Has increased women's educational attainment led to greater earnings inequality in the UK?

Author(s): Breen, Richard, 1954-;Salazar, Leire, 1975-
Date 2005
Type Working Paper
Series Estudios = Working papers / Instituto Juan March de Estudios e Investigaciones, Centro de Estudios Avanzados en Ciencias Sociales 2004/216
City: Madrid
Publisher: Centro de Estudios Avanzados en Ciencias Sociales

HAS INCREASED WOMEN'S EDUCATIONAL ATTAINMENT LED TO GREATER EARNINGS INEQUALITY IN THE UK?

Richard Breen and Leire Salazar

Estudio/Working Paper 2005/216
June 2005

Richard Breen is Official Fellow at Nuffield College, Oxford University and Fellow of the British Academy. He is also Member of the Scientific Committee of the Center for Advanced Study in the Social Sciences, Juan March Institute, Madrid. Leire Salazar is a PhD candidate at the Center for Advanced Study in the Social Sciences, Juan March Institute, Madrid.

Abstract

It is widely believed that the growth in women's educational attainment and their increasing labour force participation, together with educational homogamy, will lead to greater inequality between households in their earnings. In this paper we use data from the United Kingdom to test that assertion.

Introduction[*]

In recent decades, there has been a resurgence of interest in the topic of earnings and income inequality, not only in economics but also in sociology (Alderson and Nielsen 2002; Firebaugh 2003). Within the latter this has coincided with an emphasis on microprocesses or micromechanisms, which, among other things, may be invoked to explain links between variables at an aggregate level (Hedström and Swedberg 1998). Thus, renewed efforts have been to explain *why* and *how* inequalities in the personal distribution of resources are generated and vary over time and space, rather than merely describing them (Atkinson 1997). Even for addressing aggregate outcomes such as income or earnings inequality, the explanation is sought at a micro level, where reference is made to individual decision making processes (and to decisions of individuals in a given household).

In this paper, we seek to explain changes in earnings inequality among households in the United Kingdom by considering the effects of women's increasing educational attainment. We ask: *How has women's increasing educational attainment affected the distribution of earnings and, specifically, earnings inequality among households*[1] *and how much of the change in inequality between households can be attributed to this?*

Educational attainment does not affect individual earnings only via its impact on earnings as one part of human capital: it can also have effects on other processes that ultimately affect the earnings distribution. First, increasing educational attainment can be expected to lead to greater labour force participation among women, and thus women's earnings will become more important in explaining total inequality, regardless of whether they have an equalising or disequalising effect. Female labour force participation rates have increased – albeit to different degrees – in all industrialised countries except Sweden and

[*] We thank Sir Tony Atkinson and participants at a seminar given to the Bedford Group, Institute of Education, London in December 2004 for comments on an earlier draft of this paper and Sir David Cox for helpful discussions.

[1] Throughout the paper, the focus is only on earnings (i.e. labour income). Earnings have been found to be the most important component of income (for most people it is actually the only one) and we therefore leave aside wealth, capital gains, and other similar assets. The available evidence seems to suggest that the exclusion of these types of non-labour income would not challenge the conclusions drawn by the examination of earned income. See Atkinson 1975 for a discussion.

Finland[2] in the last three decades (Callan et al 1998). In addition, the share of households in which women contribute substantially to household income through their own earnings has rapidly increased everywhere. But whether the effect is to increase or reduce inequality will depend on the distribution of women's participation across different kinds of households, and this may vary over time, so influencing the trend in inequality. If *ceteris paribus* women belonging to households towards the bottom of the earnings distribution tend to participate more than women at the top, then there could be an equalising effect on inter-household earnings distribution. Alternatively, if women belonging to better-off households participate in paid labour in greater proportions, then it could be the case that an unequalising effect is at work.

Secondly, if members of couples tend to have similar characteristics (for instance educational attainment) that allow them to get certain returns in the labour market (earnings), and if educational homogamy increases (particularly at higher levels of education), then a reinforcement of inequalities could take place (Drobnič and Blossfeld 2001: 380-3; Blossfeld and Timm 2003: 341-2). Increases in women's educational attainment, to the extent that this equalizes the distribution of education between the sexes, can be expected to increase the proportion of educationally homogamous households. On the other hand, if the educational attainments of both sexes are increasing, the effect on the overall level of homogamy will be indeterminate, but it seems likely that we will observe a declining rate of homogamy among couples with low levels of education and an increasing rate among those with high levels.

Thirdly, increasing education and labour force participation may change the distribution of household types which may, of itself, affect inequality. Burtless (1999), for example, has pointed out the relevance of changes in household composition through changing marriage patterns on earnings inequality in the US. The two main factors contributing to this change are the decline and delay in fertility and the increase in the proportion of persons who remain unpartnered. Women's educational expansion seems likely to cause change in both of these, as the feasibility of establishing a household without a male partner increases and the opportunity cost of children also increases.

[2] In any case, in these two countries women's participation levels reach the highest possible levels.

To test the effects of women's educational expansion on changes in the distribution of earnings among households, we use data from the UK. On the one hand, the UK has experienced a very notable expansion in upper secondary and higher educational levels and important increases in the proportion of women in each successive cohort entering them (Shavit and Blossfeld 1993; Green et al 1999). On the other hand, this country showed the highest European levels of income and earnings inequality together with the steepest increase during the eighties and the early nineties (Atkinson et al 1995). This particular combination of high (female) educational expansion and high (and increased/increasing) inequality makes this an interesting test case.

In the next section of the paper we describe our data and the counterfactual estimations that we make to assess the effect of increasing female educational attainment on inequality among households in the distribution of earnings. We present some descriptive statistics, followed by the results of our counterfactuals. We then investigate some of the other factors that have led to a growth in earnings inequality, and the paper concludes with an assessment of the relative importance of increasing women's educational attainment for inequality.

Data and methods

The data we use come from the UK Family Expenditure Survey (FES). The FES is a continuous, nationally representative, cross-sectional study running from 1957 onwards and collecting detailed information on individual and household income and expenditure. We use the surveys for 1979, 1991 and 2000: these cover the range of years that witnessed the most marked increase in inequality in the UK.

Since the unit of analysis is the household, and because the link that we are trying to establish is between women's education and household earnings, we considered both earnings and educational information only for the head or for the partner of the head of the household. This means that the earnings and education of other earners in the household (e.g. children or other adult members) are ignored: thus our analysis deals only with inequality in

the earnings of heads of households and their partners. Furthermore, the FES does not provide information on educational levels as such, but rather, on the age at which full-time education ceased. Only for those individuals still in formal education at the moment of the survey (and only for the 1979 cross-section) it is possible to identify the level in which they are enrolled. We therefore use a simple categorisation of education comprising three broad levels:

Level 1 = left full-time education before age 16.
Level 2 = left full-time education between the ages of 16 and 18.
Level 3 = left full-time education after the age of 18.

