

Holford, Angus J.

Working Paper

Youth Employment and Academic Performance: Production Functions and Policy Effects

IZA Discussion Papers, No. 10009

Provided in Cooperation with:

IZA – Institute of Labor Economics

Suggested Citation: Holford, Angus J. (2016) : Youth Employment and Academic Performance: Production Functions and Policy Effects, IZA Discussion Papers, No. 10009, Institute for the Study of Labor (IZA), Bonn

This Version is available at:

<https://hdl.handle.net/10419/145143>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

IZA DP No. 10009

**Youth Employment and Academic Performance:
Production Functions and Policy Effects**

Angus Holford

June 2016

Youth Employment and Academic Performance: Production Functions and Policy Effects

Angus Holford
*ISER, University of Essex
and IZA*

Discussion Paper No. 10009
June 2016

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

Youth Employment and Academic Performance: Production Functions and Policy Effects*

We identify the effects of part-time employment, study time at home, and attitudes in school, in the production function for educational performance among UK teenagers in compulsory education. Our approach identifies the factors driving differences between the reduced form 'policy effect' of in-school employment, and its direct effect or 'production function parameter'. Part-time employment is shown to reduce performance among girls but not boys, because employment crowds out both study time at home and positive attitude in school to a greater extent for girls than boys. Part-time work also induces earlier initiation into risky behaviours for girls than boys.

JEL Classification: C35, I21, J22, J24

Keywords: labour supply, human capital, education production function, child development

Corresponding author:

Angus Holford
Institute for Social and Economic Research
University of Essex
Wivenhoe Park
Colchester, CO4 3SQ
United Kingdom
E-mail: ajholf@essex.ac.uk

* This work was carried out during PhD studies supported by an Economic and Social Research Council '1+3' studentship in Economics [reference number ES/I025499/1], with supervision and advice from Emilia Del Bono and Stephen Pudney, and postdoctoral research supported by the Economic and Social Research Council Research Centre on Micro-Social Change [reference number ES/L009153/1]. Thanks to Pedro Carneiro and Thomas Crossley for their advice as my PhD examiners, and to seminar participants at the ZEW Mannheim, IZA-SOLE Meeting in Munich, Lancaster and Huddersfield. All errors are my own. This work is based on data from the Longitudinal Study of Young People in England (Secure Data Service Access), produced by the Department for Education and National Centre for Social Research, and supplied through the Secure Data Service at the UK Data Archive. This work also uses local labour market statistics produced by the Office for National Statistics. The data are crown copyright and reproduced with the permission of the controller of the HMSO and Queen's Printer for Scotland. The use of the data in this work does not imply the endorsement of the Department for Education, National Centre for Social Research, ONS, or Secure Data Service at the UK Data Archive in relation to the interpretation or analysis of these data.

NON-TECHNICAL SUMMARY

Around 25% of 13-16 year-olds below the school leaving age in England, and 60% in the US, take some formal paid employment during school term time. Having a job while still in full-time compulsory education may improve teenagers' and young adults' welfare by enabling them to consume more, and by giving them more bargaining power over what they consume and how they spend their time. It may also improve their future educational or labour market outcomes by improving their financial literacy or communication skills, or possibly by forcing them to think about the type of job they want in the longer term, and what they need to do to get there.

However, at the same time, having a job reduces the time available for other activities, so is likely to crowd out the amount of time spent studying, or if the student is more tired, its quality. As the student has an income, it may also enable take-up of risky behaviours (e.g. alcohol, smoking and cannabis consumption). As such, whatever the direct human capital benefits of having a job while in school, the overall effect may be to reduce academic performance. This matters in the longer term because the qualifications people leave school with substantially determine their subsequent education and adult labour market opportunities.

In this paper we show that longer hours of employment by 14 and 15 year-olds in England increases take-up of risky behaviours, and reduces both the amount of study undertaken at home and outside of lessons, and students' reports that their schoolwork is important. (Having a low-skilled part-time job does not raise students' motivation for schoolwork as a means to widening their future occupational options, rather reduces motivation either through general tiredness or shifting preferences towards the activities their employment and earnings open up). This occurs to a much greater extent among girls than boys. The only time use where boys' employment resulted in greater crowd-out was sporting participation, which we show is less important for academic performance.

Overall, we find no effect of in-school employment at age 14 or 15 on GCSE performance (age-16 exams, at the end of compulsory education) for boys, but a significant negative effect of employment at age 15 for girls. For this group, an additional hour of paid employment per week reduces GCSE performance by approximately 1 grade in one subject (4.3% of a standard deviation in the total point score). About 25% of this effect can be accounted for by the study time crowded out. Those in work at age 15 on average work 6 hours per week, meaning that having a job is likely to reduce their subsequent educational and labour market opportunities substantially. This effect will be particularly severe for students near the cut-off for progression into post-compulsory education.

1 Introduction

Having a job while still in education may improve teenagers' and young adults' immediate welfare by affording them additional consumption. It potentially improves their stock of cognitive and non-cognitive human capital, in the form of financial literacy, communication skills, and lower discount rates (Oettinger, 1999; Light, 2001), but may detract from human capital accumulation by changing the student's allocation of time, crowding out time devoted to productive 'investments' in educational activities (Becker, 1965; Ruhm, 1997). This means the expected effect on test scores is ambiguous. In this paper we evaluate the effect of in-school employment on educational outcomes at the end of compulsory schooling, isolating the contributions of two mechanisms through which in-school employment affects later labour market outcomes. The first is the direct effect of the human capital acquired while working part-time in the production function for educational performance (the human capital mechanism). The second is the indirect effect caused by changes in study time at home, attitude or engagement during school hours, or other activities including risky behaviours outside of school hours (the time allocation mechanism).

The issue of part-time work during schooling matters to many individuals. In the US, approximately 60% of 14-15 year-olds take paid employment at one time (Buscha et al., 2011; Bureau of Labor Statistics, 2000), with 92% of the National Longitudinal Study of Youth (NLSY) 1997 cohort having worked at some point during High School (Hotz et al., 2002). The phenomenon is less widespread in the UK, but around 25% of 13-16 year-olds below the school leaving age in England take formal paid employment during school term time at any point. Declining participation in part-time work has been blamed by employers' organizations for young adults being increasingly ill-prepared for full-time employment, with negative implications for workforce productivity (UK Commission on Employment and Skills, 2015a, 2015b). Moreover, with the qualifications obtained from formal education substantially determining subsequent education and adult labour market opportunities, the negative or positive effects of in-school employment are likely to restrict or enhance trajectories of human capital accumulation over a much longer period for a significant proportion of every cohort of young adults (Dustmann and van Soest, 2007).

From an individual standpoint, with an increasing age of compulsory education, and increasing level of participation in Higher Education, employment experience during education may be

increasingly important to differentiate oneself in the labour market. In previous studies this form of part-time work has variously been shown either to benefit (Häkkinen, 2006; Light (2001); Ruhm, 1997; Molitor and Leigh, 2005) or to cost (Hotz et al, 2002) students in terms of adult labour market outcomes. This occurs through several competing mechanisms. These include employers' use of in-school employment as a screening device (Stiglitz, 1975), or their interpretation of part-time experience as a signal of ability to multitask (Spence, 1973) or financial constraints during schooling, and hence their social background (a form of statistical discrimination; Arrow, 1973, Baert et al., 2015). Alternatively, this 'learning on the job' may improve productivity and wages, or the employee may exploit personal connections and market information via his expanded social network (Granovetter, 1975).

We use data from the Longitudinal Study of Young People in England (LSYPE). This followed a cohort of teenagers aged 13-14 in 2004 through to their entry into post-compulsory education. The longitudinal structure and rich data on students' study time and attitude during school hours, active leisure, social activities and risky behaviours enables us to identify how the effect of part-time work and the productivity of these 'investment activities' varies according to proximity to final exams.

Our results are presented in two parts. First we evaluate the reduced-form effect of employment at age 14 and age 15 on grades at age 16. This is the "policy effect", the overall impact of an exogenous change in a specific input to the production function - part-time work - on measured performance. Several strategies have been used to identify the net effect of part-time employment on academic outcomes, including OLS regression in a contemporaneous (Lillydahl, 1990; Montmarquette et al., 2007) or cumulative specification (Ruhm, 1997; Payne, 2004), individual fixed effects or differences-in-differences (Oettinger, 1999; Sabia, 2009; Buscha et al., 2011), and instrumental variables (Stinbrickner and Stinebrickner, 2003; Häkkinen, 2006).

Our approach is most similar to that of Tyler (2003), Rothstein (2007) and Kalenkoski and Pabilonia (2010) who use state or county unemployment rates, local wage rates and state laws on teen employment as instruments for employment among US high school or college students in a cross-sectional framework. Tyler (2003) and Kalenkoski and Pabilonia (2010) both show large negative effects of in-school employment on college performance. In Rothstein (2007) the effect appears to be negligible. We too use indicators of local labour market opportunities as instruments, but at a typically smaller geographical unit. The first is the youth (age 18-24)

unemployment rate. The second is the density of small businesses in the retail, wholesale, hotel and restaurant sectors. We also use month of birth within the academic year. This determines whether restrictions on teenage employment will bind on the teenager at the time when vacancies arise.

The preceding literature is either silent on gender differences in outcomes resulting from in-school employment (focusing on males only or using a pooled sample), or finds only small differences. Our results show no effect of employment at age 14 for either boys or girls, holding age 15 employment constant. However, while no effect of age 15 employment is found for boys, an additional hour of employment per week at age 15 is shown to reduce age 16 educational performance by around 4.3% of a standard deviation for girls. On average, girls outperform boys by around 24% of a standard deviation, so this effect size means that introducing a boy and girl into a typical working week of 5-6 hours would be expected to eliminate the performance gap between them.

We make an important contribution with regard to timing of employment too. When we observe our sample at age 14, students have not yet begun the two-year courses determining their school-leaving qualifications, while at age 15 they are about halfway through. Our results suggest that part-time work does not cause a permanent setback for academic ability, but makes it harder to perform well in qualifications one is currently studying for, even well before the main intensive exam period.

Since the policy effect includes the indirect effects of changes to other inputs made by optimising individuals, it will likely differ from the “production function parameter”, or direct human capital cost or benefit of in-school employment. For example, De Fraja et al. (2010) and Datar and Mason (2008) show that students and parents substitute school inputs for their own. In this case policy effects of increased school quality are expected to understate such interventions’ benefits in the education production function. (Browning and Heinesen, 2014, discuss the conditions under which this will apply in greater detail).

In the second part of our analysis we evaluate the effect of part-time work holding constant latent variable representations of an additional input; study time at home, or attitude during school hours. We focus on adolescent self-investments and abstract from the role of parental investments. We argue that the latter bestow little marginal benefit to the development of human capital by the child’s mid-teens (Del Boca, Monfardini and Nicoletti, 2012), but we

proxy for the differences in initial conditions that parental investments create by controlling for a rich set of parental and household characteristics.

We estimate a system of simultaneous equations in which, at each point in time, part-time work is determined by exogenous variation in our instrumental variables, and one additional input (either study time, or effort and attitude during school hours) is crowded out by exogenous variation in labour supply. Both part-time work and the input at both points in time (age 14 and age 15) then enter the production function of the educational outcome measured at age 16. We thus observe the direct effect of part-time work, and that which is mediated via the additional input.

Identification of the production function parameter on each endogenous input rests on the level of this input being causally and significantly dependent on their part-time work. We show that study time at home is not significantly crowded-out by part-time work for boys. As a result, we are unable to identify the productivity of their study time at the margin driven by part-time work. We show instead that boys' main margin of adjustment, by which they accommodate their part-time work, is reducing the time devoted to 'active leisure', mainly sporting activities. Part-time work is also shown to induce participation in risky behaviours, with each hour per week raising the probability of 'ever smoking', 'ever having tried cannabis', and 'ever consuming alcohol' by approximately 1 percentage point at age 15. Effects of a similar size are found even at a year younger for girls.

We obtain well-identified and precise estimates for the production function parameter of students' attitude in school at ages 14 and 15 for boys and girls. At age 15 (during the courses determining their school-leaving qualifications but a year before the final exams), a one standard deviation improvement in this measure, at the margin crowded out by labour supply improves academic performance by 17 to 22% of a standard deviation.

2 Data and institutional background

2.1 LSYPE data

We use data from the first two waves of the Longitudinal Study of Young People in England (LSYPE), linked to data on performance in school exams in the National Pupil Database (NPD). This sample is drawn from a single academic cohort of teenagers in England who are interviewed

annually from 2004, at age 13-14.

Term time employment is captured by the questions “Do you ever do any work in a spare-time paid job, even if it is only for an hour or two now and then? (Please don’t include jobs you only do during the school holidays or voluntary work)”, and for those answering ‘yes’, “How many hours on average do you usually work in this job (or jobs) during a term time week? Please include any hours you work at the weekend during term-time”. We do not observe the type of employment undertaken, but as this question explicitly relates to school term time, the crowding out of study time or other extra-curricular activities is always a salient concern. We define two ‘investment’ factors (study time at home, and effort in school) using a further eight questions on time use, frequency of participation in key activities, and educational attitudes. These are outlined in full in Tables 5 and 6.

2.2 Regulation of employment of children in England

The rules governing the employment of children in England are set by the Department for Education (see guidelines in DCSF, 2009). Children aged at least 13 but less than the school leaving age may undertake ‘light’ work, deemed as not being harmful to their health, safety or development.¹ There are age-specific restrictions on the types and hours of work children may do. Those under 16 cannot work ‘mainly or solely’ for the sale of alcohol, for example. Those in compulsory education may work only 12 hours per week in term time, including a maximum of 2 hours on a weekday or Sunday; 8 hours on a Saturday (5 hours if under 15 years); one hour before school on a weekday; and none during school hours or after 7pm on a school night.

2.3 The exam system in England

Our measure of academic performance at age 16, the end of compulsory schooling, is the ‘GCSE and equivalent total point score’, standardized by subtracting the population mean and dividing by the population standard deviation. The majority of students obtain GCSE (General Certificate of Secondary Education) qualifications in around 10 subjects, of which at least English, Maths and Science (covering topics in all of Chemistry, Physics and Biology) are compulsory. Students are graded from A*-G, plus U (“unclassified”), and usually require five GCSEs at grade A*-C, including English and Maths to continue to post-compulsory education. Around

¹Full-time education is compulsory until the last Friday in June of the academic year when the child turns 16.

60% achieve this threshold. GCSEs and equivalent grades obtained from General National Vocational Qualifications (GNVQs) in non-academic subjects contribute to a student's overall point score.² To account for each student's trajectory of academic outcomes over time, we also control for students' total point score in low-stakes 'National Tests' in English, Maths and Science at age 11 and age 14. All these qualifications and examinations are criterion-based, measuring students' performance against a fixed standard rather than relative to their peers.

2.4 Descriptive statistics

Table 1 shows descriptive statistics relating to family background, employment and earnings at ages 14 (wave 1) and 15 (wave 2). Those in employment are positively selected on age 14 prior educational performance, and go on to perform better at the end of compulsory schooling. Those in employment also on average come from higher income households and are less likely to have a lone parent. These differences are statistically significant at the 1% level for both boys and girls at every point in time. While girls are positively selected on parental education, boys are marginally negatively selected. Overall however the table suggests that positive selection driven by social networks or 'soft skills' outweighs any negative selection driven by household financial constraints.

The propensity to work rises more markedly with age among girls, who start from a lower base. Conditional on working, girls work longer hours than boys, but receive lower earnings and hourly wages, despite girls having superior prior educational performance and being more positively selected into employment by the measures discussed above. This is consistent with boys and girls being active in distinct labour markets with different demand-side factors and human capital implications (Kooreman, 2009; Erdogan et al., 2012). For this reason we estimate our models separately by sex.

The left-hand side of Table 2 shows transition probabilities, or the proportion of sample members in employment at age 15 conditional on their initial age 14 employment status. The figures suggest some persistence in employment. Around 62% of those in employment in one year are still in employment the following year, which is more than double the population employment rate. Around 15% of the larger group of those *not* in employment in one year have entered employment by the following year, or half of the population employment rate. The right-

²The point score system is explained here: http://www.education.gov.uk/schools/performance/secondary_11/PointsScoreAllocation2011.pdf, p.3. (accessed 4th March 2014)

Table 1: Descriptive statistics by gender.

