Online Appendix for "The UK as a Technological Follower: Higher Education Expansion, Technological Adoption, and the Labour Market"

Richard Blundell, David A. Green and Wenchao Jin*

May 13, 2021

1 Alternative Production Function

A referee pointed out that a natural alternative production function given our main model could be written as:

$$Y = Y^C + Y^D \tag{1}$$

With,

$$Y^i = \theta_i F^i(U^i, S^i, K^i), i = C, D \tag{2}$$

*Blundell: University College London and Institute for Fiscal Studies, Gower Street, London, WC1E 6BT, UK, r.blundell@ucl.ac.uk. Green: Vancouver School of Economics, UBC, and Institute for Fiscal Studies, 6000 Iona Dr, Vancouver, BC V6T 1L4, Canada, david.green@ubc.ca. Jin: University College London, Gower Street, London, WC1E 6BT, UK, w.jin.11@ucl.ac.uk. We would like to thank the editor, referees and seminar participants at UBC and IFS for very helpful comments on earlier drafts. This research is part of the research programme of the ESRC Centre for the Microeconomic Analysis of Public Policy at IFS. We are grateful to the ESRC, grant number RES-000-23-1524, for financial support. The usual disclaimer applies.

where, Y is total output in the economy and Y^C and Y^D are output from firms using the centralized and decentralize technologies, respectively. Y is the sum of these two because they are alternative general purpose technologies (GPT's) producing the same output. Note that technological change, as captured in the θ's is technology specific but factor neutral within each technology. As a result,

$$\ln TFP_t = s_t^C \ln \theta_{Ct} + s_t^D \ln \theta_{Dt} \tag{3}$$

i.e., it is a weighted average of the technology specific technological change factors, with the weights being the share of total income accounted for by the centralized, s_t^C, and decentralized, s_t^C, technologies. Given that we assume these are GPT's, there is no direct measure for these shares. However, in the theory, they will be directly related to the proportion of skilled and unskilled workers in the economy, so an empirically implementable version of **??** would use the skilled and unskilled labour income shares. That is, one arrives at a formulation that is the same as the one we implement. Given that we cannot identify one of these formulations from the other, we adopt the much more common specification in which technological change is skill enhancing.

2 The Role of Capital and Constrained System Estimates

In this appendix, we discuss the role of capital and TFP in wage specifications and present results when we impose theoretically implied cross-equation restrictions. The regression equations (4) and (5) in the paper allow for general productivity growth but also incorporate a flexible skill biased technological trend. Many of the specifications estimated in the micro labour literature on technological change do not include either capital or TFP, and our specification obviously nests such an approach. In particular, if $\alpha_2 = \beta_2$ then neither

TFP_t nor K_t appear in the relative wage equation. This would imply that capital is equally complementary with skilled and unskilled labour and would occur, for example, if the production function were multiplicatively separable in K_t and the overall labour component. That is what was assumed in the seminal **?** paper and is one explanation for why most of the Skill Biased Technical Change (SBTC) literature and the polarization literature that followed use specifications that do not include capital. An alternative explanation not including capital comes from the combination of the constant returns to scale assumption and an assumption of a perfectly elastic supply of capital. It is straightforward to derive an expression for the price of capital and use it to substitute out the $\ln(\frac{K_t}{U_t})$ term in our two estimating equations. If we assume that the world price of capital is constant then here, as in the case with multiplicatively separable capital, we end up with the canonical specification for the relative wage equation with only a time trend and the relative skill supply variables on the right hand side.[1] We can, alternatively, allow the price of capital, r_t, to vary over time, implying an adjusted version of the canonical specification that includes $\ln r_t$ as a regressor. Estimates of this adjusted specification are available upon request. That specification yields very similar results in terms of the estimates of the coefficients of interest to those reported in the text. The theory underlying our specifications implies several restrictions. Weak concavity of the production function implies that $\beta_1 - \beta_2 \leq 0$. From equation (14) in the paper, the coefficient on $\ln(\frac{S_{gt}}{U_{gt}})$ in the skilled wage regression equals $\beta_1 - \beta_2$ and the estimates of that coefficient in both our OLS and IV estimates in Table 1 are negative. Second,

[1] The two different approaches for eliminating capital from the relative wage equation have different implications for the skilled wage equation. If the production function is multiplicatively separable in capital and a labour aggregate then both TFP_t and K_t enter the skilled wage equation. If, instead, the production function is not multiplicatively separable in capital and labour but capital is perfectly elastically supplied then TFP_t but not K_t is present in the skilled wage equation. As with the relative wage equation, we can include $\ln r_t$ as an added regressor in the perfectly elastic capital supply case with a time varying price of capital.

concavity implies $\alpha_1 + \alpha_2 \geq 0$. We can construct an estimate of $\alpha_1 + \alpha_2$ as, $b_2 + b_4 - (d_2 + d_4)$, which takes on values that are slightly negative in the OLS and IV estimates (-0.11 and -0.15, respectively) but are not statistically significantly different from zero in either case. Thus, here too, we cannot reject the concavity restriction. The third concavity condition, corresponding to the determinant of the Hessian, is $(b_2 \cdot d_4 - b_4 \cdot d_2) \geq 0$. This terms takes a a value of -.16 with a standard error of 0.10. Thus, from the values estimated for all three conditions, we cannot reject the null of weak concavity of the production function.

The framework implies three equality restrictions on the regression equations (4) and (5) in the paper: $b_3 + b_4 = 1$ and $d_3 + d_4 = 0$ and $b_5 = d_5$. The first restriction is clearly rejected in the OLS case while the other two are not rejected in any specification. In the first four columns of Table **??**, we present SURE and IV estimates in which we impose these restriction and show that they make very little difference to the coefficients of interest, which are the coefficients on the skill supplies and the year effect. Overall, our parameter estimates fit well (albeit not perfectly) with the requirements imposed by our assumption that we are estimating parameters associated with a well-behaved production function. The last 2 columns of Table **??** contain IV results in which we add a cubic in time, showing that this extra flexibility does not alter our results.

3 Calibration Exercises Assessing the Applicability of SBTC Models for the UK

In this appendix, we calibrate the wage equations derived from our production function using typical elasticity values from the US literature and use that to back out skill specific productivity trends for the UK in order to see if the standard model delivers reasonable predictions about underlying movements in technology.