Thus level 1 is made up of those who, for the most part, had no more than compulsory education, while level 2 contains those with O-level or equivalent qualifications, and those with A-level. Level 3 comprises individuals with a qualification higher than A-level and thus includes all those with a degree. We then add a fourth category to our educational variable, which applies to household types without a person of that sex (i.e. these are household with an unpartnered head).

As well as education we also distinguish whether or not the household head and her/his partner is working. By definition (see above) households in which neither is working will have a zero income. In reality, the earnings of these households may, in fact, be non-zero because of the presence of other earners (who may also be present in households where the head and/or his/her partner are working). In the appendix we report some statistics comparing household earnings using the restrictive definition of earners we adopt in this paper and the case in which the earnings of all earners are included. One effect of adopting the latter is that some households where neither the head nor the head's partner is working have very large Theil values, based on very small numbers. But in general, the differences between the two approaches are minor: the contribution of other earners seems to be fairly constant across different types of households, and the trend in inequality in them is much the same whether other earners are taken into account or not. In fact, the major results of our counterfactuals are robust to the choice of either definition.

Earnings are weekly wages and salaries from all (main and subsidiary if applicable) sources of employment (including income from self-employment) for the household head and partner (if any). The earnings figures throughout the paper are expressed in constant (1992) pounds sterling. Household earnings have been adjusted by household size and composition (adults and children) using the ***modified OECD scale***, which assigns a weight of 1 to the first adult, 0.5 to each additional adult (15 or older) and 0.3 to each child (under 15).

Because we focus on earnings (and not income), we have sought to exclude households whose main members (heads and partners of the heads) are retired or still enrolled in education and so we limit inclusion in our sample to households in which the head is in prime-working age (20 to 64). But we then carry out a further set of analyses in which we confine our attention to young households (where the head is between 25 and 34 years old) in order to discover whether the effects of increasing women's education are more evident in this age group than in the whole working population.[3]

The data for our analyses comprise, for each of 1979, 1991 and 2000, a 4-way table of woman's education (W_E) by man's education (M_E) by whether or not the woman works (W_W) by whether or not the man works (M_W). This is an incomplete cross-classification because 16 of the possible 64 combinations cannot be observed: furthermore, another 15 cells have zero entries because we have limited earnings to those of the household head and partner. Each cell of this table then contains three pieces of information: the number of households in it, n(ijkl), the mean earnings of those households, $\bar{x}_{ijkl}$, and a measure of earnings inequality among them, T(ijkl), where i, j, k and l index W_E, M_E, W_W and M_W, respectively.

The measure of inequality which we use is the Theil index. This meets all the desirable properties of an inequality measure[4]. If household earnings are x and households are indexed by i, then the Theil index is given by

[3] There are obvious reasons to think that this might be the case, given that changes in the distribution of education are likely to be most pronounced within this age group.

[4] See Bourguignon 1979 for a thorough discussion of different measures and their properties.

$$T = \frac{1}{n}\sum_i \frac{x_i}{\bar{x}} \ln\left(\frac{x_i}{\bar{x}}\right) \quad (1)$$

where $\bar{x}$ is the overall mean earnings. However, the most appealing property of the Theil index as a measure of inequality is its additive decomposability. This means that if the total population is exhaustively split into mutually exclusive sub-groups, then the total index (i.e. overall inequality) can be perfectly decomposed into a between-group and a within-group component. Applications of this approach are numerous, and include areas such as gender, occupational and regional inequalities. It has also been used to decompose incomes into different income sources (Shorrocks 1984). Between-group inequality is interpreted as the share of total inequality that arises through variation in the average earnings of different sub-groups: the within-group component is that part of overall inequality that is due to heterogeneity in earnings among observations within each of the sub-groups. The index decomposes as follows:

$$T = \sum_j p_j \frac{\bar{x}_j}{\bar{x}} \ln\left(\frac{\bar{x}_j}{\bar{x}}\right) + \sum_j p_j \frac{\bar{x}_j}{\bar{x}} T_j \quad (2)$$

where, for the purposes of this paper, $\bar{x}_j$ is the mean earnings in the j^{th} subgroup, T_j is the Theil value for that subgroup and p_j is the proportion in each type. The first term is the between type inequality while the second term is a weighted average of the within type inequalities. In our case, the subgroups are household types defined as the Cartesian product of the four variables W_E, M_E, W_W and M_W, and so, for our purposes, the subscript in (2) should be written ijkl rather than j, though we continue to use the latter for notational convenience .

We carry out two counterfactual analyses and one decomposition to explain change between 1979 and 1991 and between 1991 and 2000. In the first counterfactual analysis we simply allow one or more of the three terms in equation (2) (that is, p, $\bar{x}$ and T_j) to take their t+1 value while keeping the remaining terms at their t value. It is well known that the results of this method may be sensitive to the order in which the terms are allowed to take their

counterfactual values, and so we carry out the analyses using all different possible orderings. But we also use the decomposition method presented by Mookherjee and Shorrocks (1982). This is an exact decomposition of the change in inequality as follows:

$$\Delta T = \sum_j \Delta\varphi_j \bar{p}_j + \sum_j \Delta p_j \bar{\varphi}_j + \sum_j \Delta p_j \bar{\gamma}_j + \sum_j \Delta\gamma_j \bar{p}_j \quad (3)$$

Here Δ means the change between t and t+1, $\varphi_j = \frac{\bar{x}_j}{\bar{x}} \ln\left(\frac{\bar{x}_j}{\bar{x}}\right)$, and $\gamma_j = \frac{\bar{x}_j}{\bar{x}} T_j$. The bars indicate mean values over t and t+1: e.g. $\bar{p}_j = \frac{p_j(t) + p_j(t+1)}{2}$. The four terms in (3) can then be interpreted as (a) the effect, on the change in between-group inequality, of a change in the inequality in mean earnings between subgroups; (b) the effect on the change in the between group inequality of the changing distribution of subgroups within the population; (c) the effect on within group inequality of the changing distribution of subgroups; and (d) the effect, on within-group inequality, of the change in the weighted Theil value for each subgroup.[5]

Both the counterfactual and the decomposition tell us the relative importance of changes in the overall distribution of p for inequality: however, they do not tell us the importance of changes in each of the four variables that define our household types, and for this we move to another set of counterfactuals through which we seek further to decompose the total effect of the changing distribution of household types (i.e. subgroups) into the effects of change in the variables W_E, M_E, W_W and M_W and in the associations between them. For this purpose we use the Deming-Stephan method (see the appendix) which allows us to adjust the univariate, bivariate or trivariate distributions of the variables to conform to any desired configuration. Previous attempts to decompose inequality into between- and within-group components have involved a univariate distribution of groups (as in Mookerjee and Shorrocks 1982 where the groups are defined by age); or a multivariate distribution in which each variable is treated entirely separately (as in Jenkins 1995); or a multivariate distribution