	Wave 1 (age 14)		Wave 2 (age 15)	
	Boys	Girls	Boys	Girls
Employed (%)	24.9%	18.9%	28.3%	27.1%
Mean age 16 exam score ¹	281.13 (3.04)	308.25 (2.67)	283.62 (3.11)	309.51 (2.82)
Observations	7116	7250	6258	6303

By employment status	Wave 1 (age 14)				Wave 2 (age 15)			
	Boys		Girls		Boys		Girls	
Employed:	Yes	No	Yes	No	Yes	No	Yes	No
Mean hours employment ²	4.14 (0.10)	.	4.24 (0.11)	.	5.16 (0.13)	.	5.38 (0.10)	.
Mean earnings ²	£14.52 (0.41)	.	£14.24 (0.39)	.	£20.58 (0.52)	.	£18.90 (0.33)	.
<i>Family background:</i>								
Household income percentile	0.584 (0.007)	0.534 (0.005)	0.589 (0.009)	0.537 (0.004)	0.584 (0.007)	0.519 (0.005)	0.595 (0.007)	0.518 (0.005)
Higher Educated parent	17.0%	17.6%	18.7%	16.6%	16.1%	17.3%	19.0%	15.7%
Lone Parent	18.1%	23.7%	20.6%	23.2%	17.0%	24.0%	18.3%	24.4%
<i>Educational performance:</i>								
Standardized age 14 exam score ³	0.097*** (0.032)	-0.101 (0.034)	0.293*** (0.034)	0.019 (0.030)	0.096*** (0.033)	-0.090 (0.036)	0.302*** (0.031)	-0.009 (0.018)
Mean age 16 exam score ¹	365.37*** (4.17)	342.04 (2.61)	409.87*** (4.34)	379.75 (2.41)	366.42*** (4.26)	342.89 (2.89)	414.65*** (3.73)	376.89 (2.83)

Notes: Standard errors in parentheses. Population means and proportions calculated using final probability weights. Standard errors clustered by school. ¹: GCSE total point score. ²: Hours of employment and earnings are per week. ³: Key Stage 3 Average Point Score, standardized by subtracting mean and dividing by standard deviation. ***: p-value for difference in mean educational performance between those in and out of employment <0.01.

Table 2: Employment transition probabilities and sequences

	Transition probabilities				Employment sequences (percent)				
	Boys		Girls		Sequence:	00	10	01	11
	W1	W2	W1	W2					
No	0.854	0.146	0.845	0.155	Boys	62.1	9.0	12.6	16.2
Yes	0.370	0.630	0.384	0.616	Girls	65.1	7.3	15.1	12.3

Notes: **Transition probabilities** are probability of being in employment in future wave, conditional on status in current wave, using unweighted data. W1 = Wave 1, W2 = Wave 2. First digit in **sequences** refers to employment at age 14, second to age 15. e.g. 00 = Never worked. 11 = Always worked. Longitudinal probability weights are applied.

hand side of Table 2 shows the proportion of sample members undertaking each sequence of employment decisions over the two years. 36% of students have a job at some point, and 22% undergo a transition between these two years. Thus, we have the data to identify distinct effects of employment at different points in the educational process.

3 Theoretical and empirical models

We motivate our empirical specification by presenting a theoretical model of selection into employment and the effect of employment on academic performance. This model shows firstly that it is necessary to allow for distinct effects of employment at different points in the education process; secondly that employment will crowd out time devoted to study, which has a human capital payoff; and thirdly that the net causal effect (policy effect) of employment will capture both the direct human capital effect and the opportunity cost in terms of reduced investment in other activities, and so is likely to differ from the ‘production function parameter’.

We assume an individual who aims to maximise his expected utility within each time period $t = 1 \dots T$. Here $t = 1$ is the first time period in which the child is able to take in-school employment. Utility is a function of independent consumption ($w_t L_t$, where w_t is his hourly wage and L_t his hours of employment) and leisure ($(1 - L_t - S_t)$, where S_t is time devoted to study) in the current period; and expected educational performance in period T , at the end of compulsory education ($E[Y_T]$, where Y_T is a stochastic function of the individual’s human capital). A vector of socio-economic characteristics which we assume to be time-invariant (\mathbf{X}) are expected to affect individuals’ relative preferences over these elements.

We assume that the utility function is separable into the present ($f_t(\cdot)$) and future ($g(\cdot)$) oriented components, such that the individual’s maximisation problem can be written, for $t = 1 \dots T$:

$$\max_{L_t, S_t} U_t = f_t(\mathbf{L}, \mathbf{S}, \mathbf{X}) + E[g(Y_T(\mathbf{L}, \mathbf{S}, \mathbf{X}, \epsilon, v_T))] \quad (1)$$

The vectors \mathbf{L} and \mathbf{S} include current and all lagged or time varying values of the variable. We assume that for each t , $f_t(\cdot)$ is a concave function in current consumption and leisure, which are complements. The individual has a baseline human capital endowment μ_0 , which is a reduced form function of individual and household characteristics: $\mu_0 = \mu_0(\mathbf{X})$. Human capital evolves each period as a function of further investments in study and labour supply, as well as household circumstances, according to the function $h_t(\cdot)$ and a multiplicative error term $\epsilon_t \sim \ln.\mathcal{N}(0, \sigma_t^2)$ with a log-normal distribution such that $E[\ln.\epsilon_t] = 0$:

$$\mu_t = \mu_{t-1} \cdot h_t(L_t, S_t, \mathbf{X}) \cdot \epsilon_t \quad (2)$$

We assume there is a positive human capital effect of study at all times, so $\frac{\partial h_t}{\partial S_t} > 0$. We offer no

prior about the marginal human capital product of part-time work ($\frac{\partial h_t}{\partial L_t}$). Final human capital at the end of compulsory education can then be written:

$$\mu_T = \mu_0(\mathbf{X}) \cdot \prod_{t=1}^T [h_t(L_t, S_t, \mathbf{X}) \cdot \epsilon_t] \quad (3)$$

Educational performance is the function $k(\cdot)$ of final human capital and a multiplicative error term $v_T \sim \ln \mathcal{N}(0, \sigma^v)$, with a log-normal distribution such that $E[\ln v_T] = 0$:

$$Y_T = k(\mu_T \cdot v_T) \quad (4)$$

We assume that the overall effect of diminishing human capital returns to \mathbf{L} and \mathbf{S} and diminishing marginal utility of GCSE performance is to ensure the function-of-functions $g(Y(\cdot))$ is concave in \mathbf{L} and \mathbf{S} . With $g(Y(\cdot))$ concave and the evolution of human capital over time uncertain, this model predicts that the individual will underinvest in study relative to the case where the human capital outcome is deterministic. This uncertainty becomes progressively smaller as t approaches T , as the outcomes of the random draws ϵ_t in earlier periods become known. See Appendix A.1.1 for a graphical illustration of this prediction.

Assuming that $k(\cdot) = \ln(\cdot)$ and $h_t(\cdot) = e^{m(L_t, S_t, \mathbf{X})}$, with $m(\cdot)$ and $\mu(\mathbf{X})$ as linear functions, the education production function for performance at the end of compulsory education can be written and estimated in the following form. This follows Cunha and Heckman's (2007, 2009) approach to the "technology" of human capital formation in making it explicit that the *timing* of both employment and study matter for academic performance, not just their total amounts. This is a "cumulative model":

$$Y_T = \sum_{t=1}^T [\pi_t^L L_t + \pi_t^S S_t + \ln \epsilon_t] + \beta \mathbf{X} + \ln v_T \quad (5)$$

The error terms ϵ_t and v_T are log-normally distributed so the composite error term in equation (5) ($\sum_{t=1}^T [\ln \epsilon_t] + \ln v_T$) is normally distributed with mean zero.

There are three barriers to obtaining unbiased estimates of the production function parameters π_t^L and π_t^S . Firstly, study is not observed. To address this, we use a vector of latent variable representations of study and effort during school hours, in place of each S_t in equation (5). Secondly, even assuming that the correct vector of inputs is observed perfectly, we would expect

some unobservable characteristic entering $\sum_{t=1}^T [\ln.\epsilon_t] + \ln.v_T$, to positively determine selection into both employment and study. This means that $E[\sum_{t=1}^T [\ln.\epsilon_t] + \ln.v_T | \mathbf{L}, \mathbf{S}, \mathbf{X}] \neq 0$, and we will obtain positively biased estimates of π_t^L and π_t^S . Thirdly, in practice there may be measurement error in teenagers' hours of work and other investment activities, which will attenuate the OLS coefficient on these variables (Tyler, 2003). In section 6 we address these problems by estimating an education production function using instrumental variables methods and repeated measures of endogenous inputs, and discuss the associated identification issues.

Policymakers will also be interested in the partial or policy effect of employment on academic performance, including its indirect effects via inputs crowded out or facilitated by employment. The policy effect is equal to ϕ_t^L in the reduced form specification shown in equation (6):

$$Y_T = \sum_{t=1}^T [\phi_t^L L_t] + \psi \mathbf{X} + \zeta \quad (6)$$

In equation (6), ϕ_t^L may be biased by the omission of exogenous unobserved characteristics. We adopt an instrumental variables strategy to control for these. We thus obtain the overall effect of an exogenous change in employment caused by changes in labour market opportunities (\mathbf{Z}) which, conditional on other observed characteristics (\mathbf{X}), are orthogonal to educational performance. This instrumental variables specification is shown in equation (7):

$$E[Y_t | \mathbf{Z}, \mathbf{X}] = \sum_{t=1}^T [\phi_t^L E[L_t | \mathbf{Z}, \mathbf{X}]] + \psi \mathbf{X} \quad (7)$$

In equation (7), the policy effect ϕ_t^L is distinct from the production function parameter π_t^L , due to the omission of the time varying endogenous input. ϕ_t^L may thus vary over time due to (i) changes in the production function parameter π_t^L , (ii) changes in the rate at which study is crowded out by labour supply, and (iii) changes in the value of that marginal study time, π_t^S .

Our model predicts a strictly negative tradeoff between employment and study. The proof of this prediction is shown in Appendix A.1.2 (page 48). Provided that the production function parameter on study, π_t^S is non-negative, the 'partial effect' of employment will always represent a lower bound of the production function parameter ($\phi_t^L \leq \pi_t^L$). This prediction continues to hold if the specification is relaxed to permit employment and study to have cross-effects in production, i.e. to become complements (in which case the crowd-out will be relatively small) or substitutes. In Browning and Heinesen (2014), the direction of the behavioural response to

an exogenous change in one input depends on whether they are substitutes or complements as well as the cost of the activity to the individual. In our model, part-time work and study time at home are mutually exclusive time inputs, and the crowd-out is driven by the individual's preference for the third time-use category, leisure.

4 Instrumental variables

We use Full Information Maximum Likelihood (FIML) to estimate the policy effects represented in equation (7) and the production function represented in equation (5) adapted to account for endogenous inputs. These entail first-stage tobit equations for hours of employment per week, and a final-stage linear regression for GCSE performance. (Accounting for the censoring of hours of employment at zero in this way will improve the precision of our second stage estimates over those that would be obtained from a linear IV specification, without affecting its bias or consistency properties - see Appendix A.2). Estimating the production function requires a measurement model for the latent factors and an intermediate-stage linear regression equation determining each factor as a function of labour supply. The policy effects, rates of crowd-out of study time and effort, and the production function parameters are all identified through instrumental variable exclusion restrictions in the later stage equations. Specifically, we exclude (i) the child's month-of-birth within the academic cohort, (ii) the age 18-24 claimant count unemployment rate in the month of interview, and (iii) the stock of enterprises in the retail and wholesale (G) and hotel and restaurant (H) sectors of the Standard Industrial Classification (SIC) which are registered for Value Added Tax (VAT). Both the latter instruments vary over time and are measured in the Local Authority District (LAD) of residence.

4.1 Month of birth

The academic year in the UK runs from 1st September to 31st August, and each academic cohort consists of children born within the same September-August period. Month of birth will directly affect teenagers' working hours because those born earlier in the academic year will be allowed to work longer hours or in specific job types earlier than their younger peers. This means they are better placed to fill suitable vacancies which arise in September and October as older cohorts leave for university, and November and December as staff are taken on to meet pre-Christmas demand. This identifying assumption is supported by the data shown in

Appendix A.3, in which we show the seasonal cycle of vacancies in the two industrial sectors employing the greatest concentration of teenagers in the UK.

In our empirical specification, month of birth is a linear term set equal to 1 for September births, up to 12 for August births. We also include dummy variables for the month-of-interview and for interview in or after the month of the child’s birthday in all equations. This means that spurious variation in employment generated by being interviewed later in the fieldwork schedule will not contribute to identifying the causal effect of employment on academic performance.

It is well established that those born earlier in the school year retain an advantage in educational outcomes (Plug, 2001; Crawford et al., 2013 and references therein), presenting a threat to the validity of month of birth as an instrument. There are four potential mechanisms through which this effect could be causal: (i) Age-at-test: those born earlier in the school year are older when taking a test on a common date; (ii) Age of starting school: those born later in the school year are younger on a common school entry date, and so less ready to learn; (iii) length-of-schooling: if those born later in the school year also start school later in a common academic year, they will have received less schooling by the time of a test; and (iv) relative age: institutional practices such as sorting into ‘sets’ or ‘tracks’ by observed performance at a given time will compound these initial differences as those born earlier in the school year are more likely to reach the higher tracks.

Crawford et al’s (2013) analysis suggests that age-at-test is the dominant mechanism. Using data on English schoolchildren at the same age as our LSYPE sample they find that the performance gap between the oldest and youngest pupils in each cohort becomes progressively smaller with age as the relative age gap between September and August birthdays shrinks.³ These converging performances mean it is insufficient to control for prior performance at only one point in time. In this paper we shall control for the trajectory of academic performance through absolute performance at both ages 11 and 14, and assume that this leaves no residual direct effect of month of birth on academic performance.

³e.g. In June of school year 11, when most cohort members are 16, a cohort member born on 1st September has lived 6.3% longer than one born on 31st August, compared with 7.3% in school year 9 (age 14) and 9.3% in school year 6 (age 11). For the performance measure “meeting the expected standard” (an absolute measure), they show (Crawford et al., 2013, figures 3.1-3.2, pp.20-21) 26, 13, 7.8, and 6.4 percentage point gaps between August and September birthdays at ages 7, 11, 14 and 16 respectively.

4.2 Youth unemployment rate

The local authority district (LAD) age 18-24 unemployment rate is a direct measure of labour market opportunities over areas with an average population of 164,000 and size of 155 square miles. This is equivalent to 1.6 times the population but 12.5% of the average area of a US county. Kalenkoski and Pabilonia (2010) and Rothstein (2007) used the county and state *adult* unemployment rates, respectively. Our measure is considerably more localized, which we consider appropriate for these teenagers, who will be reliant on parental or public transport, or cycle-or-walkable distances. (The youngest age at which teenagers may learn to drive in England is 17). Moreover, the age restriction in our unemployment measure should increase precision, since children predominantly work in the same sectors as youths, particularly hotels, catering, wholesale and retail (Mizen et al., 1999; Hibbett and Beatson, 1995). This motivates our third and final instrument.

4.3 Stock of VAT-registered enterprises in retail, wholesale, hotels and restaurants

We obtained the number of enterprises registered for Value Added Tax (VAT, a national sales or consumption tax on most ‘non-essential’ goods and services) in each Local Authority District at the end of each calendar year, as published by the Department for Business, Enterprise and Regulatory Reform (BERR - superceded in 2010 by the Department for Business, Innovation and Skills) as the official record of business start-ups and closures (see Office for National Statistics, 2008). We divide the number of these businesses in sectors G (“Wholesale and retail trade, repair of motor vehicles, motorcycles and personal and household goods”) and H (“Hotels and restaurants”) according the 2003 Standard Industrial Classification by the mid-year population of 18-24 year-old youths, to define a measure of the density of the kinds of employers most frequently employing schoolchildren. For ease of interpretation, all descriptive statistics and regression coefficients are measured per 100 students.

The LSYPE does not provide information on the types of jobs teenagers do, but it is well documented from numerous targeted surveys that the employment of individuals in full-time education in the UK is concentrated in these sectors. Curtis and Lucas (2001) and Mizen, Bolton and Pole (1999) provide extensive reviews supporting this fact. For the precise period considered in this paper, Howieson et al. (2006) document the situation in Scotland, where the education

system, labour market conditions and legislation are very similar to England. They show the largest category of employment among those aged 14 and under is newspaper delivery (32% of workers), the majority of which will be contracted by corner shops and convenience stores. The largest categories of employment among 15 and 16 year olds are “cafés and restaurants” (17.5%), “other shops” (14%) and “chain stores” (12%).

4.3.1 Time variation in unemployment rate and VAT registrations

Table 3: Within-and-between variation in time-varying instruments

Age 18-24 unemployment rate						
Statistic:	Mean	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
Boys						
Wave 2 level (Age 15)	0.0444	0.0182	0.0280	0.0416	0.0581	0.0727
Wave 1 level (Age 14)	0.0406	0.0152	0.0243	0.0394	0.0529	0.0693
Change	0.0038	-0.0012	0.0010	0.0032	0.0062	0.0087
Absolute change	0.0045	0.0008	0.0019	0.0036	0.0063	0.0086
Girls						
Wave 2 level (Age 15)	0.0448	0.0182	0.0290	0.0408	0.0595	0.0741
Wave 1 level (Age 14)	0.0410	0.0163	0.0243	0.0394	0.0546	0.0697
Change	0.0038	-0.0012	0.0014	0.0034	0.0060	0.0087
Absolute change	0.0046	0.0009	0.0022	0.0037	0.0063	0.0087

‘G’ and ‘H’ VAT registrations per 100 18-24 year-olds						
Statistic:	Mean	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
Boys						
Wave 2 level (Age 15)	10.26	5.85	6.85	9.06	12.75	16.03
Wave 1 level (Age 14)	10.27	6.04	6.98	9.13	12.68	15.96
Change	-0.0143	-0.2565	-0.1207	-0.0183	0.1209	0.2609
Absolute change	0.1803	0.0254	0.0610	0.1209	0.2393	0.3544
Girls						
Wave 2 level (Age 15)	10.33	5.90	6.98	9.23	12.68	16.32
Wave 1 level (Age 14)	10.32	5.85	6.85	9.17	12.83	16.47
Change	-0.0060	-0.2541	-0.1193	-0.0095	0.1366	0.2611
Absolute change	0.1853	0.0272	0.0622	0.1281	0.2411	0.3532

Notes: Sample: 5122 boys, 5066 girls. ‘pc’ = ‘percentile’. Calculated using unweighted data.