Table 1: Skilled Wage and Wage Ratio Regressions: UK, 1993-2016

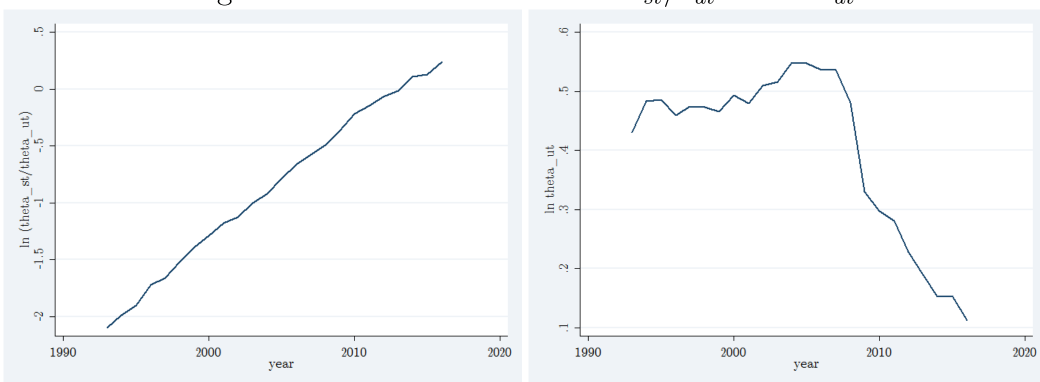
	$\ln \frac{w_{sgjt}}{w_{ugjt}}$	$\ln w_{sgjt}$	$\ln \frac{w_{sgjt}}{w_{ugjt}}$	$\ln w_{sgjt}$	$\ln \frac{w_{sgjt}}{w_{ugjt}}$	$\ln w_{sgjt}$
t	0.002	-0.006	-0.016	-0.001	-0.022*	0.006
	(0.006)	(0.006)	(0.009)	(0.010)	(0.011)	(0.010)
t^2	-0.000	0.000	-0.000	-0.000**	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\ln S_{gt}/U_{gt}$	0.056	-0.000	0.302*	-0.156	0.307*	-0.075
	(0.086)	(0.081)	(0.144)	(0.157)	(0.147)	(0.136)
$\ln \frac{TFP_t}{laborshare_t}$	0.086*	0.955***	-0.001	0.490***	0.099	0.373***
	(0.044)	(0.042)	(0.076)	(0.083)	(0.110)	(0.102)
$\ln K_t/U_t$	-0.086*	0.045	0.001	0.510***	0.104	0.145
	(0.044)	(0.042)	(0.076)	(0.083)	(0.160)	(0.149)
$\ln \tilde{S_{gjt}}/\tilde{U_{gjt}}$	0.010		0.025		0.021	
	(0.012)		(0.029)		(0.037)	
time cubic					0.000	-0.000
					(0.000)	(0.000)
$\ln \tilde{S_{gjt}}$		0.010		0.025		-0.057
		(0.012)		(0.029)		(0.133)
IVs	no	no	yes	yes	yes	yes
constraints	yes	yes	yes	yes	no	no
N	1208	1208	760	760	760	760

Notes: standard errors are shown in parentheses. The regression is at the level of 19 regions, 5-year-age-band and 3-year-period. The sample without IVs consists of 20-59 year olds. Whenever we use IVs, the sample is restricted to 20-44 year olds. The first 4 columns are the same as the first 4 columns in Table 1 in the paper except that we now impose three constrainst, and estimate using SUER and 3SLS rather than OLS and 2SLS. If we just use SURE and 3SLS and do not impose the constraints, the estimates would be very close to those in Table 1. The last 2 columns here are the 3SLS estimation of the two equations with IVs and a time cubic term; so they are the closest to the middle 2 columns in Table 1 in the paper - the only difference being the time cubic term. All specifications include complete sets of age-band dummies and region dummies. *** p<0.01, ** p<0.05, * p<0.1

To carry out this exercise, we first assume that capital is equally complementary with skilled and unskilled labour so that the capital and TFP terms drop out of equation (3) in the paper and we arrive at a specification that is the same as that used in the previous US literature. Using elasticity values common to the literature (e.g., found in **?**) of σ=1.6 and σ_a=5 and the observed trends in relative wages and relative labour supplies, we can back out an implied $\ln \theta_{st}/\theta_{ut}$ series.[2] We plot the resulting series in figure **??**, showing that it increases by more than 2 log points over the 23 years of our data. Then, given this series and observed TFP we use equation (12) in the paper to back out an implied series for θ_{ut}. That series is weakly increasing until about 2008 and then falls by more than 0.4 log points between 2008 and 2016. We view the movements of both the $\ln \theta_{st}/\theta_{ut}$ and θ_{ut} series depicted in figure **??** to be too large to be credible. We can enrich this exercise further by allowing capital to more complementary with skilled labour than unskilled labour ($\beta_2 > \alpha_2$). In our framework, $\beta 2 - \alpha 2$ is approximately the partial derivative of the log wage ratio $\ln w_s/w_u$ wrt $\ln K$, holding S, U constant. Krussell et al (2000) estimates that in the US capital equipment is more complementary with skilled labour, with $\hat{\sigma} = 0.4, \hat{\rho} = -0.5$. Their $\sigma - \rho$ roughly corresponds to our $\beta 2 - \alpha 2$. Therefore, we assume $\beta 2 - \alpha 2 = 0.9$ and back out a $\ln \theta_{st}/\theta_{ut}$ series from equation (16) in the paper. That series shows an increase of more than 5 log points over the 23 years. Again, this seems to us to be too large to be credible. Overall, the set of calibration exercises show that the basic patterns in the data, combined with standard estimated parameters from the SBTC literature yield implied skill specific productivity movements that are unrealistic. We see this as a different way of making the point that the SBTC model does not fit the UK data in our time period.

[2] We average across age bands to remove age group effects.

Figure 1: Calibrated trend of $\ln \theta_{st}/\theta_{ut}$ and $\ln \theta_{ut}$



4 The expansion of high education in the UK and education classification

The expansion of higher education over the past few decades reflects a sequence of specific policy choices made by the UK government. Since the Robbins Report in 1963, policy related to the higher education sector has been moving toward implementation of the principle that university places 'should be available to all who are qualified by ability and attainment to pursue them and who wish to do so'. The 1960s saw the foundation of more than 20 universities and dozens of polytechnics. Polytechnics were a form of higher education institution that taught both degree-level courses and below-degree-level courses, with their degrees certified by a chartered body called the Council for National Academic Awards (CNAA). A CNAA degree from a polytechnic was technically equivalent to a university degree and we treat them as equivalent in our analysis. The Education Reform Act (ERA) of 1988 changed some block grants to tuition fees (paid by Local Education Authorities for each student). In response, polytechnics increased enrolment with lower funding per student. The other major education policy change in 1988 was the replacement of CSEs and O-Levels with GCSEs as the exams that students take at age 16.[3] That reform led to an increase in educational attainment at the secondary level and hence an increase in the proportion of the young with sufficient academic credentials for potential admission to universities. In 1992, polytechnics gained the right to issue degrees and become fully-fledged universities. The reclassification of polytechnics as universities led to a jump in the number of university students in 1992; but the rapid increase in student numbers in higher education started in 1988 and continued until 1994.[4] In 1994, pressures on public expenditures and a desire to protect resources per student led the government to introduce the maximum student number control. This limited the number of full-time undergraduates at individual universities per year. As a result, the growth in student numbers slowed. This acceleration and then deceleration can be seen clearly in the BA proportion across birth cohorts in Figure **??** in Appendix A.

This paper has focused on the comparison between two education gruops: BA and HS. Here we show our main result that the BA-HS wage differential has been flat is robust to alternative definitions of education groups. In the paper, we have defined BAs as those whose highest qualification is first degree or higher, and HS as those who obtained Grade C or higher in the General Certificate of Secondary Education exam (GCSE) or equivalent and who did not have any degree-level qualification. We chose these definitions so as to be broadly comparable to college graduates and High School graduates in the US.[5]

The first alternative we investigate is to draw the bottom line of the HS group at A-levels rather than GCSEs. A-levels are subject-based exams taken

[3] Certificate of Secondary Education (CSE) and General Certificate of Education Ordinary Level (O-levels) were subject-based qualifications that students in England at the end of secondary school around age 16. CSEs are less academic, and so we count O-Levels in our definition of HS group (equivalent to GCSEs grade C or above), but not CSEs. CSEs are considered equivalent to GCSE below grade C.

[4] This has been clearly shown in Figure 2 in **?**

[5] For example, among 25-29 year olds in 2012, the US proportion of "BA" and "HS" are 35% and 56%. In the UK, the proportions according to our definition are 36% and 53%.