[5] It should be noted, however, that this method leads to an underestimate of the effect of change in p because the quantity $\bar{x}$ is itself dependent on p.

which is gradually built up, from an original univariate distribution, by the addition of successive variables (as in Cowell and Jenkins 1995). The use of the Deming-Stephan method is much more flexible than any of these, and lets us uncover a more detailed picture of the effects of counterfactual changes in the variables defining the groups. So, in our counterfactuals we can allow, for example, the univariate distribution of W_E to take its t+1 value, while preserving the distributions of the other variables, and all the associations between variables (including those involving W_E) at their t values. This avoids the consequence of more usual counterfactual methods in which a change in W_E, for example, would also induce a change in the distributions of the remaining three variables. Our strategy is (1) to allow the marginal distribution of women's education to take its t+1 value, holding all else constant at their t values; (2) we also allow the association between women's and men's education to take their t+1 values; (3) we also allow the association between women's and men's education and whether women are working or not to take their t+1 values[6]; and, (4), we let all variables and their associations take their t+1 values. The logic of this is that, in step (1) by allowing only the marginal distribution of W_E to change, we are showing what 'would' have happened had there been only a change in women's education and not in any of the other variables, nor in the behavioural consequences of women's education for household formation or labour force participation. In step (2), we add the possible effects of such behavioural change by letting the pattern of association between men's and women's education change as it is observed, so capturing changes in educational assortative household formation. In step (3) we add the effects of behavioural change (in whether women are working or not) within each household type. Together, steps (1) to (3) seem to us to capture the main mechanisms by which, directly and indirectly, changes in women's education might influence the distribution of household types in the population. From (3) we move directly to a counterfactual in which we allow the p_{ijkl} to take their observed t+1 values. The major change that this induces is that we now allow whether or not men are working to take its t+1 value.

There are two further complications in this procedure. First, in the initial step, our interest is in the effect of a change in the distribution of women's education per se and not in

[6] This is a shorthand way of saying that all the two-way associations between pairs of these three variables and their three-way interaction are allowed to take their t+1 values.

any changes in the distribution of households types that does not follow directly from this. We therefore set the marginal total for the fourth category of W_E ('not present') to its value at t (rather than t+1), and the counterfactuals for the three other categories were adjusted accordingly. Secondly, in step (2) the question arises of whether, as well as allowing the association between women's and men's education to change, we should also allow the distribution of the latter to change. Similarly, in step (3), when we allow the associations between W_E, M_E and W_W to change, should we also allow the marginal distributions of the last two to take their t+1 values? It would seem that, in (2), we should keep M_E at its t value, because change in men's education cannot be viewed as a consequence of change in women's education; on the other hand, whether or not women participate in paid work might be considered such a consequence. In fact, we carried out counterfactuals allowing for all possibilities, and the results did not prove sensitive to our choice. We therefore report the results in which the univariate distribution of W_W is allowed to change in (3), and that of M_E is allowed to change in both (2) and (3).

Descriptive results

Our sample confirms the already well-established evidence that the UK is no exception to the general expansion in women's education in most industrialised countries during the recent decades. Table 1 shows a very significant increase in medium and high educational levels in the period that our data cover. Whereas in 1979 more than 3 out of 5 women had the lowest educational level (i.e. left full-time education at the age of 15 or younger), by 2000 only 1 out of 4 had these qualifications. The proportion of those with some university education almost tripled during these two decades. Amongst the youngest cohort (those in households where the head was aged 25 to 34 at the time of the survey), women's educational expansion is even more remarkable, with a very small percentage of women holding the lowest level by the end of the period.

Table 1. *Women's educational levels over time.*

	Complete sample			Age of head: 25-34		
	1979	1991	2000	1979	1991	2000
Low	61.94	39.09	24.37	47.24	11.44	9.22
Medium	30.47	46.83	55.60	40.22	71.86	64.16
High	7.60	14.09	20.03	12.54	16.70	26.62
N	4516	4132	4468	1268	988	1052

We have argued above that women's increasing education might affect household earnings inequality through a number of possible mechanisms. Firstly, as regards household formation, we expect a) the proportion of households with an unpartnered head, and b) the proportion of educationally homogamous couples to have increased, particularly at the two highest levels. Secondly, we anticipated some growth in women's labour force participation and employment. In the tables that follow, we illustrate these changes.

Table 2 shows that the proportion of households with an uncoupled head has indeed increased since the late seventies, although there is hardly any change between 1991 and 2000. The growth in non-partnership is even more notable in the youngest cohort, although this could partly be due to increasing delays in couple formation.[7]

Table 2. *Percentage of households headed by an uncoupled person over time.*

	Complete sample			Age of head: 25-34		
	1979	1991	2000	1979	1991	2000
% Uncoupled	24.05	37.32	38.17	19.21	42.13	41.08
N	4964	5113	5243	1385	1334	1244

[7] These trends also reflect the growth in the rate at which partnerships dissolve to leave single-person headed households.

The percentage of couples in which both partners have the same educational level has evolved in the expected manner (Table 3). Homogamy at the intermediate and highest levels has steadily increased whereas in fewer couples both the man and the woman have the lowest qualifications. This pattern is even more visible amongst the youngest couples. Overall, however, the percentage of educationally homogamous marriages declined in the complete sample from 70 per cent in 1979 to 66 per cent in 2000, while for the youngest households it increased from 63 to 68 per cent. Within households in which both partners work, the correlation between their earnings has steadily increased from 0.07 in 1979 to 0.13 in 1991 and to 0.31 in 2000 among our complete sample, with the comparable figures for younger households being 0.06, 0.23 and 0.23.

Table 3. *Percentage of educationally homogamous couples by educational levels over time.*

	Complete sample			**Age of head: 25-34**		
	1979	1991	2000	1979	1991	2000
Low	50.85	28.27	15.18	35.12	2.98	2.05
Medium	15.12	29.80	38.59	20.73	59.72	49.52
High	4.06	8.61	12.25	7.15	9.07	16.51
N	3770	3205	3242	1119	772	733

Finally, we see in Table 4 the change in women's employment. The proportion of households in which the female head or partner of the head is involved in paid work (either employed or self-employed) follows a u-shaped trend, declining then rising, though by 2000 it remained slightly lower than its 1979 level. This trend is explained by the high unemployment of the early 1990s and the subsequent growth in jobs for women during the 1990s.

Table 4. *Women's employment over time.*

	Complete sample			Age of head: 25-34		
	1979	1991	2000	1979	1991	2000
% Working	57.90	50.62	55.10	57.62	47.15	55.47
N	4964	5113	5243	1385	1334	1244

Table 5 sheds more light on this trend (which will later prove important in explaining changes in earnings inequality). Here we see that there has been an overall decline in the proportion of two-earner households and in the proportion of households with one male earner. On the other hand, households with one female earner and, most noticeably, no earners, have increased their proportion of the total. The increase in the share of no earner households was very pronounced between 1979 and 1991, but it continued to rise, albeit less sharply, to 2000. The proportion of one female earner households rose slightly between 1979 and 1991 (more markedly among the younger households) and then more rapidly after 1991.

