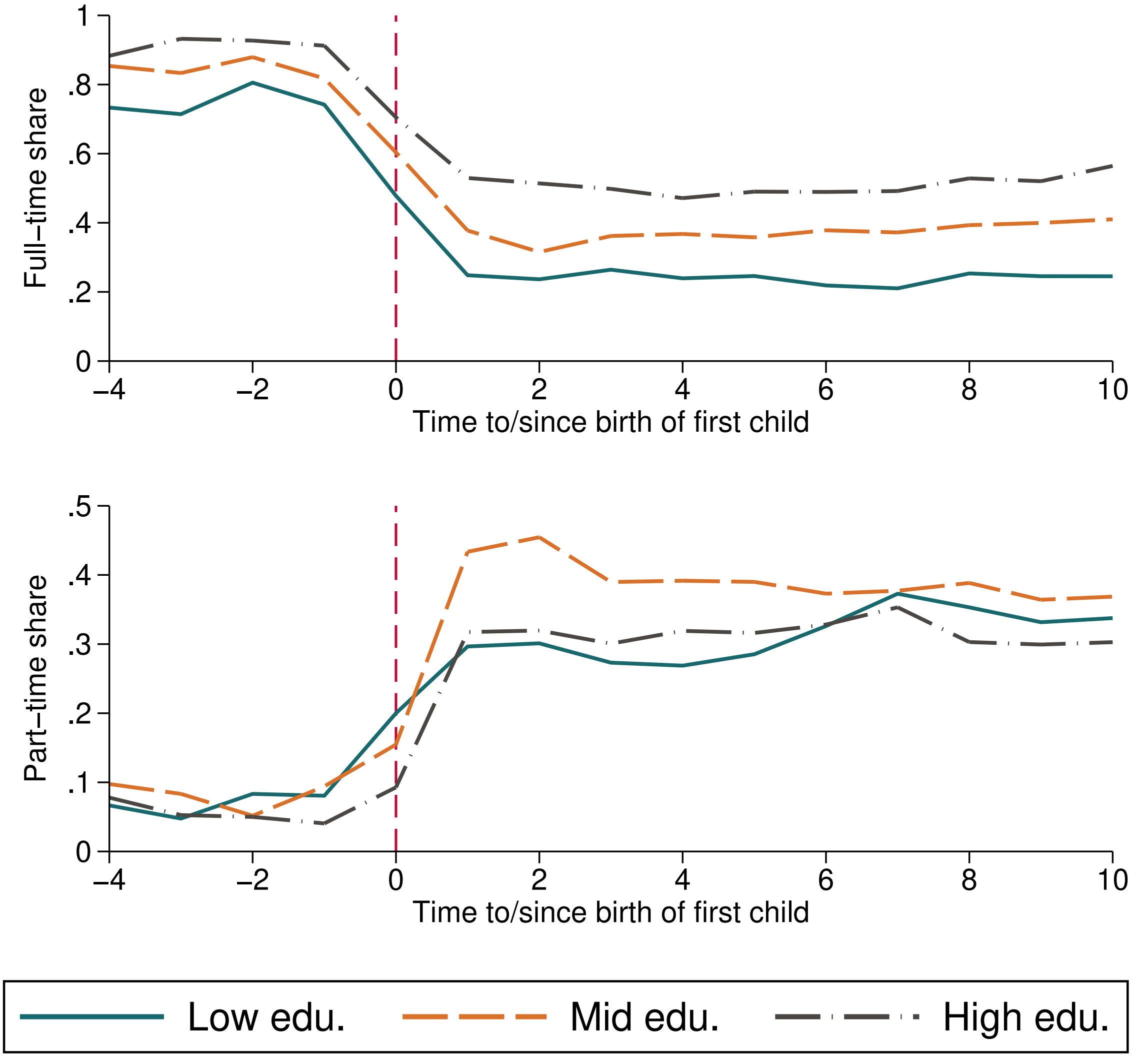
Notes: See Table 1. Unlike that table, here averages are shown for the low educated group alone.