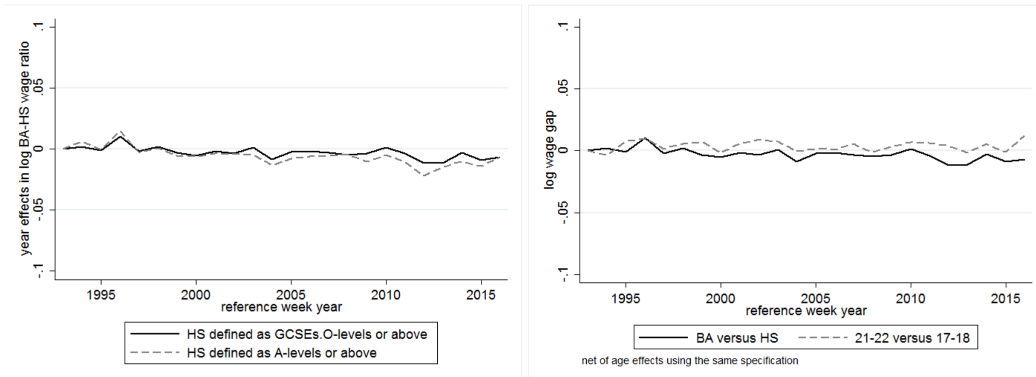


Figure 2: BA-HS log wage gap under alternative definitions

Notes: Same specification as Figure 2 in the paper. The solid line in each graph is identical to the one in Figure 2.

typically at age 18 and are a pre-requisite for university admission. Under the UNESCO's International Standard Classification of Education (ISCED 2011), both GCSEs and A-levels in the UK are classified as level 3 -"upper secondary education", and so are High School Diploma in the US. The left subgraph in Figure **??** shows that drawing the line at A-levels instead of GCSEs makes very little difference to the trend in the BA-HS wage gap.

Second, we group people by the age they left full-time education, and look at the wage gap between those who left at age 21-22 and those who left at 17-18. In Figure **??**, we show the estimated trend (net of age effects) alongside the one based on our main definition of education, which was shown in Figure 2 in the paper). Again, both trends are remarkably flat over the sample period. In summary, our main conclusion that the college wage premium has been flat since the early 90s is robust to how it's defined.

Finally, we want to address the concern that the strong increase in the BA proportion observed in the Labour Force Survey may have been over-estimated due to sampling and measurement issues. The LFS is not a compulsory survey and its response rate has been declining over time.[6] If graduates have a differential response rate to less-educated people, the LFS may yield a biased estimate for the overall BA proportion. As a sensibility check, we obtain the number of graduates from the Higher Education Student Statistics (HESA)[7]. HESA collects student information directly from each university since the early 90s, so the graduate numbers are precise. We use the total number of UK-domicile students obtaining first degrees every academic year. This is plotted as the grey solid line in Figure **??**.

Because information is collected at the time of leaving university, HESA statistics alone cannot tell us how many working-age graduates there are in total in the UK, or anything directly comparable with our Figure 1. So we use the LFS 2016 to derive its implied number of people obtaining first degrees every year. This is also tricky because the LFS doesn't tell us when people obtained each of their qualifications, only when they obtained their highest qualification. Thus, when we plot the number of people over the year they obtained their highest qualification (solid black line in Figure **??**), the number overstates the truth in recent years. This is expected because for those with postgraduate qualifications, they must have obtained first degrees in some earlier unknown years. If we omit all the postgraduates as we do in the short-dashed line in Figure **??**, then obviously we would under-estimate the truth. If we assume that all the postgraduates obtained their first degree at age 22 and add them to the last series, then we get the long-dashed line in Figure **??**. This time series happens to be very similar to the HESA trend. In fact, all three measures from the LFS and the HESA one show a strong increase in the number of new graduates over time. Together with the aging of less-educated older cohorts, this means the overall proportion of graduates in the working-age population increased rapidly since the early 90s.

[6] The response rate can be found in the ONS Labour Force Survey Performance and Quality Reports. Link

[7] The statistics start in 1994-5 and can be downloaded here.

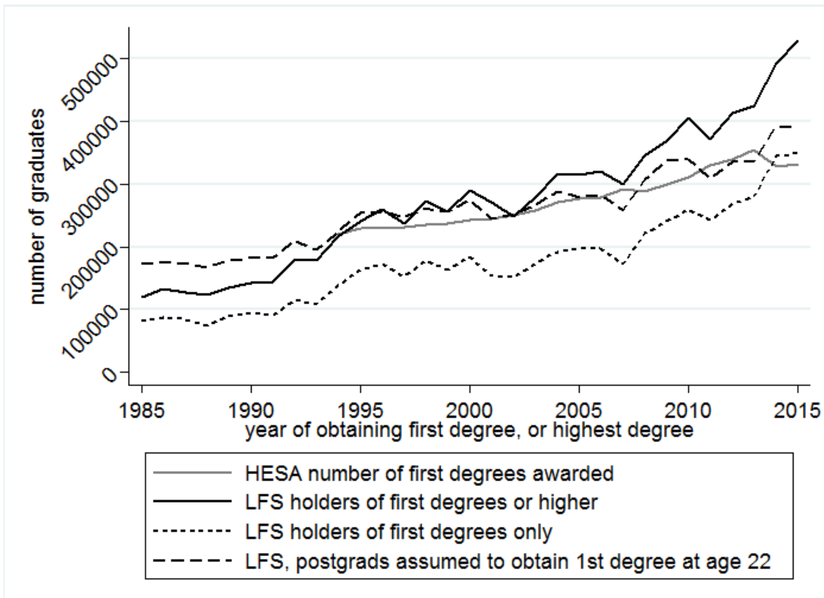


Figure 3: Number of first-degree graduates over year, HESA and LFS

Notes: the HESA series is the total number of UK-domiciled students obtaining first degrees by academic year, downloaded from here. For all the LFS series, I add up the weight of UK nationals with at least first degrees by the year they obtained their first degree or highest qualification (up to 2015). As I use 4 quarterly LFS datasets in 2016, the weight is divided by 4 to gross up to population totals. The first LFS series counts all those with first degrees or above, by the year they obtained their highest qualification. The second counts those whose highest qualification is a first degree, by when they obtained it. The third assumes that all those with higher degrees obtained their first degree at age 22 and then counts everyone by the year they obtained their first degrees.

5 Core Patterns by Birth Cohort

In section 2 in the paper, we aggregated the LFS data by 5-year age bands and year to examine time trends. Here we look at trends across birth cohorts. We aggergate the LFS data to the level of age and 5-year birht cohorts. The left subgraph of figure **??** shows the college wage premium over the life-cycle by cohort. The pattern is striking: the differential is increasing and concave over the life-cycle and there is not much difference across cohorts in either the shape or the level of the differential.

Unsurprisingly, when we regress these wage differentials on an age polynomial of order 5 and a complete set of cohort dummies, we find that the estimated cohort effects are quite flat. This is plotted in the right sub-graph of figure **??**. The same graph also plots the cohort effects in the BA proportion, which is net of age effects in the way. It is clear that the BA proportion is increasing across cohorts and the increase was particularly sharp between the 1965-69 cohort and the 1975-79 cohort. This coincides with the timing of the HE expansion. As the UK Higher Education sector expanded rapidly from 1988 to 1994, the first cohort to be directly affected was born in 1970.