Table 5. *Distribution of household types according to earners.*

	Household types			
(a) Complete sample				
	2 earners	One earner – male	One earner – female	No earners
1979	46.6	31.5	11.3	10.6
1991	37.7	29.8	12.9	19.6
2000	39.4	22.4	15.7	22.5
(b) Age of head: 25-34				
	2 earners	One earner – male	One earner – female	No earners
1979	49.2	36.0	8.4	6.4
1991	35.4	40.1	11.8	12.7
2000	39.2	26.6	16.2	17.9

Decomposition by Household Types

It is well known that during both the 1980s and 1990s earnings inequality experienced a dramatic rise in the U.K. The Theil index increased from 0.258 in 1979 to 0.454 in 1991 and 0.542 nine years later[8], an increase of 130% over two decades. Average household earnings increased steadily during the period, too. A similar pattern is evident among the younger households.

Table 6. *Average earnings and earnings inequality (Theil indexes) over time.*

	Complete sample			**Age of head: 25-34**		
	1979	1991	2000	1979	1991	2000
Average earnings	134.39	179.28	249.37	148.58	202.90	271.10
Theil	0.258	0.453	0.542	0.197	0.313	0.420
Between group	0.152	0.275	0.315	0.109	0.183	0.267
Within group	0.107	0.178	0.227	0.089	0.131	0.153
N	4964	5113	5243	1385	1334	1244

Table 6 also shows that not only is the between-group inequality larger than the within-group, but the former has increased more in absolute terms, and this is particularly noticeable among the younger households: in other words, inequalities between household types grew faster than did inequalities within them. The counterfactuals and decompositions reported in Table 7 show that the increasing between-group inequality was wholly due to the changing distribution of household types, while the growth in within-group inequality was caused by a growth in the household-type specific Theil values. Table 7 reports the observed inequality indexes, then a set of counterfactuals in which each of p, $\bar{x}$ and T, and all combinations of them, are allowed to take their t+1 values. The parallel decompositions using the Mookherjee-Shorrocks method are then shown in the final three rows, where changes in p

[8] The total Theil index in 1979 is equivalent to a Gini coefficient of 0.367; for 1991, it is equivalent to a Gini of 0.486 and for 2000 the correspondent Gini coefficient is 0.523.

affect both the between- and within-group inequality. Both the counterfactuals and the decompositions point to the same conclusions: particularly, for the whole sample, the change in inequality between household types is almost entirely driven by the change in the distribution of household types. For example, allowing only p to change in all cases brings the between-group inequality very close to its observed t+1 value. Likewise, changing the Theil values for each household type alone almost exactly reproduces the within-group inequality at t+1. Changes in the mean earnings in each type of household have little effect on either the between- or within-group inequality (not least because all household types experienced growth in their earnings). Similarly, the decompositions show that changes in p have a substantial effect on between-group inequality and, indeed, that changes in p are the single largest cause of increasing inequality.

The mechanisms that we earlier listed as possible means by which increased female educational attainment (namely, increased female labour force participation, increased educational homogamy and changes in the distribution of single and dual earner households) will all influence between-household type, rather than within-household type, inequality.[9] Thus, our concern is primarily to explain this aspect of inequality. But now we see that the change in between-household type inequality is indeed driven by the changing distribution of household types, rather than the change in their relative mean earnings. Therefore, when we carry out our Deming-Stephan counterfactuals we allow only the distribution of p to change and we keep mean earnings and the within-group Theils fixed at their t values.

[9] This is not wholly true for changes in women's participation. Certainly if more of those women who would not have worked had they had lower educational levels are now working, this will cause the share of two earner households to increase (and those of either or both of one or no earner households will decline), but if increased education leads women who, counterfactually, would have worked, to increase their hours of work, this will influence within-household type inequality, rather than the distribution of household types. We assume that this effect of increased education is likely to be small enough for us to ignore it.

Table 7. *Decomposing changes in earnings inequality (Theil index).*

Period:		1979-1991			1991-2000		
		Theil	Between-group	Within-group	Theil	Between-group	Within-group
(a) Complete sample							
Observed	t	.258	.152	.107	.453	.275	.178
	t+1	.453	.275	.178	.542	.315	.227
Change in :	p_j	.373	.264	.109	.486	.313	.173
	$\bar{x}_j$	.269	.162	.107	.455	.279	.176
	T_j	.320	.152	.168	.502	.275	.227
	p_j and $\bar{x}_j$	.384	.275	.109	.488	.315	.173
	p_j and T_j	.440	.264	.175	.532	.313	.219
	$\bar{x}_j$ and T_j	.334	.162	.172	.514	.279	.235
	p_j , $\bar{x}_j$, T_j	.453	.275	.178	.542	.315	.227
Shorrocks	$\Delta(t,t+1)$	.195	.123	.071	.089	.040	.049
Effect of changing:	p_j	.109	.106	.003	.066	.068	-.002
	φ_j	.017	.017	-	-.028	-.028	-
	γ_j	.068	-	.068	.051	-	.051
(b) Age of head: 25-34							
Observed	t	.197	.109	.089	.313	.183	.131
	t+1	.313	.183	.131	.420	.267	.153
Change in :	p_j	.279	.184	.095	.384	.254	.130
	$\bar{x}_j$	.230	.140	.090	.327	.198	.129
	T_j	.225	.109	.116	.337	.183	.154
	p_j and $\bar{x}_j$	.277	.183	.094	.394	.267	.128
	p_j and T_j	.313	.184	.129	.409	.254	.155
	$\bar{x}_j$ and T_j	.239	.140	.119	.350	.200	.152
	p_j , $\bar{x}_j$, T_j	.313	.183	.131	.420	.267	.153
Shorrocks	$\Delta(t,t+1)$	.116	.074	.042	.106	.084	.023
Effect of changing:	p_j	.167	.149	.018	.072	.072	.000
	φ_j	-.075	-.075	-	.011	.011	-
	γ_j	.024	-	.024	.022	-	.022

Table 8 contains the results of the Deming-Stephan counterfactuals applied to our data, and they show very clearly that the changing distribution of women's education (counterfactual 1) has little or no effect on the trend towards increased earnings inequality. Furthermore, the direction of its impact has been, if anything, in the direction of ameliorating the growth in inequality. When we allow for the changing distribution of men's education and the association between men's and women's education (counterfactual 2), this also has no positive effect on increasing inequality: in other words, any trends in educational homogamy have not tended to increase inequality: indeed, they too seem to have had a slight tendency to counteract the growth in inequality. When we allow for the change in women's labour force participation (in counterfactual 3) we see a slightly larger effect on increasing inequality, but this effect is much smaller than that observed in our fourth counterfactual, where we allow men's labour force participation to change. The latter seems to have been the major factor underlying the change in between-household type earnings inequality in both the 1979-1991 and 1991- 2000 period, and this is true for the whole sample and for households headed by someone aged 25-34. Even if we take the most generous definition of the impact of changing women's education on inequality (i.e. counterfactual 3, which allows for the effects of increasing education on the distribution of household types according to partnership status and on women's employment) we find that, according to the results shown in Table 8, it accounted for about three and a half per cent of the increase in between-household type inequality in the 1979-1991 period among the complete sample. In the other three cases the net effect was to reduce inequality.