Table 3 shows the within and between-individual variation in both our time-varying instrumental variables. These tables confirm that local labour market conditions, and changes in them, are

essentially orthogonal to gender. On average teenagers in our samples face a local youth (18-24) unemployment rate of approximately 4.25%, and there are just over 10 VAT-registered enterprises in the wholesale, retail, hotel and restaurant sectors for every 100 youths in this age bracket. There is significant between-variation in both instruments at both points in time, with those at the 75th percentile facing local youth unemployment that is twice as high, or approximately twice the density of suitable VAT registered employers, as those on the 25th percentile, for example.

There is also a substantial degree of variation in both measures over time within individuals. Unemployment rates and VAT-registered employer numbers worsened over time for 82.5% and 52% of our sample respectively. Just over one-third of the sample experienced an increase or decrease in the local unemployment rate of more than 0.5 percentage points, and one-quarter of the sample experienced an increase or decrease in VAT registered employers of more than 0.25 per 100 youths.

4.3.2 Threats to validity of unemployment rate and VAT-registration instruments

Identification of the distinct policy effects and production function parameters on employment at age 14 and at age 15 rests on differential time variation within individuals in both the youth unemployment rate and density of suitable employers, but both parameters are also identified through cross-sectional variation in these indicators at each single point in time. These instruments are indicators for the health of the local labour market, and may capture permanent differences in the relative affluence of the districts. Parents will endogenously have selected into their area of residence, based partly on their preference for school quality (subject to income and credit constraints), for example, and as such we may expect that children with more opportunities to work have parents with a higher preference for education and attend higher quality schools.

For this reason we condition on individual household circumstances, specifically including parental income, employment, health, education and qualification levels. To further exclude unobservable influences expected to be correlated with youth labour market opportunities and educational outcomes, such as within-locality peer effects in either the parents' or child's preferences, or the child's friendship group or leisure opportunities, we also control for the index of multiple deprivation (IMD). This captures the average levels of income, employment, health,

educational qualifications, and skills and training across a localized population of approximately 800 households, as well as indicators for barriers to housing and services; crime; and the living environment, at a specific point in time.⁴ Conditional on these, and the child's prior educational performance, we assume this area-level labour market condition has no residual direct effect on the student's educational performance, except via the teenager's hours of labour supply.

In relation to both the unemployment rate and employer density, we also assume that while the effect of labour supply on educational performance or other activities may be heterogeneous, these instruments are uncorrelated with the 'quality' of employment, and that in estimation we identify the *average* treatment effect or production function parameter.

An alternative threat to instrument validity common to both the youth unemployment rate and stock of VAT-registered enterprises in these industries, is the possibility that since the industrial structure of a local area and the youth unemployment rate are serially correlated, teenagers may adjust their behaviour at school in response to their expected employment opportunities should they leave at the end of compulsory schooling.

One mechanism through which this may occur is by selecting into taking relevant qualifications for major employers in the area, such as a vocational programs in food technology or catering, in an area with a large hospitality sector (that also employs many schoolchildren). If it is systematically easier to attain a given point score in such a qualification, our estimated effect of in-school employment on school performance will be positively biased. Point scores are allocated according to the 'size' of the qualification (the number of 'guided learning hours' deemed necessary to teach the qualification), and the 'challenge' (the difficulty of meeting the criteria for a given grade under the structure being used), and are therefore designed to capture equivalent levels of achievement (Department for Education, 2012). However, in section 7.2 we assess the evidence for whether students change their qualification choices in response to local labour market opportunities and find none.

In addition, schoolchildren with highly present-biased preferences may treat the presence of many low-skilled labour market opportunities as creating high opportunity costs of further education and low return to higher grades. This may cause them to reduce their effort, in which case our estimated effect of in-school employment on school performance, study time and attitude to schoolwork will all be negatively biased. In section 7.1 we establish bounds which

⁴We use the 2004 index, since this is the latest indicator which is predetermined to our explanatory and outcome variables.

show our conclusions regarding the sign of the effects of in-school employment on educational performance to be robust to relaxing the assumption of zero direct effect of these instruments.

4.4 Additional controls

We control for prior educational performance at age 11 and 14, household income (excluding the child’s income), home ownership, receipt of disability benefits, parents’ employment status, education and socio-economic class, household structure and non-resident siblings, child’s ethnicity, special educational needs (SEN) classification, urban-rural classification and the index of multiple deprivation (IMD) in the Lower Super Output Area (LSOA - containing 400-1200 households in an average area of 1.45 square miles⁵), and living in Greater London. Consistent with equation (7), we treat all covariates as time invariant. Several are in fact time-varying (specifically parents’ employment status, ill-health, education, and socio-economic classification; housing tenure; lone-parent family indicator; and resident and non-resident siblings), but any changes involve very few households, meaning they are almost perfectly collinear between waves. We use the age 14 observation of each covariate, to best account for initial conditions and avoid endogeneity issues. Parental income is elicited using different questions in each wave of the LSYPE, so we construct a measure of the household’s relative ‘permanent’ income over the first three waves.

5 Policy effects

5.1 The effect of employment on academic performance

Estimates of the empirical counterpart to the ‘policy effect’ equations (6) and (7), together with diagnostic tests for the strength of our instrumental variables (individually and as a battery) are shown in Table 4. As with our production function models, we assume investments take place over two periods, so $T = 2$. Results for boys show insignificant positive and negative coefficients on employment at ages 14 (wave 1) and 15 (wave 2) respectively, in both the OLS and instrumental variables specifications. For girls, the instrumental variables results show a trivial insignificant negative effect at age 14 but a very large negative effect at age 15, equivalent to 4.2% of a standard deviation in the total point score per hour of employment per week. For

⁵(Office of the Deputy Prime Minister, 2004)

English audiences, this is equivalent to one GCSE grade in one subject. This effect size is sufficiently large to expect a typical working week (5-6 hours employment) to cancel out the raw advantage in educational performance seen for girls at this age.

Our significant negative instrumental variables estimate compares with the small but precisely estimated coefficient of 0.6% of a standard deviation per hour per week from the OLS results. The difference is consistent with the unobserved characteristics driving selection into employment positively driving educational performance.

Table 4: ‘Policy effects’ of employment on GCSE performance for $T = 2$

	Boys				Girls			
	Instrumental variables		OLS		Instrumental variables		OLS	
	W1 Work Hours	W2 Work Hours	Standardized GCSE score	Standardized GCSE score	W1 Work Hours	W2 Work Hours	Standardized GCSE score	Standardized GCSE score
W2 Work Hours	.	.	-0.009 (0.012)	-0.004 (0.004)	.	.	-0.042** (0.021)	-0.006* (0.003)
W1 Work Hours	.	.	0.012 (0.015)	0.001 (0.005)	.	.	-0.002 (0.023)	-0.007 (0.005)
W2 Age 18-24 LAD unemployment rate	.	-0.866 (11.475)	.	.	.	-34.684*** (11.471)	.	.
W1 Age 18-24 LAD unemployment rate	1.290 (10.095)	.	.	.	-10.535 (12.408)	.	.	.
W2 SIC ‘G’ and ‘H’ VAT reg’ per 100 youths	.	0.191*** (0.046)	.	.	.	0.125*** (0.041)	.	.
W1 SIC ‘G’ and ‘H’ VAT reg’ per 100 youths	0.093** (0.041)	.	.	.	0.116*** (0.042)	.	.	.
Month of birth	-0.068 (0.050)	-0.031 (0.055)	.	.	-0.159*** (0.060)	-0.097* (0.051)	.	.
Age 14 ave’ point score (standardized)	.	-0.005 (0.403)	0.783*** (0.026)	0.784*** (0.027)	.	-0.573 (0.358)	0.723*** (0.027)	0.729*** (0.026)
Age 11 ave’ point score (standardized)	0.262 (0.187)	0.043 (0.418)	-0.049** (0.024)	-0.048* (0.024)	0.534*** (0.201)	1.477*** (0.382)	0.009 (0.025)	-0.004 (0.024)
<i>Individual instrument relevance</i>								
F-stat (18-24 unemp’ rate)	0.02	0.01	.	.	0.72	9.14	.	.
<i>Bonferroni-adj’ p-value</i>	(1.000)	(1.000)	.	.	(1.000)	(0.008)	.	.
F-stat (‘G’/‘H’ VAT reg’)	5.04	17.31	.	.	7.69	9.51	.	.
<i>Bonferroni-adj’ p-value</i>	(0.074)	(0.000)	.	.	(0.017)	(0.006)	.	.
F-stat (Month of birth)	1.87	0.32	.	.	6.99	3.66	.	.
<i>Bonferroni-adj’ p-value</i>	(0.515)	(1.000)	.	.	(0.025)	(0.167)	.	.
<i>Joint instrument relevance</i>								
F-stat (All instruments)	7.33	21.17	.	.	17.57	30.80	.	.
<i>p-value</i>	(0.062)	(0.000)	.	.	(0.001)	(0.000)	.	.
Observations	5122				5066			

Notes: Joint significance statistics and p-values are Wald test statistics, where indicated with Bonferroni adjustment for testing multiple hypotheses within each equation. Standard errors, clustered by school, in parentheses. Longitudinal weights applied. *: $p < 0.1$; **: $p < 0.05$; ***: $p < 0.01$. GCSE (age 16), age 14 and age 11 average point scores. Month-of-birth within academic year: Sept (oldest in year) = 1, Aug (youngest in year) = 12. **Additional controls:** Parent’s socio-economic status, parents’ employment, index of multiple deprivation (IMD), resident and non-resident siblings, lone parent family, child’s special educational needs (SEN) classification, urban-rural classification, child’s ethnicity, timing of interview.

5.2 The effect of the instrumental variables on employment

Being born later in the academic year significantly reduces hours of employment only for girls. Those born in August are predicted to work 1.75 hours less per week in wave 1, and 1.1 hours less per week in wave 2, than their September-born peers. The effects of this instrument are consistent with turning age 14 expanding employment opportunities more than turning 15, and also with boys' employment being concentrated in informal positions (such as newspaper delivery) where age restrictions are not binding, while girls' employment is concentrated in formally regulated roles such as serving in shops (Howieson et al., 2006) or with positions of responsibility such as babysitting (Kooreman, 2009).

The coefficient on the local youth unemployment rate is statistically significant only for girls in wave 2. At this point a 1 percentage point increase in age 18-24 unemployment reduces girls' predicted employment by 21 minutes (i.e. 0.35 hours) per week. The corresponding wave 1 figure is approximately one-third of this magnitude, while for boys the coefficient on the unemployment rate is quantitatively and statistically indistinguishable from zero at both points in time. Similarly in line with the results for month of birth, this gender distinction is consistent with boys' employment being more differentiated and hence insulated from adult labour market conditions.

The density of retail and wholesale (sector G) and hotel and restaurant (sector H) VAT registered enterprises is shown to be a quantitatively and statistically significant determinant of hours of employment for both boys and girls at both points in time. An additional 10 enterprises per 100 18-24 year-olds in the LAD results in between one and two additional hours of employment per week.

For both sexes at both points in time, it is the density of G and H sector VAT-registered enterprises that contributes the most identifying variation in hours of employment among our instruments, as measured by the F-statistic for its exclusion from the first stage. The F-statistic for the exclusion of all three instruments from the first stage is always statistically significant at the 0.1% level for girls, and also for boys in wave 2, but only at the 10% level for boys in wave 1. In addition, while for girls in both waves there two of the instruments are statistically significant and have an individual F-statistic close to 10, for boys only the G and H sector density is ever significant, and in wave 1 only weakly so.

The relatively weak relationship between the instruments and the endogenous explanatory vari-

able for boys will exacerbate any inconsistency in our instrumental variable estimates caused by a residual direct effect of any of the instruments on our outcome variable (Bound, Jaeger and Baker, 1995). The ‘imperfect instrumental variable’ exercise in section 7.1, is therefore a valuable robustness check, to show both the sensitivity of our estimates to using one instrument at a time (these being of substantially varying strengths), and to derive bounds for the true effect of in-school employment on academic performance under the assumption of a non-zero direct effect.

6 Production function parameters

In this section, we estimate the education production function. This means we obtain estimates of the effect of part-time employment and two other inputs to the education production function, holding the other investments constant. We also provide insight into the mechanisms through which the partial effect of part-time employment estimated above occurs, by showing how an exogenous increase in employment crowds out or facilitates other activities.

6.1 Latent inputs

We propose two inputs to the education production function: ‘Study’ (study time at home) and ‘Attitude’ (effort during school hours). We have measures for both of these inputs at time of interview in waves 1 (age 14) and 2 (age 15) of the LSYPE. To assess the effect of employment on alternative competing time uses, we also define factors for ‘Active Leisure’, ‘Social Life’ and ‘Risky Behaviours’, all of which we assume have no direct effect in the education production function, except via their effects on the quality or quantity of study time.

To be specific, our analysis relates to n individuals i at times $t = 1 \dots T$. In our empirical application we assume that $T = 2$. Each sampled individual at time t is characterised by:

- Q dimensions of ‘true’ activities $\mathbf{s}_{it} = (s_{it}^1 \dots s_{it}^Q)$. Each is measured by the discrete indicators $S_{ijt}^q, j = 1 \dots J^q$.
- Desired hours of employment l_{it} , with actual hours of employment $L_{it} \geq 0$.
- A vector of time-invariant socio-economic characteristics, \mathbf{X}_i .
- A vector of instruments which determine selection into employment, \mathbf{Z}_{it} .

- A continuous measure of educational performance in the initial (Y_{i1}) and final (Y_{iT}) periods.

We assume an ordinal quasi-linear structure for measurement of the activities:

$$\tilde{S}_{ijt}^q = \alpha_{jt}^q + \lambda_{jt}^q s_{it}^q + \epsilon_{ijt}^q \quad \text{for } j = 1..J \quad (8)$$

$$S_{ijt}^q = p \quad \text{iff } W_{jp-1t}^q \leq \tilde{S}_{ijt}^q < W_{jpt}^q \quad \text{for } p = 1..P_j^q \quad (9)$$

Here, λ_{jt}^q is the factor loading relating the observed indicator j to factor s_{it}^q . W^q s are threshold parameters. P_j^q is the number of response categories for indicator S_{ij}^q , and ϵ_{ijt}^q is the logistically distributed random response error. This means the observable indicator S_{ij}^q is linked to its unobserved continuous form \tilde{S}_{ij}^q by an ordered logit function.⁶

Table 5 defines the measures and shows the factor loadings and ordered logit cut points for each latent activity. The factor loading for the first measure is always constrained to one. The variance of each factor is shown in the bottom section. The measurement models are estimated using the pooled sample of both sexes and both waves. This ensures that a given change in the principal component of the factor corresponds to an equivalent change in absolute activity levels across groups and over time. This means that the level, crowd-out by employment, and productivity of each factor is expressed in a common unit. An interpretation of the change in activity or attitude associated with a one standard deviation increase in each factor is given in Table 6. In Appendix A.4 (page 56) we describe the distribution of the inputs between waves and sexes, to show that although significant differences exist, there is substantial overlap, so the adjustments described in Table 6 would represent a plausible and relevant change in all cases.

In our descriptive statistics (Table 7) and subsequent output each factor is constrained to have mean zero and standard deviation 1, and is treated as observed (the measurement model is not estimated simultaneously). We assume there to be no heterogeneity in the ‘quality’ of each factor and that the measurement error in these inputs is uncorrelated with the exogenous component of labour supply.

Table 7 shows no statistically significant differences in the actual study effort undertaken by those with and without a job but one exception to this for attitude to schoolwork. Those in

⁶The model and notation outlined here follows closely the structure used for a model of disability status and benefit receipt by Hancock et al. (2013).

Table 5: Factor loadings and cut points in ordered logit measurement models.