One may suspect that as the BA proportion increased so much, their quality, especially at the lower end of the BA quality distribution, may have fallen. If this is true, one may expect a fall in the wage gap at lower percentiles in the distribution. In Figure **??**, we plot the cohort effects in the wage gap at various percentiles: the 10th, the 25th, the 50th, 75th and 90th.The trend across cohorts is relatively flat for all: the difference from the 1965-69 cohort is 0.1 log terms or less in absolute terms. The 10th percentile of the BA wage relative to the 10th percentile of the HS wage appears to have fallen a bit, by around 0.07 between the 1965-69 and 1975-79 cohorts. However, this decline in the wage gap was driven by a fast increase in the real HS wage at the 10th percentile, rather than a real wage decline among BAs at the 10th percentile. As shown in the 2nd sub- graph of Figure **??**, the 10th percentile of the HS group grew by more than 15% between the 1965 and 1985 cohorts,

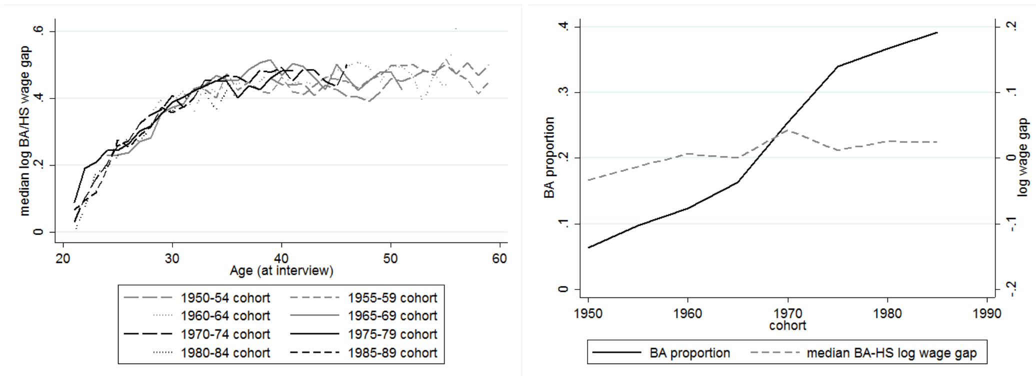


Figure 4: BA proportion and wage ratio over cohorts

Notes: We aggregate LFS data 1992-2016 up to the level of 5-year-birth-cohorts and age, where age is restricted to 20-59. We look at cohorts 1950-1985 only, so that each chort appears many years in the data. The BA-HS median wage ratio is plotted at this level in the left sub-figure. For the right sub-figure, we regress the BA proportion on cohort dummies and an age polynomial of order 5. For the BA proportion, the cohort effects are scaled to the observed porportion for 1965 cohort at 30 year old. For the wage gap, the cohort effects are normalized to 0 for the 1965 cohort.

when that of the BA group was about 10%, and growth was lower at the 25th and 50th percentiles for both groups. This decrease in within-group inequality, particularly for the HS group, looks like a natural consequence of the National Minimum Wage (NMW). The NMW was introduced in 1999 and has been raised at a faster pace than the median wage. Thus, there is no evidence of increasing supply of BAs reduing their relative wage in any part of the distribution.

6 Observable compositional changes

In this appendix, we present added investigations into compositional change effects. The first relates to the expansion of post-graduate degree holding.

The dark, solid line in Figure **??** plots the proportion of people with a postgraduate degree conditional on having a university degree. Similarly to

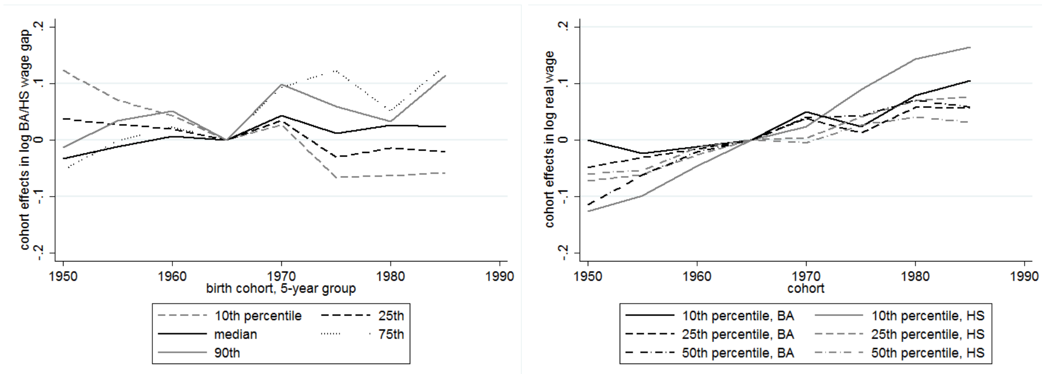


Figure 5: BA-HS wage ratio at different percentiles

Notes: We aggregate LFS data 1992-2016 up to the level of 5-year-birth-cohorts and age, where age is restricted to 20-59. We look at cohorts 1950-1985 only, so that each chort appears many years in the data. For each percentile shown in the left graph, we regress the BA-HS log wage gap on cohort dummies and an age polynomial of order 5. The cohort effects are normalized to 0 for the 1965 cohort. For the right graph, the dependent variable is the real log wage for each of the shown percentile of the education group.

what Lindley and Machin(2006) show for the US, the importance of postgraduate degrees increases for the UK in our period. Nonetheless, the proportion of postgraduates among university degree holders remains relatively low and so its change is unlikely to be a major driver of relative wage patterns. This is, in fact, what we see in the two wage gap lines in the figure. One line is a replotting of the line in figure 2 in the paper, which includes postgraduate degree holders among the university graduates,t re while the other line shows the wage gap relative to high school educated workers when we include only those with exactly a bachelor's degree and no higher. The two lines are very similar, with both showing nearly identical values in 1993 and 2016.

The second compositional shift we consider relates to immigration. The proportion of UK workers without UK nationality has more than doubled over the past two decades, from under 5% to above 10%. As immigrants are more likely to have university degrees (as confirmed in Figure **??**), the

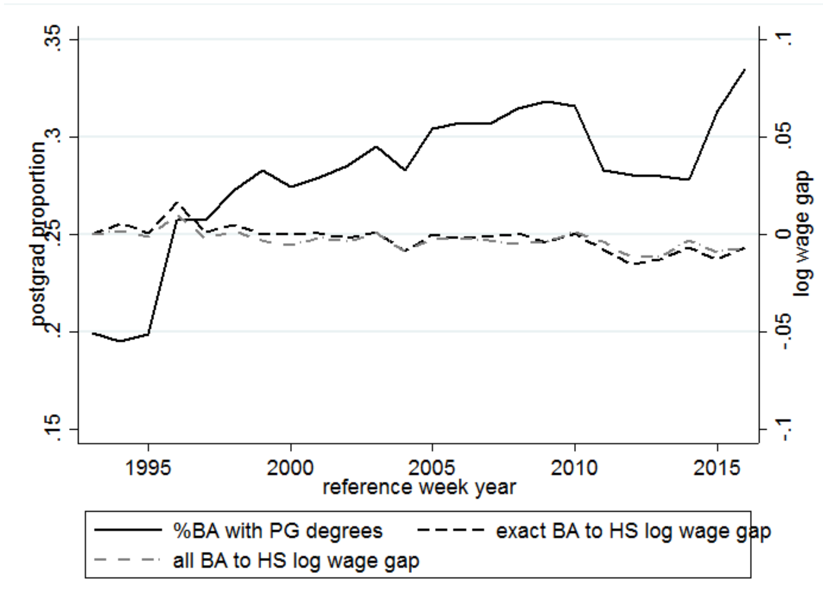


Figure 6: Year Effects for the Proportion of University Graduates with Advanced Degrees and the BA to HS Wage Ratio