Figure B.1. LABOUR SUPPLY AND WAGES FOR WOMEN OVER THE LIFECYCLE



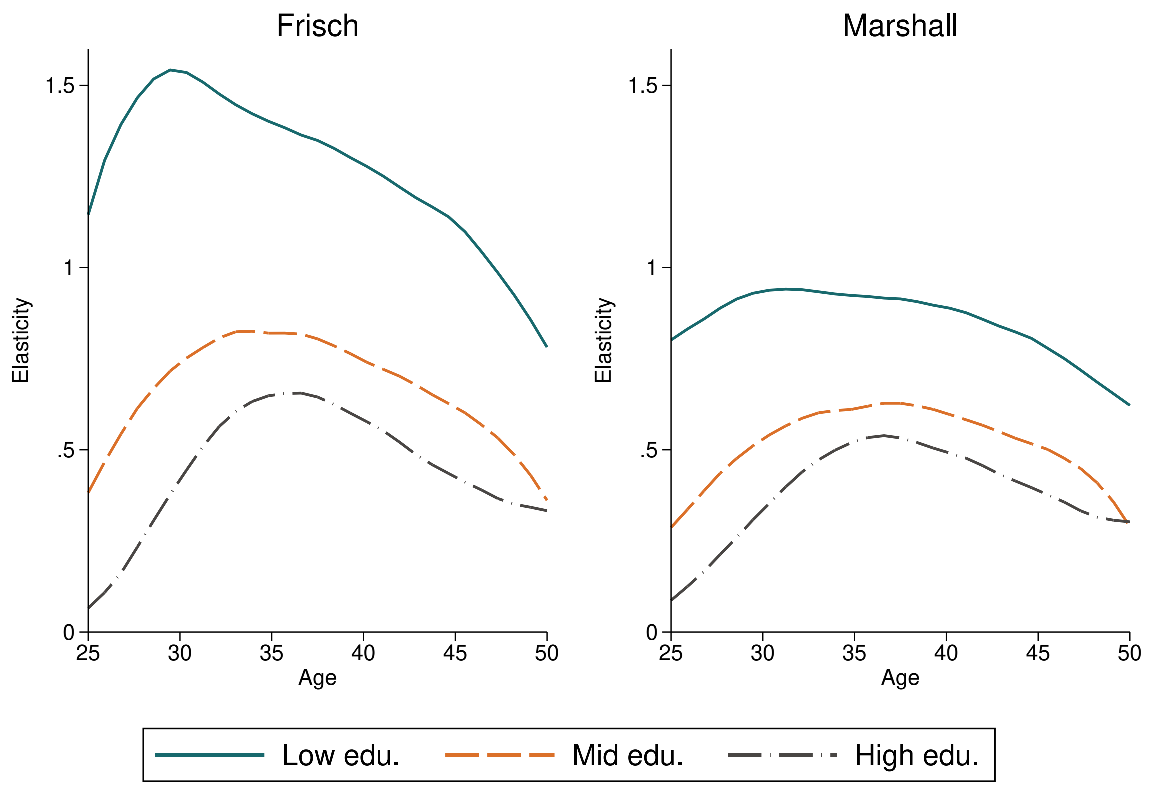
Notes: Calculations from UKHLS data. Part- and full-time shares show the fraction of women working part- and full-time respectively. The log wage subfigure shows the average log wage among women observed in-work.