Factor loadings:									
Attitude		Study		Active Leisure		Social Life		Risky Behaviour	
Work hard	1	Not truant	1	Play sport	1	Nightclub	1	Smoke	1
Worth it	0.832 (0.020)	Homework	1.453 (0.080)	School sport	0.687 (0.025)	Pub or bar	0.700 (0.046)	Alcohol	0.594 (0.022)
Work waste	0.954 (0.025)	Outside	2.500 (0.251)	Sport in gen	0.903 (0.027)	Concert	0.336 (0.018)	Cannabis	1.347 (0.074)
Interest	1.228 (0.027)	Study club	2.563 (0.249)	School clubs	0.329 (0.015)	Arcade	0.647 (0.046)		
				Play music	0.186 (-0.257)	Friends out	0.442 (0.027)		
Cut Points:									
Attitude		Study		Active Leisure		Social Life		Risky Behaviour	
Work hard		Not truant		Play sport		Nightclub		Smoke	
1	-5.028	1	-1.511	1	-0.257	1	1.303	1	3.422
2	-2.048			School sport					
3	-1.992			1	0.250				
4	-1.898			2	0.917				
5	1.404			3	3.020	Pub or bar		Alcohol	
6	1.467			4	4.684	1	2.222	1	0.126
Worth it		Homework		Sport in gen				2	0.882
1	-3.888	1	-2.101	1	-3.683			3	0.927
2	-2.996	2	-1.376	2	-2.596			4	0.971
3	0.197	3	0.002	3	-2.010	Concert		5	1.578
4	0.053	4	1.183	4	-0.699	1	0.045	6	2.118
5	0.078	5	2.530	5	-0.695			7	3.260
6	0.103			6	-0.691			8	5.797
Work waste		Outside		School clubs					
1	-4.642	1	0.497	1	0.553	Arcade		Cannabis	
2	-3.07	2	1.620	2	1.059	1	2.014	1	4.418
3	0.584	3	3.827	3	3.068				
4	0.641	4	6.028	4	4.946				
5	0.685								
6	0.740								
Interest		Study club		Play music		Friends out			
1	-5.148	1	1.099	1	1.437	1	-1.390		
2	-2.134	2	2.054			2	0.246		
3	-2.013	3	4.056			3	0.255		
4	-1.854	4	5.626			4	1.509		
5	-1.710								
6	2.953								
Variance:									
Attitude	2.414	Study	0.263	Active Leisure	4.454	Social Life	2.489	Risky	7.032

Notes: Standard errors, clustered by individual, in parentheses. Measures are defined as follows. **Attitude:** ‘Work hard’: “I work as hard as I can in school”. ‘Worth it’: “School work is worth doing”. ‘Work waste’: “The work I do in lessons is a waste of time” (recorded). ‘Interest’: “The work I do in lessons is interesting to me”. **Study:** ‘Not truant’: Student has not truanted in last 12 months (dummy variable). ‘Homework’: Estimate of hours per week spent doing homework, based on reported time spent on a computer doing schoolwork, nights per week doing so, and nights per week doing any homework. (Categories are zero, 1 to under 2 hours, 2 to under 4, 4 to under 6, 6 to under 10, and 10 or more). ‘Outside’: Works towards exams with teachers outside of lessons (frequency on 5-point scale). ‘Study club’: Attends school study clubs (frequency on 5-point scale). **Active Leisure:** ‘Play sport’: Played any kind of sport in the last four weeks (dummy). ‘School sport’: Frequency of using school sports facilities (5-point scale). ‘Sport in gen’: Frequency of doing sport (5-point scale). ‘School clubs’: Frequency of participation in school clubs or societies. (5-point scale). ‘Play music’: Played a musical instrument in the last four weeks. **Social Life:** ‘Nightclub’: Gone to a party, dance, nightclub or disco in the last four weeks. ‘Pub or bar’: Gone to a pub or bar in the last four weeks. ‘Concert’: Gone to a cinema, theatre or concert in the last four weeks. ‘Arcade’: Gone to an amusement arcade in the last four weeks. ‘Friends out’: How many times gone out with friends in last seven days. **Risky behaviours:** ‘Smoke’: Whether cohort member “ever smokes” (dummy variable). ‘Alcohol’: Frequency of alcohol consumption (6-point scale). ‘Cannabis’: Whether cohort member “ever tried cannabis” (dummy variable). ‘Don’t know’ is coded as the within-wave mean, creating up to three extra categories for the following measures: Work hard (3^{rd} , 4^{th} , 6^{th} categories), Worth it (4^{rd} - 6^{th}), Work waste (4^{th} - 6^{th}), Interest (3^{rd} - 5^{th}), Sport in gen’ (5^{th} - 6^{th}), Friends out (3^{rd}), Alcohol (3^{rd} , 4^{th} , 6^{th}).

employment are significantly more active in leisure and social activities, and risky behaviours. In general, the table also reveals higher levels of Study for girls than boys. This difference is more pronounced in Wave 2, at around 10% of a standard deviation, and indicates that before employment but after study time, girls will start with less disposable leisure time. This increases the risk that employment will crowd out their study time.⁷

⁷Another reason why this may occur is differences in time devoted to ‘household production’ (caring responsibilities or domestic chores) by gender. Appendix A.5 presents some descriptive statistics on this time use by gender and employment status. These are all consistent with girls indeed facing a tighter time constraint, but

Table 6: Interpretations of a one standard deviation change in each latent factor

Input	Measure	Interpretation	Baseline
Attitude			
<i>A 1 s.d. change on average is equivalent to all of:</i>			
Work hard	33.4 p.pt more likely to 'strongly agree' that "I always work hard in school"		26.1% strongly agree
Worth it	29.2 p.pt more likely to 'strongly agree' that "School work is worth doing".		47.3% strongly agree
Work waste	35.4 p.pt more likely to 'strongly disagree' that "The work I do in lessons is a waste of time".		36.6% strongly disagree
Interest	21.4 p.pt more likely to 'strongly agree' that "The work I do in lessons is interesting to me".		12.7% strongly agree
Study			
<i>A 1 s.d. change on average is equivalent to all of:</i>			
Not truant	6.4 p.pt less likely to have truanted in the last year.		19.2% answer yes.
Homework	1 hour additional homework per week.		Median response: ≥ 4 , < 6 hours p.w.
Outside	1 extra visit to work with teacher outside lessons per week.		Median: Never; 75 th %ile: < 1 p.w.
Study club	1 extra visit to school study clubs per week.		Median: Never; 75 th %ile: < 1 p.w.
Active Leisure			
<i>A 1 s.d. change on average is equivalent to all of:</i>			
Play sport	33.4 p.pt more likely to have played any kind of sport in the last four weeks.		53.9% answer yes.
School sport	1.5 additional occasions to use school sport facilities each week.		Median: Never; 75 th %ile: 1-2 p.w.
Sport in gen	Move from "hardly ever" participate in sport to once per week <i>or</i> move from once per week to several but not "most" days per week.		25 th %ile: 1 p.w.; Median: > 1 p.w.; 75 th %ile: "Most days"
School clubs	0.5 additional occasions to participate in other school clubs each week.		Median: Never; 75 th %ile: 1-2 p.w.
Play music	6.8 p.pt more likely to have played a musical instrument in the last four weeks.		19.9% answer yes.
Social Life			
<i>A 1 s.d. change on average is equivalent to all of:</i>			
Nightclub	35.5 p.pt more likely to have gone to a party, dance, nightclub or disco in last four weeks.		28.8% mention.
Pub or bar	14.9 p.pt more likely to have gone to a pub or bar in the last four weeks.		13.8% mention.
Concert	13.0 p.pt more likely to have gone to a cinema, theatre or concert in the last four weeks.		49.0% mention.
Arcade	15.3 p.pt more likely to have gone to an amusement arcade in the last four weeks.		15.5% mention.
Friends out	1 additional trip out with friends in last week.		Median: 1-2 p.w.; 75 th %ile: 3-5 p.w.
Risky Behaviours			
<i>A 1 s.d. change on average is equivalent to all of:</i>			
Smoke	29.8 p.pt more likely to "ever smoke".		14.0% answer yes
Alcohol	34.7 p.pt more likely to ever consume alcohol. (Conditional on some consumption: More than double frequency: bi-monthly to monthly/ monthly to 2-3 times per month/2-3 per month to 1-2 per week).		52.2% never consume alcohol 75 th %ile: once every 2 months; 90 th %ile: 2-3 p.m.
Cannabis	28.8 p.pt more likely to have "ever tried cannabis".		13.3% answer yes

Abbreviations: "p.pt" = "percentage point". "p.w" = "per week". "p.m" = "per month". "%ile" = "percentile". **Note:** Changes expressed as percentage point change in unconditional probability that condition is met due to a uniform one standard deviation change in the latent input across the entire pooled sample population, from the levels observed in the data.

The table further shows higher levels of participation in Social activities and Risky Behaviours for girls, and in Active Leisure for boys.

On inspection, there are also some interesting difference-in-differences by gender. In particular by age 15, boys have a worse 'Attitude' to schoolwork than girls if they are not working (the difference is around 13% of a standard deviation), but markedly better if they are in employment (14% of a standard deviation), mainly due to the much better attitude (25% of a standard deviation) of boys in employment compared with those not. This contrasts with the roughly equal Study (the amount outside of school hours) for those in and out of employment.

we find no suggestion that parents are any less accommodating of part-time work of girls than boys, in terms of their expectations of time to be spent on these activities.

These relationships will substantially be driven by selection. We now evaluate the extent to which these relationships are causal, and the effect of these inputs on educational performance.

Table 7: Descriptive statistics: latent variables by employment status and gender.

	Wave 1				Wave 2			
	Boys		Girls		Boys		Girls	
	Yes	No	Yes	No	Yes	No	Yes	No
Employed:								
Attitude	0.103 (0.032)	0.044 (0.016)	0.098 (0.032)	0.126 (0.016)	-0.029*** (0.025)	-0.281 (0.019)	-0.170 (0.026)	-0.151 (0.019)
Study	-0.140 (0.028)	-0.178 (0.022)	-0.126 (0.029)	-0.124 (0.019)	-0.062 (0.030)	-0.074 (0.024)	0.042 (0.028)	0.015 (0.022)
Active Leisure	0.488*** (0.026)	0.330 (0.019)	-0.026*** (0.033)	-0.302 (0.021)	0.414*** (0.026)	0.226 (0.020)	-0.205*** (0.029)	-0.488 (0.026)
Social Life	0.154*** (0.030)	-0.090 (0.018)	0.310*** (0.031)	0.018 (0.018)	0.321*** (0.028)	0.033 (0.018)	0.463*** (0.028)	0.122 (0.019)
Risky Behaviours	0.053*** (0.025)	-0.170 (0.016)	0.080*** (0.035)	-0.099 (0.016)	0.464*** (0.029)	0.237 (0.020)	0.597*** (0.036)	0.328 (0.023)

Notes: Latent inputs all have mean zero and standard deviation of 1 in the pooled sample of both sexes and both waves. Standard errors in parentheses. Population means calculated using final probability weights. Standard errors clustered by school. ***: p-value for difference in mean latent input between those in and out of employment <0.01.

6.2 Instrumental variables models exploiting tradeoffs in activities

De Fraja et al. (2010) make the foremost contribution in using instrumental variables to estimate an education production function. They consider three endogenous inputs; child, parent and school effort. They motivate the child’s birth weight, father’s social class, and school size as instruments for the respective inputs, which are determined simultaneously by the three optimizing agents.

There are theoretically viable candidates for instruments for each of the inputs assessed in the present paper. Financial incentives targeting students’ self investments have been shown to raise investments in study time. Freyer (2010) shows these to be more effective than targeting *outcomes*, because individuals may not know what investments are effective, or as shown in our theoretical model here, may be deterred by uncertainty about outcomes or the long time horizon. Other papers propose sources of exogenous variation in the opportunity cost of study or value of leisure, such as the weather (Kalenkoski and Pabilonia, 2014) or major football (soccer) tournaments during the exam period (Burgess, Metcalfe and Proud, 2011). Cross-sectional variation in opportunities for structured social activities or entertainment may determine selec-

tion into Social Life, Active Leisure and Risky Behaviours. Barron, Ewing and Waddell (2000) used variables such as the size of high school and ‘library books per student’ for this purpose. The LSYPE collects reports on whether such facilities or clubs are available, but only 12% and 8% of sample members respectively report they are not. With information on social networks, current or lagged peer-group behaviour could also be used (Lundborg, 2006; Clark and Loheac, 2007) though if these signals are observable, parents’ school choices may be endogenous.

We do not have the data to pursue exogenous variation in all the endogenous inputs simultaneously. However, as motivated in our theoretical model, with individuals behaving optimally we predict a strictly negative tradeoff of employment with study time (section 3 and Appendix A.1.2). Therefore we propose a structure in which exogenous variation in employment determines the child’s contemporaneous level of Study (or any other input), but Study does not affect employment. Both labour supply and Study then enter the production function. To identify the production function parameter on Study we require that the principal component of our latent Study indicator is strictly decreasing in labour supply. We do not impose this relationship as a restriction in the econometric framework. Instead we test whether it is borne out empirically, and interpret the coefficients obtained on Study in the appropriate light.

In our $T = 2$ case our instrumental variables are: Month of birth, the youth unemployment rate at $t = 1$ and $t = 2$, and the sectors G and H VAT registration density at $t = 1$ and $t = 2$. Month of birth and the time-specific unemployment rate and VAT registration density all identify the direct effect of employment on educational performance, and the effect of employment on the contemporaneous Study level. Exogenous variation in employment identifies the effect of Study on educational performance. The product of the latter two effects can be interpreted as the indirect effect of employment, or the portion of the overall effect that is mediated by changes in Study. Hence, the model contains employment and a single endogenous latent factor, both measured at two points in time. We describe this as a ‘single-factor multiple-wave’ specification. A model with multiple factors would be underidentified. We apply this structure in turn to our ‘Attitude’ input. We have no expected sign: employment may reduce motivation through general tiredness, or increase it through exposure to the disadvantages of jobs requiring few qualifications (Dustmann and van Soest, 2007).

To be explicit, desired hours of employment are determined by a linear labour supply function

of \mathbf{X}_i , \mathbf{Z}_{it} and prior educational performance:⁸

$$l_{it} = \boldsymbol{\theta}_t^l \mathbf{X}_i + \boldsymbol{\gamma}_t^l \mathbf{Z}_{it} + \delta_t^l Y_{i1} + \xi_{it}^l \quad (10)$$

Hours of employment equal zero if desired hours are negative, and equals desired hours otherwise:

$$L_{it} = l_{it} \quad \text{if } l_{it} \geq 0 \quad (11)$$

$$L_{it} = 0 \quad \text{if } l_{it} < 0$$

Each endogenous input s_{it}^q (including Attitude) is determined by a linear function of socio-economic characteristics, which will have determined parental and self-investments and the child's preferences, but also by hours of employment:

$$s_{it}^q = \boldsymbol{\theta}_t^q \mathbf{X}_i + \boldsymbol{\gamma}_t^q L_{it} + \delta_t^q Y_{i1} + \xi_{it}^q \quad (12)$$

In equations (10) and (12), $\boldsymbol{\theta}$ s represent vectors of coefficients, and $\boldsymbol{\gamma}_t^q$ the crowd-out (or facilitation) of factor s_{it}^q by hours of employment. The residual ξ s capture other unobservable factors. Academic performance is modelled by the following linear regression specification:

$$Y_{iT} = \sum_{t=1}^T [\pi_t^L L_{it} + \pi_t^q s_{it}^q] + \boldsymbol{\beta} \mathbf{X}_i + \delta^Y Y_{i1} + u_i \quad (13)$$

Here, $\boldsymbol{\beta}$ s and π s are (vectors of) production function parameters. u_i is a composite stochastic disturbance term ($u_i = \sum_{t=1}^T [ln.\epsilon_{it}] + ln.v_{iT}$). The residuals u_i , ξ_{it}^l and ξ_{it}^q are assumed to be positively correlated, so equations (10-11), (12) and (13) are estimated jointly, such that we can treat u_i as independent of the exogenous variation in \mathbf{L}_i and \mathbf{s}_i^q caused by variation in \mathbf{Z}_i . As a normalization restriction, the variance of each latent factor across both waves is set equal to one. We assume this structure to be 'invariant', in that neither the deterministic or stochastic components of the relationship between the inputs and outcome is affected by the fact of 'treatment', or employment (Heckman and Pinto, 2015).