This result is robust to the order in which these counterfactuals are performed. For example, if we consider the 1979-1991 change in inequality among the whole sample, then changing first men's labour force participation increases the between-group inequality from its observed value of .152 to .270. Adding women's labour force participation leaves it unchanged; adding men's education reduces it to .251 while adding women's education takes it to .264 (compared with an observed value of .275). Similarly, in the 1991-2000 period, changing first whether or not men are working alters the between group inequality from its observed 1991 value to .333. Subsequent additions of whether or not women are working (.320), men's education (.320) and women's education (.313) all have very little further impact.

Table 8. *Deming-Stephan decomposition of the changing household distribution's effects on earnings inequality (Theil index).*

(a) Complete sample						
Period:		1979-91			1991-2000	
		Theil	Between-group	Within-group	Theil	Between-group	Within-group
Observed values	t	.258	.152	.107	.453	.275	.178
	t+1	.453	.275	.178	.542	.315	.227
Change in:							
1. Women's education		.255	.148	.107	.443	.272	.172
2. 1+ men's education		.249	.144	.106	.446	.268	.178
3. 2 + women's labour force participation		.263	.156	.108	.425	.250	.175
4. 3 + men's labour force participation		.373	.264	.109	.486	.313	.173
(b) Age of head: 25-34							
Period:		1979-91			1991-2000	
		Theil	Between-group	Within-group	Theil	Between-group	Within-group
Observed values	t	.197	.109	.089	.313	.183	.131
	t+1	.313	.183	.131	.420	.267	.153
Change in:							
1. Women's education		.184	.100	.084	.311	.183	.128
2. 1+ men's education		.177	.083	.094	.311	.182	.129
3. 2 + women's labour force participation		.181	.087	.093	.288	.160	.128
4. 3 + men's labour force participation		.279	.184	.095	.384	.254	.130

The reason for the ameliorating effect of changes in the distribution of women's education is quite straightforward. The greatest inequality in earnings occurs in those households which contain men and women with the lowest level of education, while the least inequality is found among households at the middle educational level. This is in part because there are more men and women with low education (as opposed to medium and high education) in households with zero earnings, and because, in female earner households, inequality tends to be higher when education is low. So, on the one hand, changing the distribution of women's education increases the share of households containing women having the middle level of education and reduces the share containing women with the lowest level of education. For example, in 1979, 56 per cent of households contained a woman with the lowest level of education, and in such households the average Theil value was .27, while 28 per cent of households contained a woman with the middle level of education, and the

average Theil here was .19. Changing the educational distribution of women to its 1991 values put 36 per cent of households in the former category and 43 per cent in the latter: this is an obvious equality-enhancing change. Similarly, when we add the change in men's education and in the distribution of households types according to education, this reduces the share of households containing low educated men and women (and among which inequality is high) and increases the share of other households (such as homogamous couples at the middle educational level) among which inequality is lower.

On the other hand, the large unequalising effect associated with changes in the distribution of whether or not men are working arises for the following reason. As Table 5 shows, increases in the proportion of households with no earners have largely occurred at the expense of the share of households with one male earner, though households in which there is a male earner make up the majority of households in all three years (they constitute 78 per cent in 1979, 68 per cent in 1991 and 62 per cent in 2000). But this means that, in the Deming-Stephan counterfactuals in which we allow the distribution of households to gradually shift from its t to its t+1 distribution, the share of households with zero earners will remain quite near its t value until we have allowed for the change in the proportion of men who are working.

Explaining within-group inequality

We earlier noted the increasing correlation between the earnings of men and women in dual earner households: this might be considered an effect of increasing women's educational attainment if it were caused by a shift in the joint distribution of the educational level of each member of the couple. But this would be captured in our analysis as a change in the inequality between household types, and, as we have seen, changes in between-household type inequality are not driven by changes in the educational distribution of couples. In fact (and as we show below), the correlation between partners' earnings has changed in almost all combination of man's and woman's education. This will affect inequality within each type of household, rather than inequality between them.

To investigate changes in inequality within household types we turn to Table 9, panel A of which reports the Theil value for all single-earner households in each year. This shows that, in the complete sample, households with a female earner had greater inequality, in both 1979 and 1991, than their male earner counterparts; and, while there has been a general tendency for within household type inequality to grow, this growth has been more pronounced among male earner households, with the consequence that, by 2000, they were more unequal than their female headed equivalents, except among those with the lowest level of education. This is despite the decline in inequality in households with poorly educated female earners. Inequality in households where the woman had the highest level of education experienced an inverted u-trend, increasing between 1979 and 1991 and decreasing subsequently. This particular household type was the most equal (among households with a female earner) over the three points in time. Panel A also presents the same within-group inequalities for the youngest sub-sample of households. Inequality increased for all male earner households, although to a more limited degree than for the whole sample. The pattern for young female earners follows roughly that of the complete sample: the sub-group of low educated women is extremely unequal at the end of the seventies but becomes more homogeneous afterwards; inequality grew between 1979 and 1991 for women with medium education and fell slightly in 2000; finally, the same inverted u-trend that we observed for all ages female-earner households can be seen here, although in all three points in time young households with this educational level are more equal than the complete sample.

Panel B of Table 9 shows the Theil values, standard deviations of men's and women's earnings, and the correlation between them in households where there is a partner and where both the head and the partner are working (i.e. dual-earner households).[10] Among the complete sample, educationally homogamous couples (at all levels) showed similar, and not particularly high, levels of inequality in 1979, but inequality increased among all of them, though the increase was largest in households containing couples with the highest level of education. The Theil value also increased, between 1979 and 2000, in households in which the partners had, respectively, medium and high education, and in households where the man had low education but the woman had a medium level of education. The correlation between

[10] The standard deviations and correlation refer to unadjusted earnings.

Table 9: *Explaining changes in within-household type inequality.*

Panel A: Theil values in single earner households.

	Complete sample			**Age of head: 25-34**		
	Educational level of male earners					
Year	Low	Medium	High	Low	Medium	High
1979	0.111	0.122	0.117	0.078	0.122	0.098
1991	0.186	0.225	0.222	0.192	0.139	0.185
2000	0.211	0.454	0.338	0.194	0.216	0.261
	Educational level of female earners					
Year	Low	Medium	High	Low	Medium	High
1979	0.366	0.262	0.164	0.606	0.233	0.145
1991	0.329	0.301	0.232	0.510	0.327	0.181
2000	0.321	0.321	0.209	0.303	0.311	0.152

Panel B: Theil, standard deviations of men's and women's earnings, and correlation between them.