Note: The year effects use the same sample selection and regression specification as for Figure 2 in the paper.

large flows of immigrants contribute directly to the aggregate increase in the share of BAs in the workforce. But it is not clear whether we should count every immigrant with a university education as the equivalent of a university educated native born worker. As demonstrated in **?**, immigrants often work in jobs that do not match their observed skills or qualifications, implying that a simple count of the number of immigrants with a university education may over-state the contribution of immigration to the effective supply of highly educated labour. Given the size of the increase in the immigrant proportion in the past 20 years, the positive bias in the measured supply of university labour may become substantial. To address this concern, we can look at the BA-HS wage ratio among UK nationals only. Figure **??** shows that the BA-HS log wage gap is essentially flat and very similar to the trend including

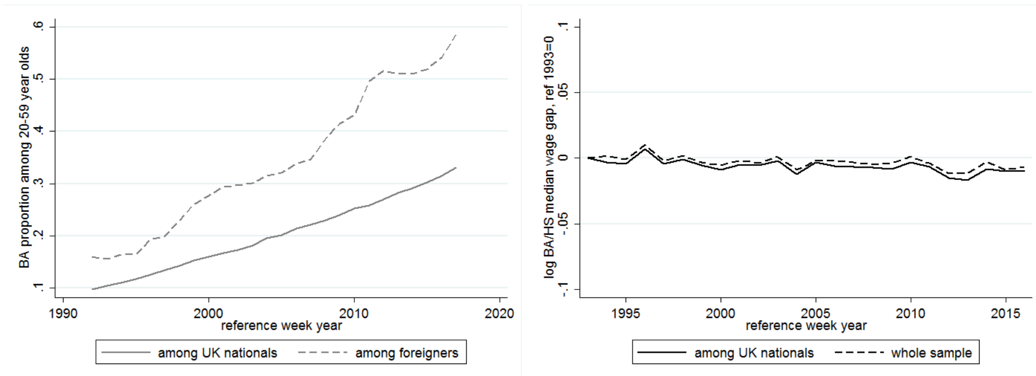


Figure 7: BA proportion and wage ratio over year, among those born in the UK

Notes: The BA proportions are not normalized. The year effects in wages are normalized to 0 in 1993. The whole sample series is the same as in Figure 2 in the paper.

immigrants.

The second observable composition dimension we investigated was between public and private sectors. Public sector employees are, on average, better educated and, with wages largely protected from direct market forces, we might expect wage differentials within the public sector to be more rigid. Given that, an expansion in the public sector might partly explain the patterns we have described. That possibility, though, falls short in two ways with respect to employment numbers. First, the proportion of workers in the public sector does not change substantially over our data period. Second, the growth in the proportion of workers with a BA is very similar between the private and public sectors.

The public-private sector dimension of movements in wage differentials is a bit more nuanced. In Figure **??**, we regress the wage differential at the year-age-band level on age dummies and year dummies and plot the year effects. The trend is slightly declining, but relatively flat. Compared to the whole economy (Figure 2 in the paper), the private sector trend is slightly

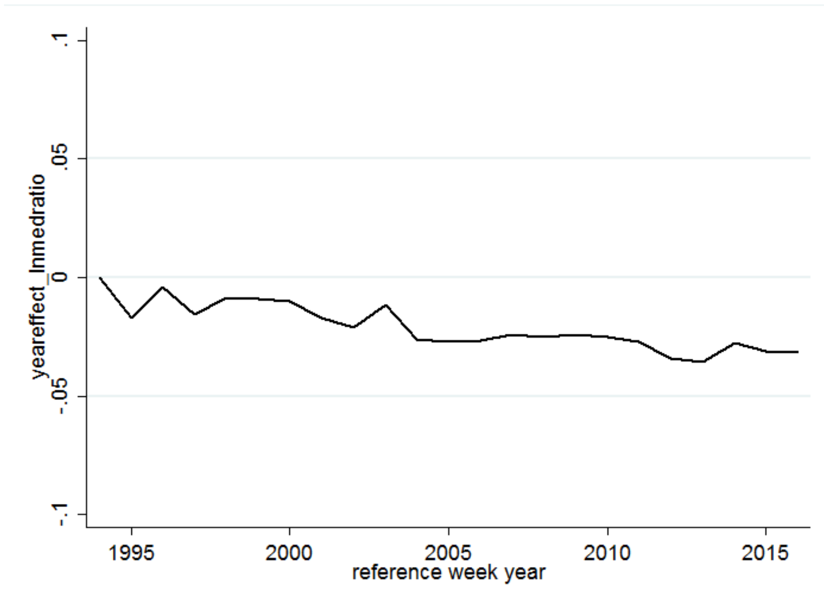


Figure 8: Time effect in BA-to-HS wage differentials, UK private-sector only

Notes: The time effect is normalized to 0 in 1994, becaue the variable on public versus private sector is available since 1994 only.

more decreasing: the change in the log wage gap over 22 years is about 0.03, rather than aobut 0.01 for the whole. This is still very small compared to what you might expect from an increase in the relative quantity of BA-to-HS (which is more than 1 full log point over the period).

One place we might look for a compositional shift is at the extensive margin: if the large increase in the relative supply of BAs combined with their constant relative wages induced a relative decline in the employment rate of BAs then this could imply changes in the relative "quality" of BA versus HS workers. In Figure **??**, we plot the estimated year effects in the employment rate of BA's and that of the HS population. The two series move very closely together over time. Thus, the lack of a relative wage response to the educational supply shift was not offset by a relative decline

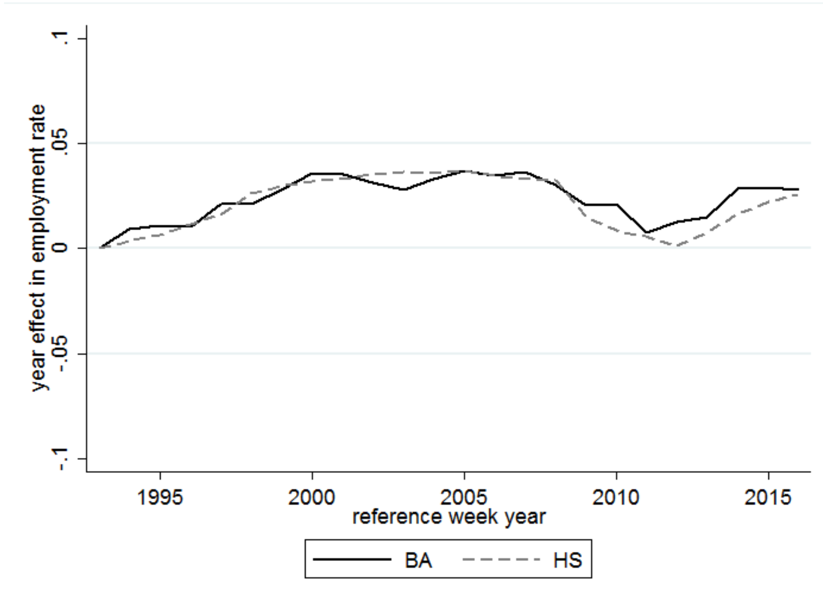


Figure 9: Time effects in employment rates among BAs and HS workers

Notes: The sample is LFS 1993-2016. The data is collapased to the level of year and 5-year age bands and education. We then regress the employment rate on a complete set of year dummies and age-band dummies. The time effect is normalized to 0 in 1993.

in employment. The change in relative employment rates is also small in the context of a near-tripling in the BA proportion over the period. Thus, we believe compositional shifts based on changes at the extensive margin are not a key driver of the main patterns.