Table 8: Contemporaneous effect of a one hour exogenous increase in employment hours on each latent factor (measured in standard deviations)

	Attitude		Study		Active Leisure		Social Life		Risky Behaviours	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
Wave 2										
Work Hours	-0.026* (0.014)	-0.055*** (0.013)	-0.007 (0.013)	-0.034** (0.014)	-0.053*** (0.013)	-0.008 (0.011)	0.016 (0.011)	0.005 (0.011)	0.032*** (0.012)	0.032*** (0.012)
p-value (unadj ¹), $p(k)$	(0.055)	(0.000)	(0.573)	(0.017)	(0.000)	(0.467)	(0.137)	(0.637)	(0.007)	(0.010)
p-value (adj ¹), p_{adj}	(0.359)	(0.000)	(0.999)	(0.129)	(0.000)	(0.995)	(0.696)	(0.999)	(0.053)	(0.076)
$r(.k)$	0.1076	0.0968	0.0980	0.0905	0.0833	0.0713	0.0919	0.0819	0.1129	0.1035
Observations	5136	5077	5136	5077	5136	5077	5136	5077	5136	5077
Wave 1										
Work Hours	-0.033* (0.018)	-0.040** (0.018)	-0.021 (0.013)	-0.009 (0.022)	-0.045*** (0.015)	-0.008 (0.015)	-0.005 (0.013)	0.019 (0.016)	0.015 (0.012)	0.030* (0.016)
p-value (unadj ¹), $p(k)$	(0.069)	(0.022)	(0.101)	(0.656)	(0.003)	(0.582)	(0.678)	(0.226)	(0.222)	(0.061)
p-value (adj ¹), p_{adj}	(0.433)	(0.164)	(0.587)	(0.999)	(0.023)	(0.999)	(0.999)	(0.871)	(0.859)	(0.395)
$r(.k)$	0.0982	0.0931	0.0821	0.0694	0.0883	0.0756	0.1097	0.0966	0.1068	0.0973
Observations	5171	5099	5171	5099	5171	5099	5171	5099	5171	5099

Notes: *: $p < 0.1$, **: $p < 0.05$, ***: $p < 0.01$. Standard errors, clustered by school, in parentheses. Adjusted p-values account for testing of multiple (10) hypotheses (five outcomes over two waves). All coefficients from single-factor-single-wave models. First stage equation for hours of employment not shown. Effect of hours of employment is identified by exclusion of labour market variables from second stage. (Local unemployment and VAT registrations in all cases, month of birth in all cases except Active Leisure, as age-within cohort is direct determinant of participation in sporting activities). Additional regressors as in Table 4

6.2.1 Estimates of tradeoffs in investments

As the first step in understanding the mechanisms behind the different policy effects estimated in section 5, we evaluate the extent to which other activities are crowded out, or facilitated, by part-time work. Table 8 shows estimates of γ_t^q (from equation 12) when the system represented in equations (10-13) and for $t = 1, 2$ is estimated for our two latent inputs and other three latent activities, and separately for boys and girls. This represents the contemporaneous crowd-out (or crowd-in) of each factor (s_{it}^q) by employment in each time period.

Each column shows first the coefficient on hours of employment in the single-factor, single wave, instrumental variable regression estimated using Full Information Maximum Likelihood with a tobit first stage. This is followed by its standard error and p-value obtained directly from the estimation procedure described. Since, for each gender, we are testing 10 null hypotheses (that of zero effect of labour supply on all of the five additional inputs in the education production function, each at two points in time), we would expect, on average, to find one statistically significant relationship at the 10% level that is in fact down to chance. (Alternatively stated, we can expect to make one type I error by erroneously rejecting a null hypotheses at the 10%

⁸Specifically, age 11 and 14 performance enter equations for dependent variables measured at $t = 2$. Only age 11 performance is included for dependent variables measured at $t = 1$, because for the majority of students, the examinations determining the age 14 score occurred after the wave 1 interview.

significance level). Hence, the final row in each panel shows the p-value adjusted for this multiple hypothesis testing. This uses the following modified Bonferroni Adjustment method, with the adjusted p-value calculated as follows:

$$p_{adj} = 1 - (1 - p(k))^{g(k)} \text{ where } g(k) = M^{1-r(.k)} \quad (14)$$

Here M equals 10, the number of outcomes being tested, $p(k)$ is the unadjusted p-value for the k^{th} outcome and $r(.k)$ is the mean of the (absolute) pairwise correlations between all the outcomes other than k . Both $p(k)$ and $r(.k)$ are shown in the table.⁹

Three clearly robust results emerge from Table 8. Firstly, having a job makes girls *care less about* their schoolwork, at age 15 (their Attitude is reduced). Secondly, working part-time causes boys to reduce their engagement with ‘active leisure’ activities at both ages 14 and 15. Thirdly, in-school employment induces both boys and girls to increase their uptake of risky behaviors at age 15.

Taking these results in turn, the coefficients at face-value suggest that Attitude is always endogenous to hours of employment. An additional hour per week crowds out Attitude by 2.6% and 5.5% of a standard deviation in wave 2, and 3.3% and 4.0% in wave 1, for boys and girls respectively, so the crowd-out is always larger for girls than boys. This means the effects of general tiredness, or a shift in preferences towards other activities opened by employment and earnings, outweigh any positive effect on motivation for education oriented human capital through making the role of education in widening future job options more salient (Dustmann and van Soest, 2007). However only the wave 2 result for girls is robust to the multiple hypothesis test, with their rate of crowd-out remaining significant at all conventional levels.

An additional hour of employment per week reduces boys’ participation in active leisure activities by 4.5% and 5.3% of a standard deviation at ages 14 and 15, equivalent to, for example, 6 hours employment per week reducing the probability of having participated in some form of sport in the last four weeks by 10 percentage points. The equivalent results for girls are quantitatively and statistically indistinguishable from zero. This conclusion retains statistical

⁹See Sankoh, Huque and Dubey (1997), pp.2534-2535, for discussion and Aker et al. (2012) for an economics application. While the number of outcomes considered here is large (10), the average correlation between them is small (≈ 0.1). Under these conditions this procedure performs well in Sankoh et al’s (1997) simulations, rejecting a false null hypothesis 5.9% of the time, given a maintained 5% experimentwise error rate.

significance at the 10% level after multiple testing, as does the finding that an additional hour of employment per week increases participation in risky behaviours by 3.2% of a standard deviation at age 15 for both boys and girls. Referring back to Tables 5 and 6, this coefficient can be interpreted as a typical employment burden of 6 hours per week raising the probability of ‘ever smoking’, ‘ever having tried cannabis’, and ‘ever consuming alcohol’ by approximately 6 percentage points each, or by 50%, 50%, and 20% respectively.

By contrast, the effects in relation to non-risky social activities are never statistically different from zero. This suggests that parents are willing to finance or facilitate these, but teenagers need their own income to access riskier consumption goods.

There is tentative evidence (initially significant at the 10% level) that having a job induces *early* (age 14) initiation into risky behaviours for girls. The coefficient is similar to that seen in at age 15, and double the rate seen for boys at age 14.

Taken alone, the rate of crowd-out of girls’ study by employment in wave 2 (3.4% of a standard deviation per hour per week) is also statistically significant, with $p = 0.017$, but after accounting for multiple testing it is marginally insignificant at the 10% level. This means the argument that having a job also makes girls *do less* study must be made tentatively. No other coefficients for study are significant, but the rate of crowd-out is larger for girls than boys both in the second wave and on average. This wave 2 difference may be explained partly by differences in travel time, and partly by girls’ relative lack of previous experience in combining employment with other activities (55% of girls working in wave 2 had not previously worked, compared with 44% of boys - see Table 2).

6.2.2 Estimates of production function parameters

The rank condition for identification of the parameter π_t^q (the effect of input q on final performance, holding constant labour supply) in the production function equation (13) requires that the coefficient on labour supply in equation (12) determining input q is not zero: $\gamma_t^q \neq 0$. The only specification in which γ_t^q is statistically different from zero and of the expected sign at the 10% level or less for both $t = 1, 2$ is Attitude for both boys and girls. Wave 2 Study, for girls, also changes significantly with employment at the 5% level. The Study production function parameter for girls is therefore identified in a contemporaneous ($t = 2$ only) specification. In the cumulative single-factor multiple-wave specification, these parameters are sensitive to the

wave 2 observations' collinearity with the underidentified wave 1 observation.

In Table 9 we show the estimated production function parameters π_t^L and π_t^q for each q , in cumulative ($T = 2$ with $t = 1, 2$) and contemporaneous ($T = 2$ with $t = 2$ inputs only) specifications, and a reduced form version accounting for wave 1 inputs only ($T = 2$ but $t = 1$ inputs only). In the top section, the 'policy effect' estimates for labour supply (ϕ_t^L) from the corresponding specifications are shown for ease of comparison. For each of our latent inputs, only coefficients in **bold** are identified by the criteria described above.

Table 9: Single-factor instrumental variable production function parameters

	Boys			Girls		
	Cumulative (Both waves)	Contemporaneous (Wave 2 only)	'Reduced form' (Wave 1 only)	Cumulative (Both waves)	Contemporaneous (Wave 2 only)	'Reduced form' (Wave 1 only)
Work Hours only (Policy effect)						
Work Hours Wave 2	-0.009 (0.012)	-0.008 (0.010)	.	-0.042** (0.021)	-0.044*** (0.017)	.
Work Hours Wave 1	0.012 (0.015)	.	0.001 (0.013)	-0.002 (0.022)	.	-0.042*** (0.016)
<i>Observations</i>	5122	5122	5122	5066	5066	5066
Attitude						
Work Hours Wave 2	-0.007 (0.012)	-0.003 (0.010)	.	-0.028 (0.020)	-0.021 (0.024)	.
Work Hours Wave 1	0.015 (0.015)	.	0.008 (0.013)	0.004 (0.021)	.	-0.031** (0.015)
Attitude Wave 2	0.170*** (0.020)	0.241*** (0.017)	.	0.224*** (0.038)	0.374 (0.247)	.
Attitude Wave 1	0.031 (0.021)	.	0.159*** (0.025)	0.122** (0.053)	.	0.239*** (0.016)
<i>Observations</i>	5019	5094	5130	4976	5042	5058
Study						
Work Hours Wave 2	-0.007 (0.012)	-0.007 (0.009)	.	-0.035* (0.019)	-0.042*** (0.015)	.
Work Hours Wave 1	0.010 (0.015)	.	0.004 (0.013)	-0.004 (0.021)	.	-0.042*** (0.016)
Study Wave 2	0.135*** (0.013)	0.169*** (0.038)	.	0.169*** (0.016)	0.127*** (0.015)	.
Study Wave 1	0.013 (0.014)	.	0.143*** (0.035)	0.098*** (0.017)	.	0.070*** (0.020)
<i>Observations</i>	5122	5150	5185	5066	5089	5111

Notes: **Bold:** Effect of exogenous change in hours of employment on contemporaneous latent factor is statistically significant at the 10% level or less (For cumulative specification see Table 8). Standard errors in parentheses. *: $p < 0.1$, **: $p < 0.05$, ***: $p < 0.01$. Longitudinal weights applied. **Additional covariates:** Parent's socio-economic status, parents' employment, child's prior educational performance, resident and non-resident siblings, lone parent family, child's special educational needs (SEN) classification, child's ethnicity.

Considering first the specification in which all parameters are identified in the cumulative specification, the coefficients on Attitude for both boys and girls are positive and large in magnitude (compared with those obtained in the OLS specifications) for wave 2 in both the cumulative and contemporaneous specifications, and wave 1 in the reduced form specification.

In wave 2 but controlling for behaviour at wave 1, a one standard deviation change in Attitude

to schoolwork at the margin crowded out by labour supply increases educational performance one year later by 17% and 22% of a standard deviation respectively for boys and girls. There is sufficient within-individual time-variation in employment and attitude to schoolwork that the wave 1 parameters on Attitude in the cumulative specification are both well identified and precisely estimated, though only the parameter for girls is statistically significant (an effect size of 12% of a standard deviation).

The contemporaneous specification yields a similar conclusion for boys, but a larger and considerably less precise parameter for girls. The wave 1 parameter in both reduced form specifications is considerably larger than that in the cumulative. It is similar in magnitude instead to the wave 2 estimates, suggesting that the wave 1 coefficient in these specifications absorbs the effect of the later input.

Considering next the instance where the parameter on the endogenous input is well-identified at only one point in time (Study for girls), the contemporaneous specification shows an effect size of wave 2 Study, at the margin crowded out by in-school employment, of 12.7% of a standard deviation.

Intuitively, there are quantitatively important changes in moving from the policy effect (top panel in Table 9) to the production function parameter (middle and lower panels of Table 9) only in instances where (i) in-school employment has a significant positive or negative effect on the latent input and (ii) the parameter on this latent input in the production function for educational performance is significantly different from zero.¹⁰

For girls, controlling for Attitude moderates the wave 2 policy effect by one-third (cumulative specification) and one-half (contemporaneous) of its size. This means the production function parameter is no longer statistically different from zero. While the overall effect of an additional hour's employment per week is to reduce GCSE scores by 4.2% of a standard deviation, its direct effect is only 2.8% of a standard deviation. The loss of statistical significance is not due to any loss in precision. Alternatively stated, approximately 33% of the negative 'policy effect' of in-school labour supply at age 15 is accounted for by the marginal reduction in Attitude to schoolwork induced by this labour supply.

¹⁰Alternatively stated, if employment has no effect on the input, then no matter how valuable or costly the input, controlling for its mediating effect will make no difference to the estimated effect of employment on academic performance. Similarly, if the input has no effect on educational performance, then no matter how much this input changes as a result of in-school employment, no difference will be made to the estimated effect of employment on academic performance.

The still-substantial residual negative estimated production function parameter on employment will partly be due to omitted *types* of inputs (we can control for only one additional endogenous input in each specification), but also largely due to omitted *timing* of inputs. For example, study during regular term-time (which we observe measures for) may be a poor predictor of intensive revision close to high-stakes exams, which may be highly productive but we do not observe. If labour supply is highly autocorrelated and later crowds out unobserved revision time the estimated effect of labour supply will be downward biased.

For boys, controlling for Attitude produces production function parameters which are all larger (less negative or more positive), than their policy effect counterpart. These results provide no evidence to shift the conclusion that the direct effect of having a job on school performance at the end of compulsory schooling for boys is zero.

7 Robustness to Imperfect Instrumental Variables

7.1 Bounding the effect of part-time work on performance, attitude and study

As discussed in section 4, there are reasons why month of birth, the local youth unemployment rate, and the density of suitable employers in the retail and wholesale (G) and hotel and restaurant (H) sectors, may have a residual direct effect on educational performance, or an indirect one if students adjust their attitude to schoolwork and their study effort in response to anticipated labour market opportunities were they to leave school.

In this section we use the method of Nevo and Rosen (2012) to test the robustness of the main conclusions of this paper; a significant negative effect of girls' in-school employment at age 15 on school performance at age 16, but a zero effect for boys; to the presence of 'Imperfect Instrumental Variables' (IIVs), which violate the standard exogeneity condition. This exercise also enables comparison of the conclusions obtained using each of our instruments individually.

Nevo and Rosen (2012) define three assumptions necessary to form bounds for the effect of an endogenous regressor on the outcome variable, having relaxed the assumption of zero direct effect of the instrument:

- Assumption 1: The instrument is negatively correlated with the endogenous regressor.

The signs of both the conditional and unconditional correlations between our instruments and Hours of Work are testable. Month of birth and the age 18-24 unemployment rate are indeed negatively correlated with hours of work. The density of sector G and H VAT registrations is positively correlated with Hours of Work, so without loss of identifying power or interpretation, for this section we define our instrument as the negative of this measure.

- Assumption 2: The signs of correlations of the instrument and the endogenous variable with the structural error term have to be the same.

Comparing the instrumental variable and OLS regressions suggests that Hours of Work has a positive correlation with the structural error term, and this assumption would conform to the standard endogeneity critique of studies of in-school employment (Baert et al, 2015), that those in work are positively selected on unobserved characteristics.

Month of birth will have a positive correlation with the structural error term only if age-at-test is the dominant mechanism for relative-age effects in educational performance, such that those born later in the school year can naturally be expected to catch up with their older peers, and this is not adequately captured by controlling for prior educational performance at two previous points in time.

Both an *increased* unemployment rate and *reduced* density of VAT registered employers (an increase in the negative of our instrument) can be argued to increase the opportunity cost of leaving school at age 16, and so prompt greater effort or a more positive attitude to schoolwork, leading to stronger educational performance.

We note that both these assumptions are consistent with our instrumental variables estimate being downward biased, in this application away from zero, and the true parameter lying in the interval between the OLS and IV estimates. Nevo and Rosen's (2012) third assumption enables this bound to be tightened:

- Assumption 3: The absolute correlation of the instrument with the structural error term is smaller than that of the endogenous regressor with the structural error term.

This final assumption requires not that month of birth, the local youth unemployment rate and density of potential employers have *zero* direct effect on educational performance (or attitude, or study effort), rather that any residual direct effect, having conditioned on a rich

set of individual and household characteristics, other area circumstances and the child’s prior educational performance at two points in time, will be *less* important than the unobservable characteristics driving positive selection into employment.

Nevo and Rosen’s (2012) procedure is, for each imperfect instrumental variable being used, first to obtain IV coefficients using that instrument alone, and then using as the instrument a weighted sum of the endogenous variable and imperfect instrument, which is exogenous by construction. This ‘adjusted’ instrument ($Z^\#$) is calculated as follows:

$$Z^\# = \sigma_Z X - \sigma_X Z \tag{15}$$

Here Z is the unadjusted instrument, X is Work Hours, and σ_Z and σ_X are their standard deviations. The ‘standard’ and ‘adjusted’ coefficients on Hours of Work then form intervals, and the intersection of these intervals is the identified set. Confidence regions for the identified set are then defined as follows (Stoye, 2009, see Fitzsimons and Malde, 2014 for an application):

$$CI = [\hat{\theta}_l - \frac{c_\alpha \hat{\sigma}_l}{\sqrt{N}}, \hat{\theta}_u + \frac{c_\alpha \hat{\sigma}_u}{\sqrt{N}}] \tag{16}$$

Here $c_\alpha = 1.96$ for desired 5% probability of type I error, and $\hat{\sigma}_l$ and $\hat{\sigma}_u$ are standard errors for the estimated coefficients $\hat{\theta}_l$ and $\hat{\theta}_u$ which form the lower and upper bounds of the identified set, respectively. N is the sample size.