Figure B.2. Relationship between birth of first child and labour supply



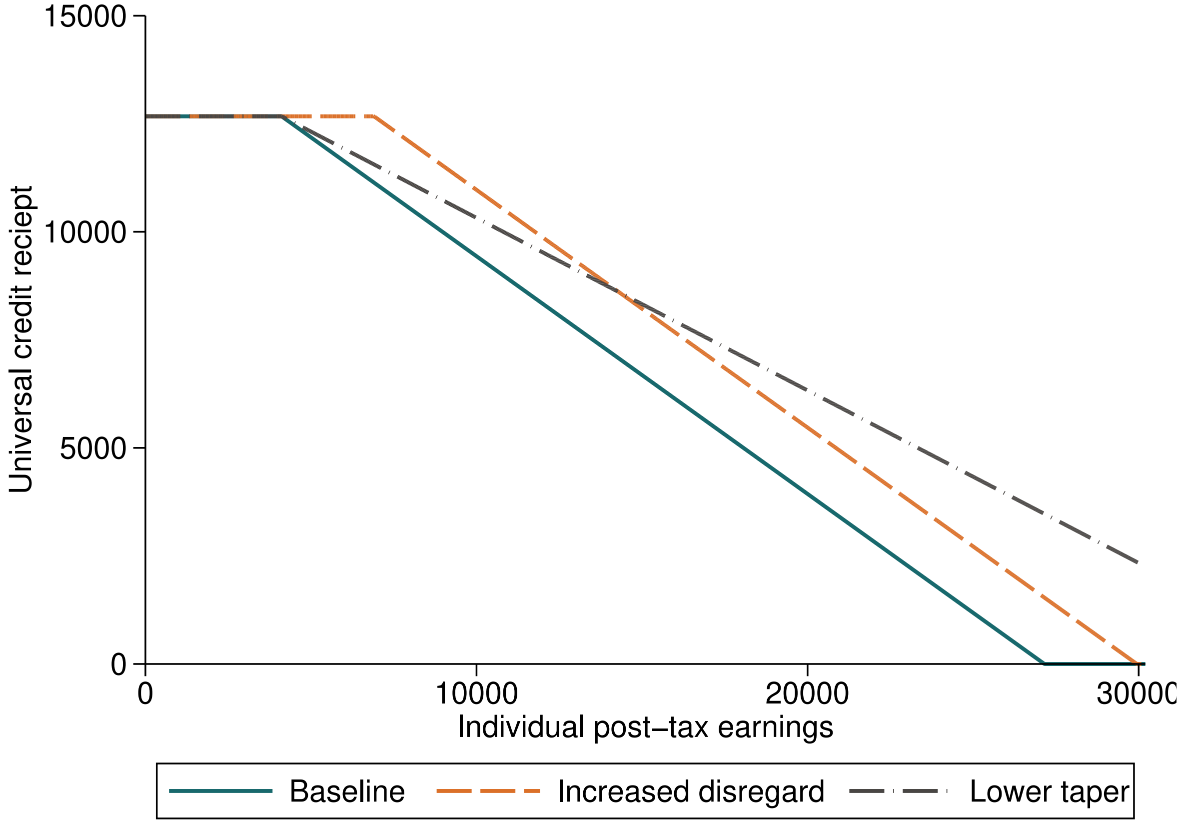
Notes: Calculations from UKHLS data. Part- and full-time shares show the fraction of women working part- and full-time respectively. Birth of first child is measured based on the first survey wave in which a household is observed transitioning from having no children to having one or more children. Negative values are constructed by identifying these households in prior waves. To improve sample size, we also use age of the oldest child in the household to measure time since birth of first child among those households who we do not observe transitioning from having no children to having one or more children.