7 Unobservable compositional changes: bounds

Implementation of a bounding approach rests on some (preferably minimal) assumptions about the model of wage determination. We will consider a simple but very standard model in which the wage for person i in education group j is given by:

$$\ln w_{ict} = \sum_{j=1}^{3} D_{ijt}\beta_{cj} + \sum_{j=1}^{3} D_{ij} f_{cj}(age_{it}) + \sum_{j=1}^{3} D_{ij}\lambda_j\eta_i + \epsilon_{ict} \tag{4}$$

where c indexes the person's birth cohort, D_{ij} equals 1 if person i is in education group j, and zero otherwise, f_{cj} is a cohort-and-education-group-specific age profile of wages, normalized to 0 for age 30 and ϵ_{ict} is an idiosyncratic error that is independent across time and people and of all other right hand side components in the regression. The specification incorporates a person-specific ability factor, η_i, the effects of which differ across education groups according to loading factors, λ_j. Importantly, both the distribution of η_i and its factor loadings are stationary across cohorts. This model is extreme in its assumption of only one ability factor, but it is also very standard and allows us to see clearly the effects of selection.

We are interested in the price per efficiency unit of workers with a given type of education ($\beta_{cj} + f_{cj}(age_{it})$ in (**??**)). This is unobservable because we do not observe the median wage for a composition constant group, Below we will adopt some assumptions and bounds on the composition-constant median wage for each education group.

We shall assume that the values of the λ's and other parameters are such that for each cohort, the three education groups correspond to three contiguous, non-overlapping ranges of ability. In particular, the groups are defined by two cohort-specific thresholds A_{uhc}, A_{hdc}. University graduates are those with $\eta > A_{uhc}$; high-school grads have $A_{hdc} < \eta \leq A_{uhc}$; and high-school dropouts have $\eta \leq A_{hdc}$. In theory, such a hierachical model of selection could be rationalized by a Roy model where individuals choose education levels by comparing their expected net present value of wages and of costs, and assuming $\lambda_u > \lambda_h > \lambda_d$ and that the costs of obtaining education are weakly decreasing in ability. In addition, the hierachical model fits the idea that university admission is largely rationed by prior attainment.

Consider a situation in which the university proportion increases between cohorts c and c+1, because there is less rationing. This corresponds to a decline in the value of A_{uhc}. Importantly, some individuals who would not get a university degree if they were born with their respective ability in cohort c will get a degree if they belong to cohort c+1 but no one is induced to make the opposite switch. That is, there will be flows in only one direction. Let's call the set of individuals who would get a degree if they face the conditions in cohort c+1 but not if they were in cohort c, "joiners". Their ability distribution has a range with a top value of A_{uhc} and so it lies entirely below that of the rest of university graduates in cohort c+1. The latter group have abilities that are high enough for them to enter university even when the costs were higher (as they were for cohort c). We will call them "stayers".[8] Obviously, the joiners' ability distribution lies above that of those who remain in the HS group in cohort c+1.

The observed wage distribution of BAs in cohort c+1 is a combination of that of the joiners and that of the stayers. Under our assumptions, if the number of BA's increases across cohorts then that must reflect an inflow of joiners but no outflow. That means we can use the observed median wage for BA's in the first cohort as corresponding to the median wage of the stayers. In the second cohort, we can form two extreme bounds based on what we assume about the joiners. In the first, we could assume that all the joiners have lower ability than the median stayer. We could then form one extreme estimate of the median wage for stayers by first trimming a number of observations equal to the number of joiners from the bottom of the observed wage distribution for the second cohort and then getting the median of the remaining observations. For example, if the size of the BA group increases from 20 to 30 percentage points of the population between cohort c and cohort c+1 at a given age, then we trim the bottom one third of the BA wage distribution of cohort

[8]Calling them stayers and joiners is a slight abuse of terminology since we are considering different cohorts and so there are no individuals actually staying or joining. Instead, these groups correspond to different ranges in the stationary η distribution.

c+1 and the median of the remaining distribution is the upper bound of the median of the stayers. Another extreme bound could be formed by similarly trimming the top third of the cohort c+1 distribution and getting the median for the remaining sample. However, under an hierarchical model of the kind we are discussing, the best the joiners could be is as good as the stayers (if they were better than the stayers, they would be in the sector already). If they are as good as the stayers then the observed median wage for BA's in cohort c+1 would be the same as the median wage for the stayers. Thus, the observed median forms the other bound on the cohort c+1 median wage for the stayers. The next two pages explain mathematically why the trimming method and the observed median are THE upper and lower bounds under the hierachical model. Differencing these bounds for the stayers' median wage in cohort c+1 from the observed median wage for cohort c then gives us bounds on the movements in the price for BA labour for a composition constant group.

Because people (or, more properly, ability values) can be induced to switch into or out of higher education but not both at the same time, we can decompose the distribution function for BA wages in cohort c+1 into a component related to the distribution function for the "stayers" and a component for the "joiners":

$$\begin{aligned}\Pr(\ln W_{uc+1} < w|\eta > A_{uhc+1}) &= p_{uc+1}\Pr(\ln W_{uc+1} < w|\eta > A_{uhc}) \\ &+ (1 - p_{uc+1})\Pr(\ln W_{uc+1} < w|A_{uhc} \geq \eta > A_{uhc+1}), \forall w \quad (5)\end{aligned}$$

where, p_{uc+1} is the proportion of the university educated in cohort c+1 who are stayers. Equation (**??**) holds for any wage level w, but we are interested in a particular level: the median wage in cohort c+1 for the university sector stayers, denoted as $\tilde{w}_{uc+1}$.

We can write $\tilde{w}_{uc+1}$ as,

$$\tilde{w}_{uc+1} = \beta_{c+1u} + f_{c+1u}(age_{it+1}) + \lambda_u med(\eta_i + \epsilon_{ic+1t+1}|\eta_i > A_{uhc}) \quad (6)$$

Assuming stationarity of the η and ϵ distributions across cohorts, differencing this relative to the median conditional university wage in cohort c at the same age, age^* would yield,

$$\tilde{w}_{uc+1} - med(\ln W_{uct}|\eta_i > A_{uhc}) = \beta_{c+1u} + f_{c+1u}(age^*) - \beta_{cu} - f_{cu}(age^*) \quad (7)$$

That is, by comparing wage movements for people with the same set of $\eta's$ (the ones corresponding to choosing to get a university degree under either set of costs), we could obtain an estimate of the change in the actual wage profile across cohorts.