Here our endogenous regressor, Hours of Work, is non-negative rather than continuous, and we implement the first-stage as a tobit, rather than linear regression. In creating our adjusted instruments we multiply the instruments by the standard deviation of *observed* hours of work, which is smaller than that of its linear and continuous counterpart, the latent ‘desired employment’. This means our adjusted instruments place a smaller relative weight on the imperfect instrument, and hence larger relative weight on the endogenous regressor. This means the upper bound we obtain for the effect of hours of employment on academic performance is looser; closer to but still more negative than the OLS coefficient; than implementation of this procedure with a continuous regressor would allow. This intuition about the statistical properties of our adapted IIV estimator (and more generally, implications of using a tobit first stage), is confirmed in the results obtained on simulated data which we present the results in Appendix A.2.

The results are shown in Table 10. Considering first the GCSE score for boys, neither the ordinary nor adjusted IV estimates with single instruments are ever significantly different from zero or from each other, although in each case the point estimate for the ordinary method is smaller (more negative) than the adjusted. The confidence region for the identified set implies that the true damaging effect of part-time work is between 0.35% and 1.01% of a standard deviation per hour worked. Turning to girls, the confidence region implies an effect size of at least 1% percent of a stand deviation per hour per week, and no more than 3.7%.

It is also instructive to compare the coefficients obtained using each individual instrument at a time. Assigning their differences to systematic bias rather than to random variation implies that month of birth (the weakest of our three instruments, but that which produces the tightest bounds), is performing best in controlling for unobserved heterogeneity.

Accounting for imperfect instrumental variables, maintains a statistically significant crowd-out by part-time work of ‘Attitude’ or effort during school hours, for both boys and girls. We find an upper bound for this negative effect of 1.4% and 1.5% of a standard deviation, and lower bounds of 2.5% and 3.9%. With respect to study the whole of the 95% confidence region no longer excludes zero for either boys or girls. The lower bound for girls, at -2.95% of a standard deviation per hour, is slightly less damaging than that implied by our preferred estimates, in the main body of the paper, while for boys the effect size was always negligible.

Altogether, these results show that our main conclusion; a statistically significant negative effect of girls’ in-school employment at age 15 on school performance at age 16; is robust to relaxing the assumption that our instrumental variables have zero direct effect on academic performance. However, the magnitude of the effect may be as little as one-quarter of that shown in our headline results. An effect size of 1% of a standard deviation per hour worked per week is equivalent to the average employment workload (among those working) of 6 hours per week reducing GCSE performance by one-and-a-half grades in one GCSE subject. This alone is likely to compromise the future educational or labour market opportunities only of individuals very close to the threshold for continuation in full-time education (5 A*-C grades including English and Maths). However, for both boys and girls, the result that employment has a negative effect on individuals’ attitude to schoolwork is also robust to this exercise. For

Table 10: Using the ‘Imperfect Instrumental Variables’ method to define bounds for the effect of Hours of work on academic performance, attitude and study time. Contemporaneous specification with tobit first-stage.

Outcome: Instrument	Boys					
	GCSE score		Attitude		Study	
	Ordinary	Adjusted	Ordinary	Adjusted	Ordinary	Adjusted
Month of birth	-0.010 (0.009)	-0.004 (0.004)	-0.025* (0.014)	-0.010 (0.006)	-0.006 (0.013)	-0.001 (0.006)
18-24 unemployment rate	-0.013 (0.009)	-0.003 (0.003)	-0.026* (0.014)	-0.007 (0.005)	-0.008 (0.014)	0.002 (0.006)
‘G’ / ‘H’ VAT registration density	-0.010 (0.009)	-0.004 (0.004)	-0.030** (0.014)	-0.014** (0.006)	-0.006 (0.013)	0.001 (0.006)
Intersection of bounds	-0.010	-0.004	-0.025*	-0.014**	-0.006	-0.001
95% confidence region	[-0.0101	-0.0035]	[-0.0255	-0.0138]	[0.0062	-0.0013]
Observations ^a	5056	5054	5054	5054	5112	5109

Outcome: Instrument	Girls					
	GCSE score		Attitude		Study	
	Ordinary	Adjusted	Ordinary	Adjusted	Ordinary	Adjusted
Month of birth	-0.036** (0.017)	-0.010*** (0.004)	-0.039*** (0.015)	-0.008 (0.005)	-0.029* (0.016)	0.001 (0.006)
18-24 unemployment rate	-0.048*** (0.017)	-0.008** (0.003)	-0.051*** (0.015)	-0.005 (0.005)	-0.040*** (0.015)	0.006 (0.005)
‘G’ / ‘H’ VAT registration density	-0.041** (0.017)	-0.013*** (0.004)	-0.045*** (0.015)	-0.015*** (0.005)	-0.029* (0.015)	0.001 (0.006)
Intersection of bounds	-0.036**	-0.013***	-0.039***	-0.015***	-0.029*	0.001
95% confidence region	[-0.0365	-0.0105]	[-0.0919	-0.0150]	[-0.0295	0.0011]
Observations ^a	5054	5055	4996	5008	5054	5054

Notes: Confidence regions for the identified set are as defined by Stoye (2009), $CI = [\hat{\theta}_l - ((c_\alpha \hat{\sigma}_l)/\sqrt{N}), \hat{\theta}_u + ((c_\alpha \hat{\sigma}_u)/\sqrt{N})]$, where $c_\alpha = 1.96$ for desired 5% probability of type I error, $\hat{\sigma}_l$ and $\hat{\sigma}_u$ are standard errors for the estimated coefficients θ_l and θ_u , and N is the sample size. ^a Observations are for the equation defining the bound in question, i.e. using VAT registrations in all cases except the upper bound for Study, which uses month of birth. *: $p < 0.1$, **: $p < 0.05$, ***: $p < 0.01$. First stage equation for hours of employment not shown. Additional regressors as in Table 4

a proponent of part-time work by teenagers, at best an additional hour of employment per week still reduces Attitude by 1.5% of a standard deviation. This is a significant reduction in teenagers’ view of the importance of education for their longer-term opportunities, which may concern educators more than the effects on exam results themselves.

7.2 An alternative explanation for the damaging effect of late versus early employment

A key contribution of this paper is to separate out the effects of in-school employment taken at different points in time on subsequent GCSE performance. While in-school employment never has a significant effect for boys, our main conclusion for girls is that (conditioning on hours worked at the other point in time) having a job at age 14 has no effect on GCSE performance, while having a job at age 15 is highly damaging to academic performance. We attribute this to the higher productivity of inputs crowded out by employment closer to exam time and during the courses for the qualifications determining our performance measure.

An alternative explanation concerns the timing of employment opportunities and expectations about those which will be available in the future. In the previous section, we established that increased current employment opportunities directly reduce current attitude and effort towards schoolwork. We reasoned this was through raising the expected opportunity cost of staying in school after age 16, and this result potentially biased the estimated effect of employment on academic performance downwards, though not sufficiently wholly to account for its negative effect in our headline estimates.

In a dynamic setting, *improving* employment opportunities in the local area will concentrate individuals' labour market activity at a later point in time. If individuals expect these opportunities to continue to improve (i.e. they expect changes in opportunities to be serially correlated), thereby raising the expected opportunity cost of staying in school at 16, then this concentration of labour market activity at the later point in time will be also be associated with lower effort and attitude toward schoolwork.

We compiled a dataset of the deviation from the national average in the change from one year earlier of (i) the youth unemployment rate and (ii) the density of sector G and H VAT-registered enterprises, for every LAD in England from 1995-2006. We implemented the Wooldridge (2002) test for serial correlation in panel data (`xtserial` in Stata, Drukker, 2003). We could not reject that changes in the youth unemployment are serially uncorrelated ($F(1, 325) = 0.171$, $p = 0.6792$), but found that changes in the density of retail, wholesale, hotel and restaurant businesses are highly serially correlated ($F(1, 325)=9.981$, $p = 0.0017$). Hence, any strategic changes in behaviour must be in areas specific to the opportunities anticipated to be available in these sectors, rather than the strength of the labour market in general. Since we observe

students' choice of optional subjects to be studied at age 15 and 16, we can establish whether there is evidence for any strategic response in the propensity to study catering or food technology, or indeed any conditional correlation with local employment opportunities, which may also bias our results if some qualifications are easier than others in which to obtain a high grade.

7.2.1 Subject choice and changes in employment opportunities

Table 11 below presents tests of whether students make their GCSE subject choices strategically, in anticipation of their employment opportunities when they leave school. The dependent variable is an indicator equal to one if the student, in school year 10 (the penultimate year of compulsory schooling and first year in which optional courses are taken) is studying for a GCSE in Food Technology or a National Vocational Qualification in Leisure and Tourism, or Hospitality and Catering.

Table 11: Effect of changes in local labour market conditions on GCSE subject choice

Explanatory variable:	Change in stock of sector G/H VAT registered enterprises per 100 youths at $t = 2$			
	Boys		.	Girls
Logistic regression odds ratio	1.387	(0.416)	.	0.752 (0.199)
Linear probability model coefficient	0.039	(0.037)	.	-0.057 (0.053)
Observations	5112		.	5055
% taking a 'food subject'	15.54		.	28.13

Notes: Standard errors in parentheses. *: $p < 0.1$, **: $p < 0.05$, ***: $p < 0.01$. Longitudinal weights applied. **Additional covariates:** Index of multiple deprivation (IMD), parent's socio-economic status, parents' employment, child's prior educational performance, resident and non-resident siblings, lone parent family, child's special educational needs (SEN) classification, child's ethnicity.

We see no significant effects of changes over time in the availability of retail, wholesale, hotel and restaurant jobs, there is no significant effect on the propensity to choose food-orientated subjects at GCSE level in either our logit or linear probability models. Moreover the coefficients obtained for boys and girls are of opposite signs. This suggests no threat to the integrity of the policy effects and production function parameters we estimate for hours of work in wave 2 versus wave 1, driven by anticipation of leaving full-time education in response to an improving labour market in this specific field.

8 Summary and conclusions

A large proportion of teenagers and young adults in the UK and North America take some formal paid employment while still in full-time education. In this paper we first evaluated the net effect of teenage employment on performance in high stakes qualifications by English pupils in compulsory education. We estimated OLS and instrumental variables cumulative specifications for age 16 academic performance. These identified the net effects of part-time employment at different points in time, including indirect effects caused by unobserved changes in other activities which may result from changes in labour supply. We identified a significant negative effect of the labour supply of girls at age 15 only. For them, an additional hour of employment reduces academic performance by 4.2% of a standard deviation of our chosen measure. This effect is quantitatively large, and sufficient to substantially narrow subsequent educational opportunities.

We then sought to understand the mechanism behind this effect by controlling for additional endogenous inputs, assuming a causal structure for the tradeoff in time-use between employment and study and other potentially productive or damaging activities. This structure enables us to identify the parameter for the contribution of these additional inputs, at the margin crowded out or facilitated by employment, in the production function for educational performance. We provide evidence which supports the gender differential in the net effect of in-school employment being due to employment depressing girls' attitude to and motivation for schoolwork to a greater extent than boys. This attitude in turn is a stronger determinant of academic success for girls than boys. More tentatively, while students' own study effort is shown to be similarly valuable for boys and girls, it appears that, at least at age 15, this is crowded out to a significantly greater extent by the employment of girls than boys. However, even when study time at home and outside of lessons is accounted for in this way, the estimated production function parameter on girls' labour supply at age 15 remains large in magnitude. This suggests that employment is still crowding out important unobserved inputs closer in time to the final high-stakes exams.

We noted that our sample faced a system of continuous assessment over a two-year period, contributing to a significant minority of the overall marks in most GCSE subjects, with the remainder determined by final exams sat at the end of compulsory schooling. Although our conclusions are qualitatively very similar to those obtained the last time the system was governed by final exams only (Dustmann and van Soest, 2007), it is possible that the ongoing move from

continuous assessment to final-exam-only assessment of GCSEs from 2013-2017 will reduce the penalty to having a job by reducing the relative importance of long-term inputs of study in favour of intensive revision shortly before these high-stakes exams. A comparison of our results with those obtained from a second, more recent, LSYPE cohort ('LSYPE2', for which only the wave 1/age 14 data is yet available) will enable future research to evaluate the effect of assessment structure on the impact of part-time work and long-term study.

In policy terms, our results suggest that any step to reduce the hours of part-time employment among schoolchildren in the final two years of compulsory education, through enforcement or revision of the current permitted hours, would improve the age-16 academic performance of girls but not boys. While there is dynamic state dependence in employment; it is easier to keep a job than obtain one (see Appendix A.6, page 61); we do not make a case for raising the age at which children are allowed to take a job: the micro benefit of not having a job at age 14 making it harder to get a job at age 15 (when the damage to academic performance is done) will not translate into a macro benefit if the entire cohort faces the same initial restriction. Moreover, although we find no negative effects on age 16 academic performance of having a job at age 14, this work does not address dynamic questions, such as whether earlier employment experience either increases students' ability to accommodate this without crowding out other productive activities, or instead influences their preferences to initiate a poorer trajectory of investment activities.

It should be borne in mind that those taking part-time employment are on average somewhat positively selected, and that the crowding out of study observed here actually moderates the significant gender gap by which girls outperform boys at age 16 in the UK. There may also be a role for prior employment experience in improving subsequent employment opportunities beyond compulsory education. Nevertheless, we have found no evidence for a positive role for in-school employment in producing the education-oriented human capital captured by our exam performance measure. Insofar as in-school employment induces early initiation into risky behaviours and a reduction in students' valuation of education-oriented human capital, and that these effects may be expected to persist, we have also shown evidence for an additional mechanism, beyond age 16 exam scores, through which subsequent human capital accumulation may be restricted.

References

- Aker, Jenny C., Boumnijel, Rachid., McClelland, Amanda., Tierney, Niall., February 2012. Zap it to me: The impacts of a mobile cash transfer program. Mimeograph.
- Arrow, Kenneth J., 1973. The theory of discrimination. In: Ashenfelter, O., Rees, A. (Eds.), *Discrimination in Labour Markets*. Springer.
- Baert, Stijn., Rotsaert, Olivier, Verhaest, Dieter., Omeij, Eddy, 2015. A signal of diligence? student work experience and later employment chances. IZA Discussion Paper 9170, Bonn, Germany.
- Barron, John., Ewing, Bradley., Waddell, Glenn., 2000. The effects of high school athletic participation on education and labor market outcomes. *Review of Economics and Statistics* 82 (3), 409–421.
- Becker, Gary S., 1965. A theory of the allocation of time. *Economic Journal* 75, 493–517.
- Bound, John, Jaeger, David A., Baker, Regina M., 1995. Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American Statistical Association* 90 (5).
- Browning, Martin., Heinesen, Eskil., 2014. Study versus television. *IZA Journal of Labor Economics* 3 (2), 1–16.
- Bureau of Labor Statistics, 2000. Report on the youth labor force.
- Burgess, Simon., Metcalfe, Robert., Proud, Steven., 2011. Student effort and educational attainment: Using the England football team to identify the education production function. CMPO Working Papers 276, Bristol, United Kingdom.
- Buscha, Franz., Maurel, Arnaud., Page, Lionel., Speckesser, Stefan., 2011. The effect of high school employment on educational attainment: A conditional difference-in-differences approach. *Oxford Bulletin of Economics and Statistics* 74 (3), 380–396.
- Clark, Andrew E., Lohéac, Youenn., 2007. “It wasn’t me, it was them!” Social influence in risky behaviour by adolescents. *Journal of Health Economics* 26 (4), 763–784.
- Crawford, Claire., Dearden, Lorraine., Greaves, Ellen., 2013. When you are born matters: Evidence for England. Institute for Fiscal Studies Report R80, London, United Kingdom.

- Cunha, Flavio., Heckman, James., 2007. The technology of skill formation. *American Economic Review Papers and Proceedings* 97 (2), 31–47.
- Cunha, Flavio., Heckman, James., 2009. Human capital formation in childhood and adolescence. *CESifo DICE Report* 7 (4), 22–228.
- Curtis, Susan., Lucas, Rosemary., 2001. A coincidence of needs? Employers and full-time students. *Employee Relations* 23 (1), 38–54.
- Datar, Ashlesha., Mason, Bryce., 2008. Do reductions in class size “crowd out” parental investment in education? *Economics of Education Review* 27, 712–723.
- Department for Children, Schools and Families, 2009. Guidance on the employment of children. Department for Children, Schools and Families: https://www.education.gov.uk/publications/eOrderingDownload/Child_employment09.pdf, accessed 24-08-2012.
- Department for Education, January 2012. Performance points: A practical guide to key stage 4 and 5 performance points. URL: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/160734/performance_points.pdf, accessed 10th August, 2015.
- Drukker, David M., 2003. Testing for serial correlation in linear panel-data models. *Stata Journal* 3 (2), 168–177.
- Dustmann, Christian., van Soest, Arthur., 2007. Part-time work, school success and school leaving. *Empirical Economics* 32 (2), 277–299.
- Fitzsimons, Emla., Malde, Bansi., 2014. Empirically probing the quantity-quality model. *Journal of Population Economics* 27, 33–68.
- Freyer, Roland., 2010. Financial incentives and student achievement: Evidence from randomized trials. *NBER Working Papers* 15898, 1–91.
- Granovetter, M. S., 1973. The strength of weak ties. *American Journal of Sociology* 78, 1360–1380.
- Häkkinen, Iida., 2006. Working while enrolled in a university: does it pay? *Labour Economics* 13, 167–189.