Complete sample

	Men´s Education											
Women's	Low				Medium				High			
Education	T	s_m	s_w	r	T	s_m	s_w	r	T	s_m	s_w	r
Low												
1979	0.082	74.2	56.1	0.04	0.092	96	59.5	0.08	0.108	164.6	75.1	-0.27
1991	0.087	116.4	72.2	0.10	0.397	684.8	88.7	-0.01	0.252	287.7	132.8	0.07
2000	0.154	284.5	112.5	0.03	0.134	181.9	147.8	0.09	0.078	188.9	118.9	0.22
Medium												
1979	0.085	96.5	65.4	-0.07	0.089	108.6	74.1	0.05	0.105	158.3	73.7	-0.08
1991	0.118	151.8	94.7	0.09	0.116	186.6	118.4	0.12	0.090	219.8	120.7	-0.08
2000	0.125	203.3	150.2	0.16	0.138	231.2	157.6	0.21	0.222	567.4	144.7	0.18
High												
1979	0.056	89.6	74.1	-0.24	0.065	130.8	93.1	0.11	0.089	122.5	89.3	0.13
1991	0.126	195.2	123.1	-0.08	0.101	191.3	149	-0.02	0.171	369.5	183.0	0.12
2000	0.072	185.5	165.2	-0.18	0.183	478	199.8	0.10	0.170	531.2	341.8	0.43

Age of head: 25-34

	Men's Education											
Women's	Low				Medium				High			
Education	T	s_m	s_w	r	T	s_m	s_w	r	T	s_m	s_w	r
Low												
1979	0.086	79.3	52.2	0.05	0.072	83.4	58.6	0.05	*	*	*	*
1991	0.054	75.4	48.4	0.33	0.109	98.1	84.6	0.13	*	*	*	*
2000	0.117	146.3	83.4	0.33	0.065	112.8	91.5	0.02	*	*	*	*
Medium												
1979	0.087	70	69	0.06	0.092	94.8	72.1	0.05	0.057	147	65.4	-0.47
1991	0.115	105.7	84.9	0.19	0.095	116.7	94.8	0.13	0.084	152	117.8	0.09
2000	0.053	153	97.6	-0.18	0.103	148	127.6	0.21	0.097	289.8	131.7	0.10
High												
1979	**	**	**	**	0.049	93.9	78.5	-0.34	0.084	115.9	81.4	0.20
1991	*	*	*	*	0.102	140.9	140.7	0.29	0.104	209	141.5	0.20
2000	*	*	*	*	0.218	677.5	219.9	0.06	0.107	304.7	196.7	0.32

Note:
T = Theil value
s_m = standard deviation of men's earnings
s_w = standard deviation of women's earnings
r = correlation, men's and women's earnings
* 5 or less observations
** 9 observations

the couple's earnings increased in homogamous households with medium and high levels of education (and in all households that contained a highly educated man), but the correlation declined in low-low households. The variation in earnings is generally larger the higher the level of education (as we might have expected given that the level of earnings varies with education in a similar way), and this is true of both sexes, but it is also the case that the standard deviation of one partner's earnings varies according to his or her partner's level of education.

But such increasing correlations do not automatically lead to more inequality. The correlation in earnings in households in which both partners had high education grew a lot between 1991 and 2000 (from 0.12 to 0.43) but their Theil value remained unchanged. Furthermore, the correlation is higher here than anywhere else, but the Theil is not particularly big. So while an increasing correlation will tend to cause more inequality, there have been offsetting factors that mean that, empirically, we do not always see a positive association between the correlation and the Theil index.

For the youngest households it is hard to establish a clear picture because of the small sample sizes, but it is certainly the case that inequality tends to be lower in these young households than in the complete sample (which fits with the results shown in Table 6), and, unsurprisingly, the standard deviations of earnings are lower. As in the complete sample, the Theil index is greater in 2000 than in 1979 in all educationally homogamous households and in all combinations of medium or highly educated men married to medium or highly educated women. The correlation in earnings has increased in all educationally homogamous households.

Dual earner households

Much attention has been focused on dual earners as a likely source of increasing earnings inequality, which might be induced not only by greater educational or earnings homogamy, but also by the increased importance of women's earnings. Table 10 shows the

results of a Shorrocks decomposition of within-group household earnings inequality in dual-earner households by factor component (Shorrocks 1982) where factors, for our purposes, are the relative contributions of earnings coming from men and women.[11] Inequality in this type of household is largely a function of men's earnings; for the complete sample, women's earnings only account for a third (at most) of total inequality and their contribution does not follow a single pattern over time (it decreases and then increases again). In the youngest households, women's contribution to inequality tends to be higher: in 1979 and 1991 it accounted for around 40% of the total variance, although by 2000 their relative importance declined and became similar to that observed in the complete sample of households. This picture is consistent with two of the facts that we learned from panel B of Table 9: a) that the distribution of men's earnings became significantly more unequal over the two decades and b) that although women's earnings inequality also experienced an important increase, the actual level of inequality remained systematically lower than men's.

Table 10: *Men's and women's contributions to household earnings inequality (in percentages) in dual-earner households.*

	Complete sample	
	Men's contribution	Women's contribution
1979	66.28	33.72
1991	81.45	18.55
2000	70.21	29.79
	Age of head: 25-34	
	Men's contribution	Women's contribution
1979	60.83	39.17
1991	56.08	43.92
2000	72.89	27.11

Note: These calculations refer to unadjusted earnings

[11] Shorrocks (1982) shows that, for a large class of inequality measures of which the Theil index is one, the relative contribution to inequality of each factor, say X, of which the measure on which inequality is computed - say Y - is the sum, can be expressed as the covariance between X and Y divided by the variance of Y.

Conclusions

We suggested three mechanisms by which increasing levels of education attainment among women might lead, *ceteris paribus*, to a growth in inequality between households in their earnings: these are changes in female labour force participation; increasing educational homogamy, and thus an overall increase in the correlation of partners' earnings; and a change in household formation behaviour, with more single earner households and more households with fewer or no dependent children. These three mechanisms have been mentioned by previous authors, though they have not been associated with what we believe to be one of their main causes – namely the growth in women's educational attainment.

Our results show that increasing earnings inequality in Britain between 1979 and 1991 was mainly due to growing inequality between types of household defined according to the educational level and employment status of the head and his/her partner (if any), while, between 1991 and 2000, the much lesser increase was about equally due to between-group and within-group change.[12] The three mechanisms we outlined would all affect between-group inequality, mainly by changing the distribution of household types in the population, and our initial counterfactuals and our decomposition using the Mookherjee and Shorrocks (1982) method showed that all the between-group change could indeed be attributed to change in the distribution of household types, while all of the within group change was due to change in the Theil values in each type of household. But our second set of counterfactuals, using the Deming-Stephan algorithm, showed that changes in the distribution of women's education played little, if any, role in accounting for changes in the inequality between household types, which was overwhelmingly the result, in both periods, of the changing proportion of male household heads (or male partners of household heads) who were not working.