We cannot observe $\tilde{w}_{uc+1}$ because we are comparing across cohorts and so cannot see who has ability levels that would result in their choosing the university degree in the different rationing situations. But we can obtain bounds for it. Returning to equation (**??**), we can obtain an estimate of p_{uc+1} based on changes in the size of the u group between cohort c and c+1 combined with the argument that people (or, rather, ability levels) either enter or leave the group but not both. We know that the second term on the right hand side of (**??**) ($\Pr(\ln W_{uc+1} < \tilde{w}_{uc+1}|\eta > A_{uhc})$) equals 0.5 by the definition of $\tilde{w}_{uc+1}$, and the left hand side corresponds to a quantile of the conditional distribution of wages for the u group in the c+1 cohort, and so is calculable from the data. That only leaves the last term ($\Pr(\ln W_{uc+1} < \tilde{w}_{uc+1}|A_{uhc} \geq \eta > A_{uhc+1})$) unknown and unknowable. However, since it is a probability, we can bound it on one side as $\Pr(\ln W_{uc+1} < \tilde{w}_{uc+1}|A_{uhc} \geq \eta > A_{uhc+1}) = 1$, which corresponds to the marginal people who obtain a degree in cohort c+1 but would not have done so in cohort c having wages that place them below the median wage for the group who would get a degree in either cohort. Based on this, we can get an upper bound on $\tilde{w}_{uc+1}$ by solving,

$$\Pr(\ln W_{uc+1} < \tilde{w}_{uc+1}|\eta > A_{uhc+1}) = \frac{1}{2}p_{uc+1} + (1 - p_{uc+1}), \quad (8)$$

This is equivalent to trimming the bottom $(1 - p_{uc+1})$ proportion of observations from the c+1 university wage distribution and obtaining the median of the remaining sample.

Since the abilities of university "joiners" between cohort c and c+1 are assumed to be entirely below the abilities of the "stayers", a joiner's wage can be higher than a stayer's only when the joiner has a particularly positive shock ϵ_{it} or the stayer has a particularly negative shock. As the idiosyncratic shock is assumed to be independent of ability, it follows that the joiners' wage distribution is first order stochastically dominated by that of the stayers. Mathematically,

$$\Pr(\ln W_{uc+1} < \tilde{w}_{uc+1} | A_{uhc} \geq \eta > A_{uhc+1}) \geq \Pr(\ln W_{uc+1} < \tilde{w}_{uc+1} | \eta > A_{uhc}) \tag{9}$$

Using the right side of this expression as the lower bound on $\Pr(\ln W_{uc+1} < \tilde{w}_{uc+1} | A_{uhc} \geq \eta > A_{uhc+1})$ in (**??**) implies that the right hand side of (**??**) just equals 0.5. That is, the other bound is the c+1 median itself.

Meanwhile, we can implement a similar exercise for the HS group. In this case, though, if the BA group grows between cohort c and c+1 this must be directly matched with an emigration of individuals from the top of the HS ability distribution between those cohorts. In trimming terms, this means that one bound can be obtained by appending a number of workers equivalent to the increase in size of the BA group to the top of the cohort c+1 wage distribution for HS workers. At the same time, if the Drop-out group shrinks then, under the single factor Roy model, they must have moved to the bottom of the ability distribution in HS and we would trim a number of workers equivalent to the decrease in size of the Drop-out sector from the bottom of the cohort c+1 HS distribution. Doing both the BA and Drop-out related trimming and appending yields a new adjusted HS sample in cohort c+1 that corresponds to one bound on the wages for the HS group stayers. Taking the difference between the median wage in that sample and the actual median wage for HS workers in cohort c yields an upper bound on the change in the log wage profile at a given age for HS workers. Consider the benchmark case where the upper bound scenarios for the BA and HS workers correspond to one another (i.e., the movements out of the top of the HS distribution become the movements into the bottom of the BA distribution). We can then obtain one bound on the movement in the university - high school wage differential by taking the difference between the upper bound on the movement in the university median and the upper bound on the movement in the high school median. The other bound is the actual change in the median wage ratios shown in Figure **??**.

We repeat the sample trimming exercise for each cohort using the 1965-69 cohort as the base of comparison (cohort c in our example). The resulting quality-adjusted wage differentials are reported in the left panel of Figure **??**. The second panel shows cohort effects derived in the same manner as in the earlier figures. The cohort effects show an increase in the adjusted upper bound differential between the 1965-69 and 1970-74 cohorts. Given that the other bound is the actual change in the median wage ratio, the implication is that under this ability model, one cannot argue that selection on unobservables obscured what was actually a decline in the true wage differential. For the difference between the 1965-69 and 1975-79 differential, one bound shows a near zero change and the other shows a 4 percent decline. Thus, here there is some room to argue that selection is hiding a true decline in the ratio, but that decline is still very small compared to a doubling of the proportion of the population with a BA. For the post-1980 cohorts, the bounds include larger declines - about 15% relative to the 1965-69 cohort. However, a glance at the profiles in the left panel suggests the need for some caution in interpreting the cohort coefficients. The age profiles for the various cohorts no longer look parallel once the extreme bound trimming is implemented, implying that the age at which we evaluate the cohort differences can alter our conclusions. But, overall, our conclusion from this exercise is that, under this model of ability, selection on unobservables cannot explain why we do

Figure 10: UK Median BA-to-HS wage ratio, adjusted to the education split of 1965 cohort

Notes: For each age and cohort, we adjust the wage distribution by using the proportions observed for 1965 cohort as reference points. For example, if the observed proportion of BAs is higher than that for the 1965 cohort at the same age, we would trim the bottom of the observed BA distribution.

not see a large decline in the education wage differential for the cohorts with the largest increase in their education level.

8 Implications of Exogenous Skill Biased Technological Change with Managerial Tasks

In this appendix, we examine the implications of an exogenous skill biased technological change in the context of a standard production function that incorporates two skill levels and two broad types of tasks. The model exposition is similar in nature to that used in the **?** paper on technology diffusion and the labour market. In particular, we consider a model in which one technology is in use at a time. Output, Y, is produced according to the Cobb-Douglas production function:

$$Y = M^{\alpha} L^{1-\alpha} \tag{10}$$

where, M is hours of managerial labour, L is hours of production labour, and α is a parameter. Each task is performed by a combination of skilled and unskilled labour, with the labour aggregated through CES functions:

$$M = [aS_M^{\sigma} + (1-a)U_M^{\sigma}]^{1/\sigma} \tag{11}$$

and

$$L = [bS_L^{\rho} + (1-b)U_L^{\rho}]^{1/\rho} \tag{12}$$

where, $\frac{1}{1-\sigma}$ is the elasticity of substitution between skilled and unskilled labour in managerial tasks; $\frac{1}{1-\rho}$ is the elasticity in labouring tasks; a and b are parameters; S_M is the amount of skilled labour in the managerial task; and U_L is the amount of unskilled labour in the basic labouring task. We assume that skilled labour is relatively more productive in the managerial tasks (i.e., $a > b$) and that skilled and unskilled labour are more substitutable in the labouring task (i.e., that $\rho > \sigma$).

We assume that the numbers of unskilled and skilled workers in the economy are given exogenously in any period and that each worker supplies a fixed endowment of labour inelastically. Market clearing in the labour market corresponds to the total number of workers with each skill level in the economy being equal to the sum of the numbers employed in the various occupations and technologies:

$$S = S_L + S_M$$

and,

$$U = U_L + U_M$$

Workers of each skill type can choose freely whether to work as a manager or a labourer and so there will be one skilled wage, w_s and one unskilled wage

w_u.

In this framework, a skill-biased technological change can be represented as an increase in a, i.e., an increase in the productivity of S workers as managers. This captures both that the technological change favours S workers and that it is related to management tasks. Note that we are assuming that the technological change arrives exogenously and alters the production function of firms without them choosing whether or not to adopt the new technology.