Figure B.3. ESTIMATED LABOUR SUPPLY ELASTICITIES OVER THE LIFECYCLE



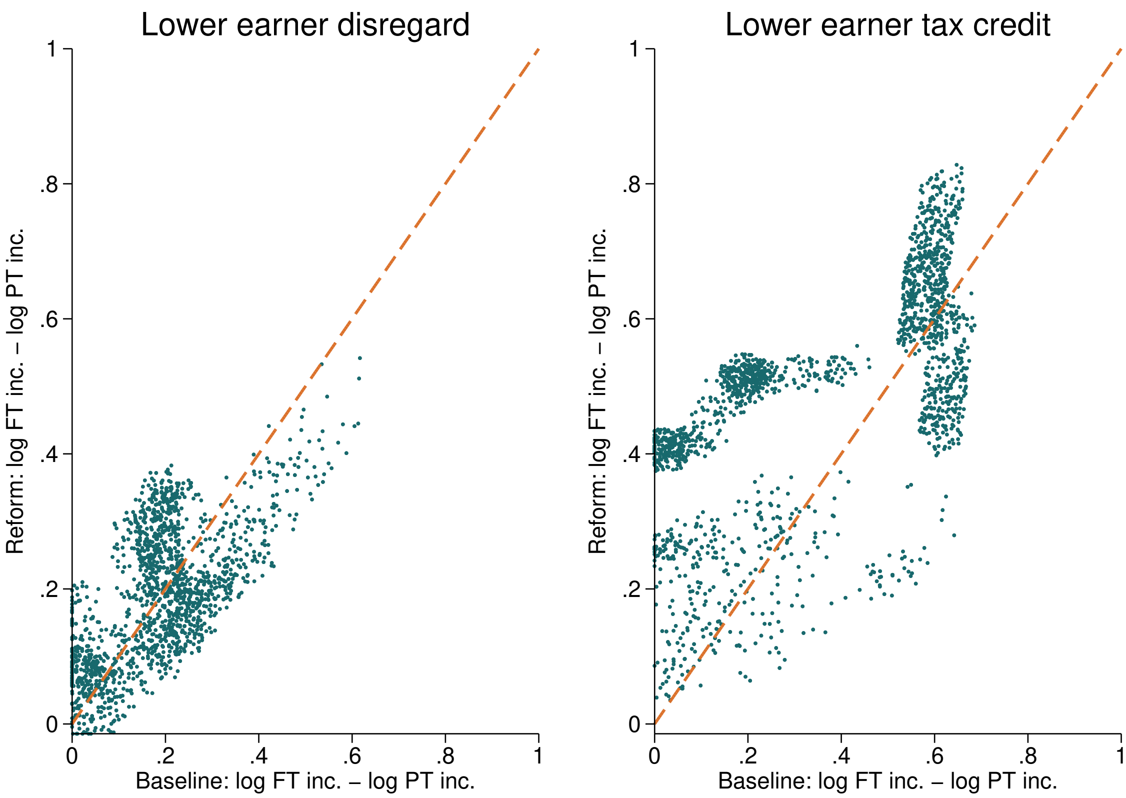
Notes: Elasticities shown are with respect to hours - i.e., the proportional change in total hours in response to a change in wages. Marshall elasticities are calculated by perturbing the entire profile of wages from the indicated age onwards; they therefore document the labour supply response this period to an expected, permanent increase in wages, and reflect both substitution and income effects. Frisch elasticities are calculated by perturbing wages in one period only. We provide transfers to compensate for wealth effects. Frisch elasticities document the labour supply response this period to an expected, temporary increase in wages.

Figure B.4. UC ENTITLEMENT FOR AN EXAMPLE SINGLE MOTHER UNDER DIFFERENT REFORMS



Notes: All values shown are annualised (UC is in fact assessed at a monthly frequency). Example household is a single woman with weekly rent of £100 and two children, at least one of whom was born before April 2017 and so the family is eligible for the 'family element' in UC. The horizontal axis shows the after tax (but pre-transfer) earnings of the woman. 'Baseline' is the actual 2022-23 tax and benefit system.

Figure B.5. Financial incentives to move from part- to full-time work under different policy reforms



Notes: In both subfigures, the horizontal axis shows the change in a woman's log household income if she moved from part-time to full-time work under the baseline system - a measure of the intensive margin (substitution) incentive. The vertical axis shows the same statistic under the reform system (lower earner disregard or lower earner tax credit for the left and right figures respectively). The orange dashed line is the 45 degree ray. Women above the line see a strengthened incentive to move from part- to full-time work following the reform; women below the line a weakened incentive. For visual clarity, women whose returns to part- and full-time work are the same under the baseline and reform system (and so lie on the 45 degree ray) are excluded from the sample. We also add a small amount of noise to each observation to make it easier to see observations that are on top of one another. Returns to work are calculated using simulated incomes at age 30. For the reform systems, returns to work are calculated by simulating that the reform is unexpectedly introduced at age 30 (thereby avoiding the reform having any impact on behaviour).