Although it is true that the correlation between the earnings of heads of households and their partner has increased over the period (and very sharply between 1991 and 2000), this is not due to increased educational homogamy, because the growth in the correlation has

[12] The changes among our sample of young households were somewhat different, but here we concentrate on the results from the complete sample.

occurred in many types of two-earner households, and has been particularly pronounced among couples, both of whom have high levels of education. This indicates that there is some process of assortative mixing, beyond educational homogamy, occurring. One possibility is that our measure of education is neither reliable nor fine enough to capture the processes of educational homogamy. On the one hand, the categorization of years of education is likely to be a poor proxy for the educational qualifications that are important in helping people to get jobs and in determining their earnings; on the other, broad qualifications themselves may not discriminate fully in this respect. So, for example, considerations of which subject someone has obtained a degree in, and where they got it from, will undoubtedly also play a role in shaping earnings, and it may be that marriage partners meet within specific educational settings (such as a particular University or a particular course) or that being in such a setting helps to increase their likelihood of subsequently meeting and establishing a household with someone who comes from the same or a similar setting. One way in which this might occur is if a certain employer (or a large share of the employers in a particular business) recruits alumni from the same, or the same kind, of educational establishment or course. Another possibility is that, as the age of forming a stable partnership increases, partners may increasingly be found from among those with similar earnings, or earnings potential (if, for instance, this leads them to share similar lifestyles), and this may operate irrespective of, or, more likely, in addition to, a commonality in educational attainment. In other words, this would provide for increasing homogamy within those with the same level of education. Yet a third possibility is assortative mating on the basis of unmeasured characteristics which might correlate with earnings. These explanations are not mutually exclusive, but which, if any, of them are operating could only be answered if we knew how people meet the partners with whom they form households: to the best of our knowledge, this is a notably under-researched area. But we end by reiterating our main result: given the data at our disposal we find that conjectures about the unequalizing effects of increasing women's educational attainment and labour force participation are not born out in the UK over the period of rapidly growing inequality during the last 20 years of the 20th century.

REFERENCES

Alderson, A.S. and F. Nielsen 2002. "Globalization and the Great U-Turn: Income Inequality Trends in 16 OECD Countries". *American Journal of Sociology* 107: 1244-99.

Atkinson, A.B. 1975. *The Economics of Inequality*. Clarendon Press. Oxford.

Atkinson, A.B., L. Rainwater and T.M. Smeeding. 1995. "Income Distribution in OECD Countries: Evidence from the Luxembourg Income Study". *OECD Social Policy Studies* 18. Paris

Atkinson, A.B. 1997. "Bringing Income Distribution in from the Cold". *Economic Journal* 107(441): 297-321.

Blossfeld, H.-P. and A. Timm (eds.). 2003. *Educational Systems as Marriage Markets in Modern Societies. A Comparison of Thirteen Countries*. Kluwer Academic Publishers. Dordrecht.

Bourguignon, F. 1979. "Decomposable Income Inequality Measures". *Econometrica* 47(4): 901-20.

Burtless, G. 1999. "Effects of Growing Wage Disparities and Changing Family Composition on the U.S. Income Distribution". *European Economic Review* 43: 853-65.

Callan, T., Nolan, B., O'Neill, D. and O. Sweetman. 1998. "Female Supply and Income Inequality in Ireland". *Economics Department Working Papers Series*. NUI Maynooth.

Cowell, F.A. and S.P Jenkins. 1995. "How much inequality can we explain? A methodology and its application to the United States". *The Economic Journal* 105: 421-30.

Deming, W.E. and F.F. Stephan. 1940a. "On a Least Squares Adjustment of a Sampled Frequency Table when the Expected Marginal Total are Known". *The Annals of Mathematical Statistics* 11(4): 427-44.

Deming, W.E. and F.F. Stephan. 1940b. "The sampling procedure of the 1940 population census". *Journal of the American Statistical Association* 35: 615-30.

Drobnič, S. and H.-P. Blossfeld 2001. "Careers of Couples and Trends in Inequality", pp. 371-86 in Blossfeld, H.-P. and S. Drobnič (eds.) *Careers of Couples in Contemporary Society: From Male Breadwinner to Dual-Earner Families*. Oxford University Press. Oxford.

Firebaugh, G. 2003. *The New Geography of Global Income Inequality*. Harvard University Press. Harvard.

Green, A., A. Wolf and T. Leney. 1999. *Convergence and Divergence in European Education and Training Systems*. Institute of Education. University of London. London.

Hedström, P. and R. Swedberg. 1998. *Social Mechanism. An Analytical Approach to Social Theory*. Cambridge University Press. Cambridge, U.K.

Jenkins, S.P. 1995. "Accounting for Inequality Trends: Decomposition Analyses for the UK, 1971-86". *Economica* 62: 29-63.

Mookherjee, D. and A.F. Shorrocks 1982. "A Decomposition Analysis of the Trend in UK Income Inequality". *The Economic Journal* 92: 886-902.

Shavit, Y. and H.-P. Blossfeld (eds.). 1993. *Persisting Inequality: Changing Educational Attainment in Thirteen Countries*. Westview Press. Boulder, Co.

Shorrocks, A.F. 1982. "Inequality Decomposition by Factor Components". *Econometrica* 50 (1): 193-211.

Shorrocks, A.F. 1983. "The Impact of Income Components on the Distribution of Family Incomes". *Quarterly Journal of Economics* 98 (2): 311-26.

Shorrocks, A.F. 1984. "Inequality Decomposition by Population Subgroups". *Econometrica* 52 (6): 1369-85.

APPENDIX 1

Compete Sample	*1979*	*1991*	*2000*
Percentage of households in which `others´ work:			
All	21	21	15
Households in which neither head nor partner work	19	13	9
Male earner households	17	27	13
Female earner households	30	19	16
Dual earner households	22	21	18.5
Mean earnings coming from `others´:			
All	72	186	217
Households in which neither head nor partner work	76	179	199
Male earner households	77	204	261
Female earner households	73	209	211
Dual earner households	69	162	207
Variance in earnings coming from `others´:			
All	2238.5	23098.2	36046.7
Households in which neither head nor partner work	2329.2	21837.8	17158.7
Male earner households	2355.1	28660.9	38981.7
Female earner households	1961.4	23165.8	31458.9
Dual earner households	2239.9	16751.1	41080.0