- Hancock, Ruth., Morciano, Marcello., Pudney, Stephen., Zantomio, Francesca., 2013. Do household surveys give a coherent view of disability benefit targeting? A multi-survey latent variable analysis for the older population in Great Britain. ISER Working papers 2013 (05), Colchester, United Kingdom.
- Heckman, James., Pinto, Rodrigo., 2015. Econometric mediation analyses: Identifying the sources of treatment effects from experimentally estimated production technologies with unmeasured and mismeasured inputs. *Econometric Reviews* 34 (1), 6–31.
- Hibbett, A., Beatson, M., 1995. Young people at work. *Employment Gazette* 4, 169–177.
- Hotz, V. Joseph., Xu, Lixin Colin., Tienda, Marta., Ahituv, Avner., 2002. Are there returns to the wages of young men from working while in school? *The Review of Economics and Statistics* 84 (2), 221–236.
- Howieson, Cathy., McKechnie, Jim., Semple, Sheila., 2006. The nature and implications of the part-time employment of secondary school pupils. Scottish Executive Social Research.
- Kalencoski, Charlene Marie., Pabilonia, Sabrina Wulff., 2010. Parental transfers, student achievement and the labor supply of college students. *Journal of Population Economics* 23 (2), 469–496.
- Kalencoski, Charlene Marie., Pabilonia, Sabrina Wulff., 2014. Does high school homework increase academic achievement? IZA Discussion Paper 8142, Bonn, Germany.
- Kooreman, Peter., 2009. The early inception of labor market gender differences. *Labour Economics* 16, 135–139.
- Light, Audrey., 2001. In-school work experience and the returns to schooling. *Journal of Labor Economics* 19 (1), 65–93.
- Lillydahl, Jane H., 1990. Academic achievement and part-time employment of high school students. *Journal of Economic Education* 21 (3), 307–316.
- Lundborg, Petter., 2006. Having the wrong friends? Peer effects in adolescent substance use. *Journal of Health Economics* 25 (2), 214–233.
- Mizen, Philip., Bolton, Angela., Pole, Christopher., 1999. School age workers: The paid employment of children in Britain. *Work, Employment and Society* 13 (3).

- Molitor, Christopher J., Leigh, Duane E., 2005. In-school work experience and the returns to two-year and four-year colleges. *Economics of Education Review* 24, 459–468.
- Montmarquette, Claude., Viennot-Briot, Nathalie., Dagenais, Marcel., 2007. Dropout, school performance and working while in school. *Review of Economics and Statistics* 89 (4), 752–760.
- Nevo, Aviv., Rosen, Adam., 2012. Identification with imperfect instruments. *Review of Economics and Statistics* 94 (3), 659–671.
- Oettinger, Gerald S., 1999. Does high school employment affect high school academic performance. *Industrial and Labor Relations Review* 53 (1), 136–151.
- Office for National Statistics, January 2008. nomis official labour market statistics: VAT Registrations/Deregistrations by Industry. URL: <http://www.nomisweb.co.uk/articles/338.aspx>, accessed 7th August, 2015.
- Office of the Deputy Prime Minister, 2004. The English Indices of Deprivation 2004: Summary (revised). <http://webarchive.nationalarchives.gov.uk/20120919132719/www.communities.gov.uk/documents/communities/pdf/131206.pdf>, accessed 25-07-2013.
- Payne, Joan., 2004. The impact of part-time jobs in years 12 and 13 on qualification achievement. *British Educational Research Journal* 29 (4), 599–611.
- Plug, Erik J.S., 2001. Season of birth, schooling and earnings. *Journal of Economic Psychology* 22, 641–660.
- Rothstein, Donna., 2007. High school employment and youths’ academic achievement. *Journal of Human Resources* 42 (1), 194–213.
- Ruhm, Christopher J., 1997. Is high school employment consumption or investment? *Journal of Labor Economics* 15 (4), 735–776.
- Sabia, Joseph., 2009. School year employment and academic performance of young adolescents. *Economics of Education Review* 28 (2), 268–276.
- Sankoh, A.J., Huque, M.F., Dubey, S.D., 1997. Some comments on frequently used multiple endpoint adjustment methods in clinical trials. *Statistics in Medicine* 16, 2529–2542.
- Spence, Michael, 1973. Job market signaling. *Quarterly Journal of Economics* 87, 355–374.

- Stiglitz, Joseph E., 1975. The theory of screening, education, and the distribution of income. *American Economic Review* 65, 283–300.
- Stinebrickner, Ralph., Stinebrickner, Todd R., 2003. Working during school and academic performance. *Journal of Labor Economics* 21 (2), 473–491.
- Stoye, Jörg., 2009. More on confidence intervals for partially identified parameters. *Econometrica* 77 (4), 1299–1315.
- Tyler, John H., 2003. Using state child labor laws to identify the effect of school-year work on high school achievement. *Journal of Labor Economics* 21 (2), 381–408.
- De Fraja, Gianni., Oliveira, Tania., Zanchi, Luisa., 2010. Must try harder: Evaluating the role of effort in educational attainment. *Review of Economics and Statistics* 92 (3), 577–597.
- Del Boca, Daniela., Monfardini, Chiara., Nicoletti, Cheti., 2012. Self investments of adolescents and their cognitive development. IZA Discussion Paper 6868, Bonn, Germany.
- Wooldridge, Jeffrey., 2005. Simple solutions to the initial conditions problem in dynamic, non-linear panel-data models with unobserved heterogeneity. *Journal of Applied Econometrics* 20, 39–54.
- Wooldridge, Jeffrey M., 2002. *Econometric Analysis of Cross Section and Panel Data*. MIT Press, Cambridge (MA).

A Appendix

A.1 Graphical illustrations of model predictions

A.1.1 Uncertainty leads to underinvestment in Study

In section 3 we stated that the theoretical model predicts that uncertainty in the evolution of human capital and in GCSE performance will lead to underinvestment in study, compared to the case with no uncertainty. Figure A1 shows a graphical illustration of this prediction, which is an application of Jensen's inequality.

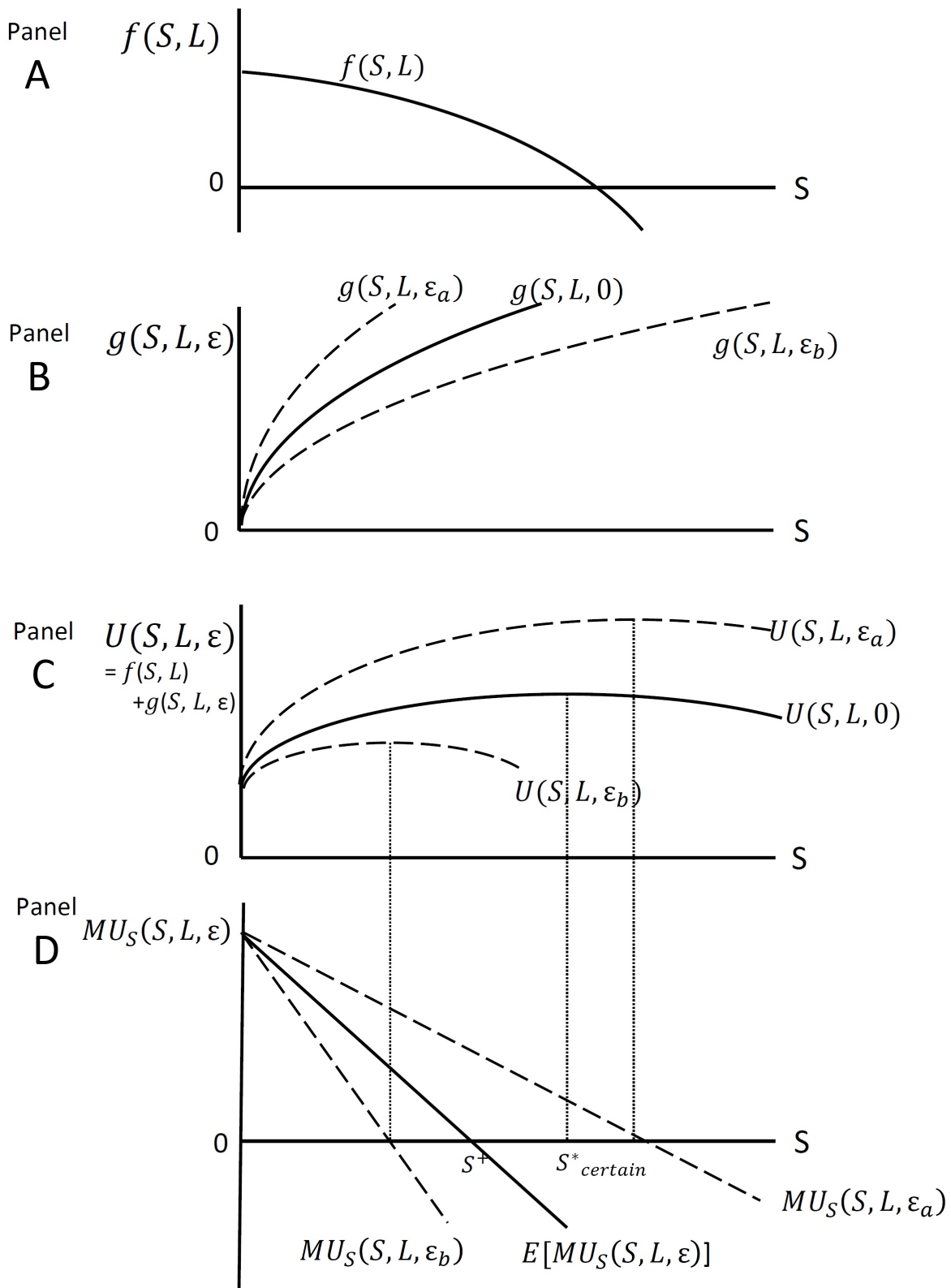
Panel A shows the deterministic relationship between the present-orientated component of utility $f(S, L)$ and study (S). This has its peak at $S = 0$. Because consumption and leisure are complements $f(S, L)$ is decreasing in S at an accelerating rate.

Panel B shows the relationship between the future-orientated component of utility $g(S, L, \epsilon)$ and study (S). Its expected value $g(S, L, 0)$ is illustrated by the middle line. The upper and lower lines $g(S, L, \epsilon_a)$ and $g(S, L, \epsilon_b)$ show the utility from two possible realizations (labelled ϵ_a and ϵ_b) of the vector of shocks to human capital (ϵ) and GCSE performance conditional on human capital (v), which occur with equal probability.

Panel C shows the sum of the graphs in panels A and B. The point $S_{certain}^*$ shows the optimal choice of study for the case with no uncertainty. At this point $U(E[g(\cdot)])$ is at its maximum and $MU(E[g(\cdot)]) = 0$. Panel D plots the net marginal utility of study for the expected realization of g , labelled $MU(E[g(\cdot)])$, and the expected net marginal utility of study, labelled $E[MU(g(\cdot))]$. The latter is the mean of the functions $MU(g(S, L, \epsilon_a))$ and $MU(g(S, L, \epsilon_b))$.

It can be seen that expected marginal utility $E[MU(g(\cdot))]$ is always below the marginal utility of the expected outcome $MU(g(S, L, 0))$, and $E[MU(g(\cdot))] = 0$ at a lower Study (S^+) than $MU(g(S, L, 0)) = 0$ (at $S_{certain}^*$). This means the individual will underinvest in Study relative to the case with no uncertainty. As the random draws ϵ_1 and ϵ_2 become known, the overall variance of the residual uncertainty becomes smaller. Uncertainty is reduced still further if the ϵ_t s and v are positively serially correlated, and the individual can update his prior about the likely sign and magnitude of the shock to human capital.

Figure A1: Uncertainty leads to underinvestment in Study



A.1.2 Strictly negative tradeoff of employment and study

This appendix shows that in our model, the effect of employment on study time must be negative. Hence, the coefficient on employment when study is omitted from a regression model will represent a lower bound of the human capital effect of employment on academic performance. Treating labour supply as fixed the first order condition for the choice of study time is:

$$\frac{\partial E[g(\mathbf{S}, \mathbf{L})]}{\partial S_t} + \frac{\partial f_t(\mathbf{S}, \mathbf{L})}{\partial S_t} = 0 = [\cdot] \quad (\text{A1})$$

Using implicit differentiation, the partial derivative of study with respect to labour supply that maintains this optimal position is as follows:

$$\frac{\partial S_t}{\partial L_t} = - \frac{[\frac{\partial [\cdot]}{\partial L_t}]}{[\frac{\partial [\cdot]}{\partial S_t}]} = - \frac{[\frac{\partial^2 E[g(\mathbf{S}, \mathbf{L})]}{\partial S_t \partial L_t}] + [\frac{\partial^2 f_t(\mathbf{S}, \mathbf{L})}{\partial S_t \partial L_t}]}{[\frac{\partial^2 E[g(\mathbf{S}, \mathbf{L})]}{\partial S_t^2}] + [\frac{\partial^2 f_t(\mathbf{S}, \mathbf{L})}{\partial S_t^2}]} \quad (\text{A2})$$

We now show that for an interior solution both the denominator and numerator of this expression must be strictly negative, meaning the partial derivative itself is negative.

Figure A2 plots the marginal utility (MU) of study for the present- and future-orientated components of utility ($f(\mathbf{S}, \mathbf{L})$ and $E[g(\mathbf{S}, \mathbf{L})]$), holding labour supply constant. The expected MU of study in the future-orientated component is always positive ($\frac{\partial E[g(S)]}{\partial S} > 0$), but decreasing in study ($\frac{\partial^2 E[g(S)]}{\partial S^2} < 0$). The expected MU of study in the present-orientated component is always negative ($\frac{\partial f(S)}{\partial S} < 0$) but decreasing in study ($\frac{\partial^2 f(S)}{\partial S^2} < 0$). For an interior solution it must be the case that at $S = 0$, $\frac{\partial E[g(S)]}{\partial S} > -\frac{\partial f(S)}{\partial S}$, and therefore in the (S, MU_S) space the function $\frac{\partial E[g(S)]}{\partial S}$ must cut $\frac{\partial f(S)}{\partial S}$ from above at the optimal point. Hence, $[\frac{\partial^2 E[g(\mathbf{S}, \mathbf{L})]}{\partial S_t^2}] + [\frac{\partial^2 f_t(\mathbf{S}, \mathbf{L})}{\partial S_t^2}]$, the denominator in equation A2, is less than zero.

Figure A3 plots the MU of study for the present- and future-orientated components of utility over a range for labour supply. The expected MU of study in the future-orientated component is always positive ($\frac{\partial E[g(S)]}{\partial S} > 0$) but may be increasing, constant or decreasing in L depending on whether study and labour supply are complements in production. The expected MU of study in the present-orientated component is always negative ($\frac{\partial f(S)}{\partial S} < 0$), and because consumption and leisure are complements, decreasing in labour supply ($\frac{\partial^2 f(S)}{\partial S \partial L} < 0$). For an interior solution it must be the case that at $L = 0$, $\frac{\partial E[g(S)]}{\partial S} > -\frac{\partial f(S)}{\partial S}$, so again in the (L, MU_S) space, the function $\frac{\partial E[g(S)]}{\partial S}$ must cut $\frac{\partial f(S)}{\partial S}$ from above at the optimal point. Hence, $[\frac{\partial^2 E[g(\mathbf{S}, \mathbf{L})]}{\partial S_t \partial L_t}] + [\frac{\partial^2 f_t(\mathbf{S}, \mathbf{L})}{\partial S_t \partial L_t}]$, the

numerator in equation A2, is less than zero. Hence, $\frac{\partial S_t}{\partial L_t}$, the negative of the ratio of two negative expressions, is itself negative.