To understand the impact of this change note that, working from the firm's first order conditions, it is straightforward to show that the wage skill ratio is,

$$\begin{aligned} \frac{w_s}{w_u} &= \frac{a}{1-a}(\frac{S_M}{U_M})^{\sigma-1} \\ &= \frac{b}{1-b}(\frac{S_L}{U_L})^{\rho-1} \end{aligned} \tag{13}$$

Rearranging these expressions slightly, we get:

$$\frac{a}{1-a}\frac{S_M^{\sigma-1}}{S_L^{\rho-1}} = \frac{w_s}{w_u} = \frac{b}{1-b}\frac{U_M^{\sigma-1}}{U_L^{\rho-1}} \tag{14}$$

In the context of this model, in order to match the main data pattern of an increase in S accompanied by no change in $\frac{w_s}{w_u}$, equation (**??**) shows that we need an increase in a of just the right size so that the skill biased demand increase just balances the relative supply shift. We view it as somewhat implausible that there were an exogenous set of technological changes that just balanced the supply shifts over an extended period of time, but we cannot reject that this could have occurred. Instead, we ask about the further implications of such changes if this were the mechanism driving our main data patterns. Examining (**??**), note that if a increases then the ratio of the number of skilled workers who are managers to the number who are labourers must also increase in order to match the unchanging wage ratio. This is the opposite of the implication from our endogenous technological choice model in which the expansion in S is accompanied by a decreasing proportion of S workers who are managers.

9 Results on education expansion and wages in other countries

Our analysis fits with results in **?**. She uses two European surveys to examine wage and education patterns in 12 European countries between 1994 and 2009. Many of the economies in her data are in our sample of countries with substantial educational growth in this period.[9] and she shows that the proportion of the population who are tertiary education graduates for all of these countries pooled together goes up by 50% across the birth cohorts she studies. The dependent variable in the main exercise in the paper is the wage premium to having a tertiary or other post-secondary education relative to a high school diploma. This is regressed on a relative educational supply variable, a variable intended to capture skill biased demand shifts, and a complete set of country, year, and birth cohort effects. The results indicate statistically significant but very small relative supply effects with a 10% increase in the relative number of post-secondary to secondary graduates being associated with a 1.2% decline in the log wage ratio for the two groups in their OLS estimates. In addition, the relative demand effect is very small and not statistically significant from zero. Thus, Crivellaro's results with a set of 12 European economies matches closely with our results for the UK: substantial increases in education have little effect on the wage ratio and there is also little evidence of an ongoing skill biased demand shift.

Our results also fit with findings in some other papers examining wage differentials and education increases in other economies. **?** examines these

[9] The countries in her data are: Austria, Belgium, Germany, Denmark, Spain, Finland, France, Greece, Ireland, Italy, Portugal, and the UK.

patterns for Taiwan, which underwent a dramatic boom in creating new post-secondary institutions between 1990 and 2000. As a result of that boom, between 1990 and 2010, the number of post-secondary graduates increased by a 600%. Yet over that same period, the difference between the mean log hourly wage for university graduates and workers with less than a university education was quite flat. That university wage premium was approximately 0.6 in 1980, 1990, 2000, and 2010. Chen interprets this outcome within an exogenous skill biased technological change model. As we argued earlier, for such a model to generate a flat premium trend requires a lucky, exact balance of relative supply and exogenous demand shifts. We believe that our model, in which the flat profile provides a more natural explanation. **?** and **?** examined a similarly large increase in education levels driven by policy changes in South Korea in the 1980s and 1990s. Between 1990 and 2005, the proportion of high school graduates who enrolled in a post-secondary programme increased from approximately 30% to 80%. **?** show that the post-secondary wage premium declined in the 1980s but was flat during the substantial educational expansion that started in the mid-1990s. They show that the latter patterns coincided with an increase in expenditures on IT and conclude that the flat premium reflected an endogenous technological change model. These trends could fit with our model, with the initial decline in the wage premium in the 1980s corresponding to a period before the economy entered the cone of diversification. The post-1994 period is then the period of transition to taking up more skill-biased technologies, as evidenced by the coinciding increase in IT expenditures. Finally, **?** examine wage impacts of an earlier large increase in post-secondary attainment in Norway, taking advantage of regional variation in the creation of universities in the 1970s. They show that the regions where new universities were added had a significant jump in the education level of their workforce but that the wage differential between university and high school educated workers either stayed flat or increased. They interpret this within the context of an endogenous technological change model and show evidence that the productivity of skilled workers increased in relative terms in the regions with new universities.

10 OECD Data on Wage Differentials

In this appendix, we present the results from a simple exercise based on the data on educational attainment and wage differentials from **?**. As mentioned in the text, we focus on the set of OECD economies that have a lower proportion of their population than the US with a tertiary education in the initial year of the data (1997) and experience a growth in that proportion by at least 40% by 2010. In the table, below, we present estimates for this set of countries from a regression on a constant and a linear trend of the wage ratio between the mean annual earnings of all workers aged 25 to 64 with a tertiary education and the mean annual earnings of workers with an upper secondary education being their highest education level. Of the 11 countries meeting our criterion, 7 have trend coefficients that are not statistically significantly different from zero, 2 have positive and significant coefficients, and 2 have negative and significant coefficients.

References

Borghans, L. and B. ter Weel (2008). Understanding the technology of computer technology diffusion: Explaining computer adoption patterns and implicaitons for the wage structure. *Journal of Income Distribution 17*(3), 37–70.

Card, D. and T. Lemieux (2001). Can falling supply explain the rising return to college for younger men? a cohort-based analysis. *The Quarterly Journal of Economics 116*(2), 705–746.

Table 2: Regressions of Wage Differential on a Time Trend by Country

Country	Const	t
Belgium	30.02**	0.15
	(2.19)	(0.23)
France	47.25**	0.036
	(4.24)	(0.46)
Ireland	47.84**	1.47
	(12.2)	(1.2)
Korea	31.82**	1.86*
	(7.26)	(0.73)
New Zealand	22.26**	-0.28
	(2.9)	(0.3)
Norway	31.98**	-0.33**
	(0.91)	(0.11)
Poland	72.66**	-0.18
	(6.58)	(0.68)
Spain	19.63**	1.72**
	(2.88)	(0.3)
Sweden	32.57**	-0.58**
	(1.06)	(0.11)
Switzerland	57.76**	-0.21
	(2.12)	(0.21)
UK	58.59**	0.045
	(2.81)	(0.29)

Authors's calculations based on data from **?**. Standard errors in parentheses.
*, ** statistically significant at the 1% and 5% significance levels, respectively.

Carneiro, P., K. Liu, and K. G. Salvanes (2014). How does labor demand react to changes in the supply of skills? evidence from schooling reforms.

Carpentier, V. (2006). Funding in higher education and economic growth in france and the united kingdom, 1921-2003. *Higher Education Management and Policy 18*(3), 1–26.

Chen, U. (2013). The impact of the expansion in post-secondary education on the university wage premium in taiwan. Technical report, University of Michigan.

Choi, K.-S. and J. Jeong (2005). Technological change and wage premium in as small open economy: The case of korea. *Applied Economics 37*(1), 119–131.

Choi, S. (2015). When everyone goes to college:the causal effect of college expansion on earnings. *Social Science Research 50*, 229–245.

Crivallero, E. (2016). The college wage premium over time: Trends in europe in the last 15 years. *Research in Labor Economics 43*, 287–328.

Dustmann, C., T. Frattini, and I. P. Preston (2013). The effect of immigration along the distribution of wages. *The Review of Economic Studies 80*(1), 145–173.

Katz, L. F. and K. M. Murphy (1992). Changes in relative wages, 1963-1987: Supply and demand factors. *The Quarterly Journal of Economics 107*(1), 35–78.

OECD (2012). Education at a glance: Oecd indicators. Technical report, OECD.