Complete Sample	***Definition of Household Earnings***					
	1979		1991		2000	
	Def. 1	Def. 2	Def. 1	Def. 2	Def. 1	Def. 2
Total average earnings	134	150	179	199	249	265
Total Theil	0.257	0.217	0.453	0.403	0.542	0.500
Households in which neither head nor partner work						
Mean households earnings	0	15	0	13	0	9.5
Variance	0	1469.4	0	1997.4	0	1422.4
Theil	0	1.854	0	2.335	0	2.646
Male earner households						
Mean households earnings	140	153.5	212	243.5	323	342
Variance	5838.1	6361	30278.3	33583.5	239792	244461.9
Theil	0.129	0.120	0.244	0.221	0.410	0.382
Female earner households						
Mean households earnings	93	119	140	160	212	231
Variance	6746.6	7630.8	13982.5	14770.9	31139.1	31616.9
Theil	0.356	0.271	0.338	0.289	0.327	0.288
Dual earner households						
Mean households earnings	171	185	260	275	364.5	382
Variance	6215.9	6120.9	41001.3	40416	76426.3	77109.7
Theil	0.098	0.084	0.173	0.155	0.185	0.169

Note: Definition of household earnings:
Definition 1 = Earnings from head of the household and partner (if present)
Definition 2 = Earnings from head of the household, partner (if present) and others (if present)

Age of head: 25-34	*1979*	*1991*	*2000*
Percentage of households in which `others´ work:			
All	4	12	4
Households in which neither head nor partner work	3	9	1
Male earner households	6	23	7
Female earner households	9	5	5
Dual earner households	2	1	2
Mean earnings coming from `others´:			
All	74	230	304.5
Households in which neither head nor partner work	103.5	294	154
Male earner households	73	224	332
Female earner households	72	206	352
Dual earner households	71	218	220
Variance in earnings coming from `others´:			
All	1477.1	30753.7	56516.4
Households in which neither head nor partner work	312.8	57618.8	894.2
Male earner households	1259.1	29520.7	74595.3
Female earner households	1425.5	21216.9	36580.8
Dual earner households	2229.6	3643.1	41830

Age of head: 25-34	***Definition of Household Earnings***					
	1979		1991		2000	
	Def. 1	Def. 2	Def. 1	Def. 2	Def. 1	Def. 2
Total average earnings	149	152	203	219	271	277.5
Total Theil	0.197	0.192	0.313	0.298	0.420	0.412
Households in which neither head nor partner work						
Mean households earnings	0	5	0	17	0	1
Variance	0	641.8	0	5303.9	0	42.9
Theil	0	3.403	0	2.758	0	4.739
Male earner households						
Mean households earnings	142	147	217	250	299	312
Variance	5509.7	6104.8	19736.3	24064	58350.8	58931.4
Theil	0.116	0.120	0.180	0.177	0.254	0.241
Female earner households						
Mean households earnings	104.5	113	173	180	242	254
Variance	6874.1	7160.9	18009.7	18690.3	33450.8	35757.8
Theil	0.352	0.324	0.315	0.311	0.291	0.287
Dual earner households						
Mean households earnings	180	181	269.5	271	388	390
Variance	6832.6	6768.7	19248	19100.4	68585.7	68973.7
Theil	0.098	0.096	0.123	0.121	0.159	0.159

Note: Definition of household earnings:
Definition 1 = Earnings from head of the household and partner (if present)
Definition 2 = Earnings from head of the household, partner (if present) and others (if present)

APPENDIX 2

The data for our counterfactuals comprise a 4-way table of woman's education (W_E) by men's education (M_E) by whether or not the woman works (W_W) by whether or not the man works (M_W). The last two of these are dichotomies, but the education variables each have four categories: low, medium and high education plus, because our units of observation are households, a category of 'not present' for those households where the head of the household does not have a partner. For the same reason, the 4-way table is incomplete because -16 out of the possible 64 combinations of the variables can never be observed: these are known as 'structural zeroes'.

To compute the counterfactuals we use the Deming-Stephan algorithm (sometimes called Iterative Proportional Fitting), which provides a means of adjusting the marginal distributions of a contingency table while preserving the pattern of associations, as captured by odds ratios, among those variables (Deming and Stephan 1940a, b).[13]

Let f_{ij} be the frequencies of a contingency table with rows i=1,…,I and columns j=1,..,J and define $f_{i+\equiv}\sum_j f_{ij}$ and $f_{+j} \equiv \sum_i f_{ij}$ to represent the row and column totals of the table. Given target row and column distributions, f^*_{i+} and f^*_{+j}, the Deming-Stephan algorithm adjusts the observed frequencies by a series of iterations, each of which has two steps, as follows:

$$(1)\quad f_{ij}^{11} = f_{ij}\frac{f^*_{i+}}{f_{i+}}, \quad f_{ij}^{12} = f_{ij}^{11}\frac{f^*_{+j}}{f_{+j}^{11}}$$

$$(2)\quad f_{ij}^{21} = f_{ij}^{12}\frac{f^*_{i+}}{f_{i+}^{12}}, \quad f_{ij}^{22} = f_{ij}^{21}\frac{f^*_{+j}}{f_{+j}^{21}},$$

and so on until convergence to the adjusted frequencies.

The Deming-Stephan algorithm is very easily implemented on a 2-dimensional table and so we carried out our estimations by reshaping the 4-way contingency table as the appropriate 2-way table. In *counterfactual 1* we allow the marginal distribution of women's education to change. So we reshaped the 4-way E_W by M_E by W_W by M_W (4 by 4 by 2 by 2) table into a 2-way E_W by X1 table, where X1 is the 16 category variable capturing all the E_M by W_W by W_M combinations. This 4 by 16 table thus includes the 16 structural zeroes, but this is not a problem because the Deming-Stephan procedure, of necessity, preserves any zero cell values. We thus make counterfactual estimates of the change between t and t+1 by taking the E_W by X_1 table observed at t and applying the algorithm using the observed X1 (column) marginal totals and the counterfactual E_W (row) margin, which is given by the E_W distribution at t+1. Because the sample sizes vary between t and t+1 we use the t sample size and thus the

[13] Odds ratios express the relative chances of an observation being found in category j rather than in j' of one variable, conditional on being located in category i rather than i' of another variable.

counterfactual row totals are given by the t+1 row proportions applied to the t row total. But in this particular counterfactual our interest is in the effect of a change in the distribution of women's education per se and not in any changes in the distribution of households types that does not follow directly from this. We therefore set the marginal total for the fourth category of E_W ('not present') to its value at t, and the counterfactuals for the three other categories were adjusted accordingly.

Counterfactual 2 involved reshaping the 4-way table into a 2-way X_2 by X_3 table, where X_2 captures all the combinations of E_W and E_M while X_3 captures W_W by W_M: thus the table has 16 rows and 4 columns. The row variable measures both the educational distributions of men and women and the association between them (including the share of households without a man or without a woman). As before, the t+1 row distribution is used to form the counterfactual row distribution while the column distribution is left at its observed value. In *counterfactual 3* the row variable is then expanded to 32 categories with the addition of the distinction between households in which the woman is working and those in which she is not.

The reshaping of the table and the Deming-Stephan algorithm itself are easily carried out on any computer program that allows the user to write macros.