Figure A2: Showing that $[\frac{\partial^2 E[g(\mathbf{S}, \mathbf{L})]}{\partial S_t^2}] + [\frac{\partial^2 f_t(\mathbf{S}, \mathbf{L})}{\partial S_t^2}] < 0$

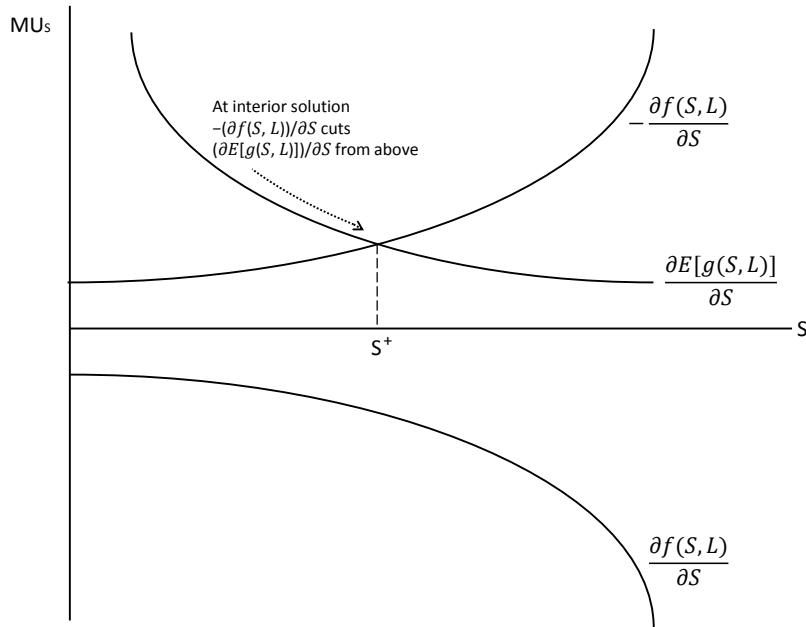
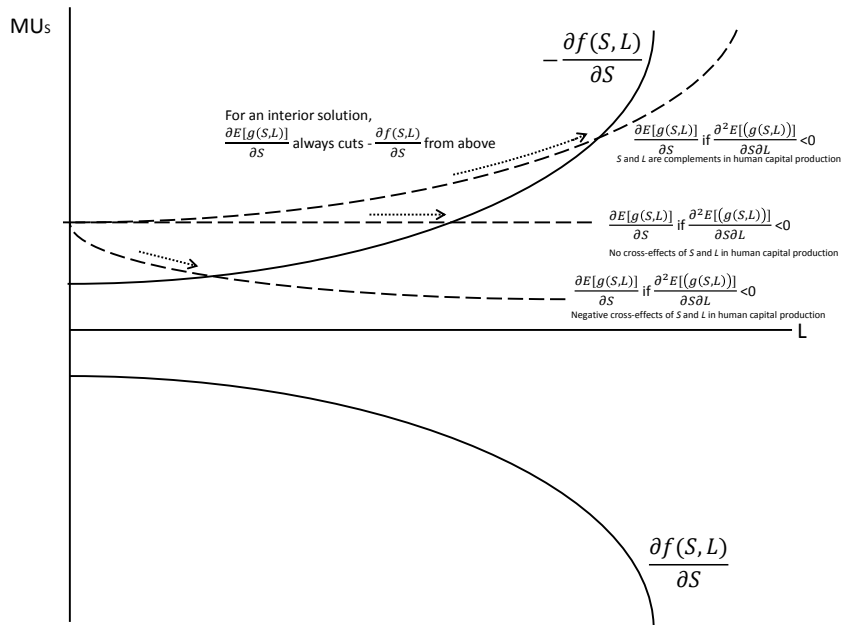


Figure A3: Showing that $[\frac{\partial^2 E[g(\mathbf{S}, \mathbf{L})]}{\partial S_t \partial L_t}] + [\frac{\partial^2 f_t(\mathbf{S}, \mathbf{L})}{\partial S_t \partial L_t}] < 0$



A.2 Statistical properties of our estimators with a tobit first stage

A.2.1 FIML versus linear IV

In this paper all our instrumental variable estimates have been obtained using full information maximum likelihood (FIML) with a tobit first stage for hours of employment at each point in time. In this section we first present simulations to compare the statistical properties of this estimator with a linear instrumental variables estimator.

We use 10000 replications, drawing a structural error term representing ability (a), desired hours of work (x^*) and an instrument (z) from a normal distribution with means zero, standard deviations of 0.5, 1 and 1, and correlations of -0.25, zero, and -0.5. In other words we assume a negative ability bias and a strong, valid instrument.

$$\begin{pmatrix} a \\ x^* \\ z \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0.25 & -0.125 & 0 \\ -0.125 & 1 & -0.5 \\ 0 & -0.5 & 1 \end{pmatrix} \right)$$

We then set actual hours of work (x) equal to zero if desired hours of work are negative, and construct an outcome variable $y = x + a$, such that the ‘true’ coefficient on hours of work is 1, but we expect a negative omitted variables bias if ability is excluded. Table A1 shows the results obtained from 10,000 replications of the OLS regression, implementation of the linear instrumental variables estimator, and the FIML estimator with a tobit first stage:

Table A1: Coefficients obtained from simulated data with valid instruments

Estimator	Replications	Mean coeff'	Std. dev. of coeff's	5th percentile	95th percentile
OLS	10,000	0.8167	<i>0.0118</i>	0.7976	0.8361
Linear IV	10,000	1.0007	<i>0.0284</i>	0.9537	1.0471
FIML, tobit 1st stage	10,000	1.0003	<i>0.0215</i>	0.9650	1.0360

The results confirm that under conditions meeting the assumptions of the instrumental variables estimator, estimation of a tobit first-stage confers no benefit in terms of consistency or bias, but a marginal gain in precision. The 90% confidence interval, between the 5th and 95th percentiles

of coefficients obtained from these replications, is 24% narrower, a margin of 0.0224, or 2.25% of the true coefficient,

A.2.2 Imperfect instrumental variables with a tobit first stage

We now present simulations to evaluate the performance of the Nevo and Rosen (2012) ‘imperfect instrumental variables’ estimator at tightening the bounds for the true effect of an explanatory variable when, as in our case, the endogenous regressor is non-negative rather than continuous, and the first stage is estimated using a tobit rather than linear regression.

We adopt the same conditions as above, except for setting the correlation between our instrumental variable and ability to -0.1, rather than zero. Hence the instrument is negatively correlated with hours of work ($\rho = -0.5$) has the same sign of correlation with ability (the structural error term) as does hours of work (negative), but the absolute value of the correlation of the instrument with the structural error term is smaller than that of hours of work with the structural error term (-0.1 versus -0.25), meeting all three of the assumptions specified by Nevo and Rosen (2012).

$$\begin{pmatrix} a \\ x^* \\ z \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0.25 & -0.125 & -0.05 \\ -0.125 & 1 & -0.5 \\ -0.05 & -0.5 & 1 \end{pmatrix} \right)$$

Table A2 below shows the results obtained from 10,000 replications of the OLS regression, the linear instrumental variable specification and its ‘adjusted for imperfect instruments’ counterpart, the tobit first-stage instrumental variable specification and two ‘adjusted for imperfect instruments’ counterparts. In all cases the ‘true’ coefficient, by construction, is 1. In constructing the adjusted instrument in the first, the instrument is weighted by the standard deviation of ‘latent desired employment’ (1) and in the second by the standard deviation of observed employment.

These simulations show that with an imperfect instrument, not only does FIML with a tobit first stage produce more precise estimates than the linear instrumental variables method, but the bias is smaller (approximately 11% rather than 20%). In this example, implementing Nevo and Rosen’s (2012) imperfect instrumental variables procedure tightens the lower bound for the

Table A2: Coefficients obtained from simulated data with imperfect instruments

Estimator	Replications	Mean coeff'	Std. dev. of coeff's	5th percentile	95th percentile
OLS	10,000	0.8165	<i>0.0117</i>	0.7974	0.8360
Linear IV					
Ordinary	10,000	1.2003	<i>0.0305</i>	1.1512	1.2506
Adjusted	10,000	0.9315	<i>0.0139</i>	0.9089	0.9546
FIML, tobit 1st stage					
Ordinary	10,000	1.109	<i>0.0205</i>	1.0753	1.1427
Adjusted (std. dev of x^*)	10,000	0.9541	<i>0.0143</i>	0.9308	0.9779
Adjusted (std. dev of x)	10,000	0.9112	<i>0.0127</i>	0.8902	0.9327

coefficient from 0.8165 (the OLS coefficient) to 0.9315. As expected, adapting their procedure for the tobit first stage produces a more conservative estimate of the lower bound (closer to the OLS), of 0.9112. This occurs because the standard deviation of observed hours of work is lower than that of latent desired hours of work, which are effectively being modelled in the tobit first stage, and this leads to a higher weight than assumed by the Nevo and Rosen estimator, being placed on the endogenous regressor in creating its second-stage fitted value. The alternative procedure, using the estimated standard deviation of latent desired employment, produces a tighter lower bound, but in the main body of this paper we choose the more cautious approach.

A.3 Seasonality of job vacancies for young adults

The month of birth instrument used in this paper is motivated by two conditions:

1. Employment opportunities increase discontinuously on reaching one's birthday. This is because the regulation of teenage employment conditions on age in completed years, rather than school year.
2. A disproportionate number of vacancies arise in the first few months of the school year (September-December), compared with later on. The intuition for this is that:
 - (a) Older cohorts of students give up their jobs at the end of their summer holidays as they leave for university or choose to focus on their more demanding requirements for the new academic year.
 - (b) New positions are created in the autumn, in anticipation of increased consumer demand prior to Christmas.

The former condition is necessary, but not sufficient, because if vacancies arise at a fairly even rate throughout the school year, those who are relatively young will be at no particular disadvantage. The second condition means that within a school cohort, those who have already had their birthday or whose birthday will occur sooner, will be ahead in the queue to fill these vacancies.

Figures A4 and A5 present some evidence on point (2) in graphical form. This come from data on the number of vacancies notified to the 'JobCentre Plus' (JCP) in each town in the UK each month from May 2002-April 2006, and again for May 2006-November 2012.¹¹ As explained in the main body of this paper, sectors G ("Retail and wholesale...") and H ("Hotels and restaurants") are particularly relevant for the employment of schoolchildren and young adults, so we selected just these two sectors. For each of these time periods, we kept only those towns with an unbroken monthly series (no missing or suppressed data).

Figure A4 shows the coefficients on 11 dummy variables for months February-December, for year- and town- fixed effect regressions of log monthly notifications to JCPs. The height of the

¹¹See <https://www.nomisweb.co.uk/query/construct/summary.asp?mode=construct&version=0&dataset=50>, accessed 8th March, 2016. We use separate series because of an administrative change in May 2006 which reduced the number of vacancies being notified retrospectively. These were occurring where specific positions were filled by a speculative JCP referral to the employer, before the position was formally notified to the JCP (See <https://www.nomisweb.co.uk/articles/177.aspx>, accessed 8th March 2016, for a more detailed explanation).

Figure A4: Notifications of vacancies to Job Centre Plus: Sector G: Wholesale and retail trade; repair of motor vehicles, motorcycles and personal and household goods. (Difference in logs compared over January of same year).

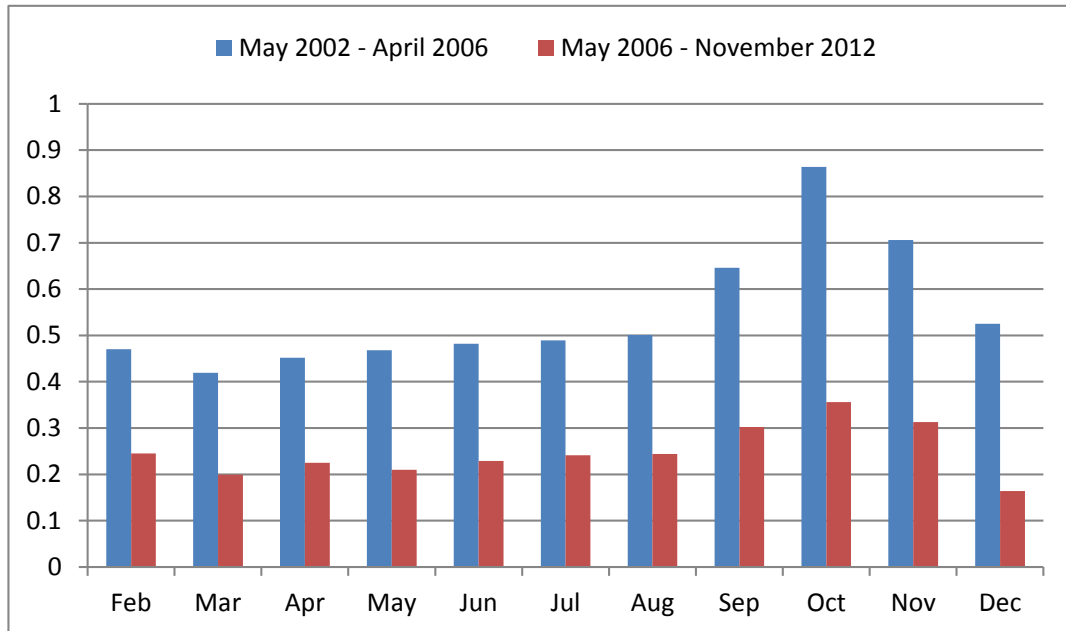
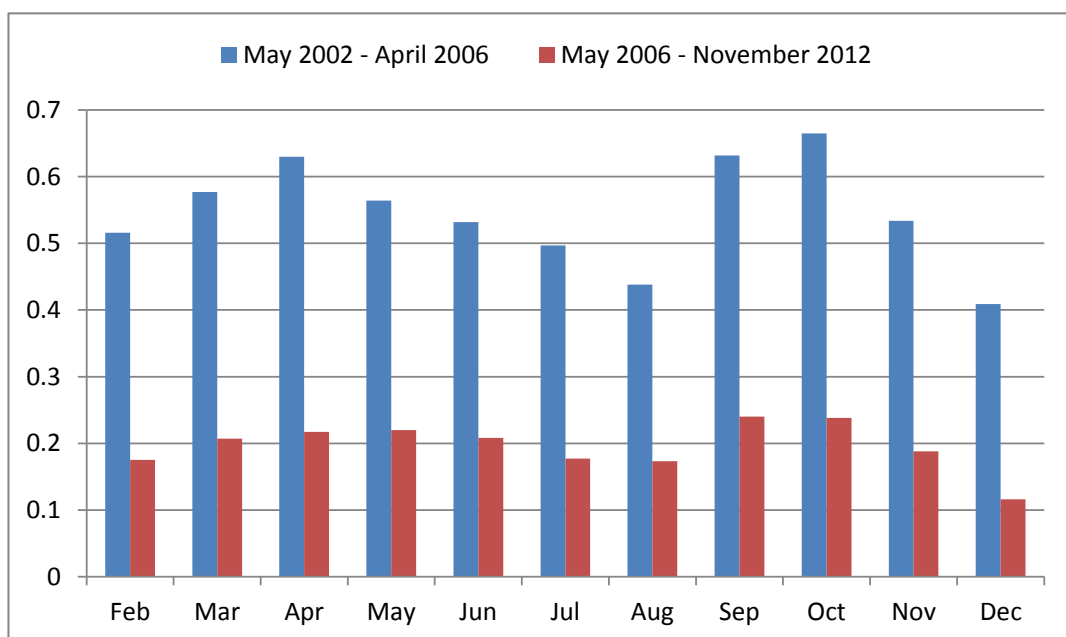


Figure A5: Notifications of vacancies to Job Centre Plus: Sector H: Hotels and restaurants. (Difference in logs compared over January of same year).



bar can thus be interpreted roughly as the proportional increase JCP notifications compared with January in the same year. In both series, there is a clear spike in September, October, and November. For the pre-May 2006 series, the volume of vacancies rises from about 45-50% above the January level to 63, 88 and 70% respectively. The differences are smaller in the post-May 2006 data, but the autumn spike is clearly still apparent.

Figure A5 repeats the exercise for Hotels and restaurants. The cycle of vacancy notifications here has twin peaks, in the autumn and spring, but this pattern is still sufficient to give those born earlier in the school year an advantage in filling those vacancies.

A.4 Distribution of latent inputs

Recall that the measurement model (equations 8-4.9, shown in Table 5) was estimated using the pooled sample of both waves and sexes. This ensures that a standard deviation change in each factor represents a common unit in terms of absolute activity levels, the interpretations of which are provided in Table 6. In turn, this means that the parameters representing the crowd-out or facilitation of each activity by employment, and their production function parameters, are directly comparable between waves and sexes.

Figures A6-A10 show the distribution of the principal component of each latent factor by wave and sex. Each histogram has 12 bins, and the four histograms within each figure share a common scale on both axes. The similarity or overlap between groups mean that the proposed changes in activity levels will represent a relevant and plausible adjustment for each group.

The distributions are most similar for Study (Figure A6) and Attitude (Figure A10), though in the former case girls are slightly more concentrated in the middle and less at the bottom of the distributed, and in the latter case, marginally more prevalent at the top. Differences are more apparent for Social Life; girls are slightly more frequent at the foot of the distribution, but also have a mass point at one standard deviation above the mean; and Active Leisure, where the distribution for boys is always to the right of that for girls (the modal density being just over one standard deviation higher), though there is considerable overlap between the sexes. Risky behaviours are highly positively skewed, though in a very similar way for boys and girls. For all groups the modal density is at no risky behaviours.

Figure A6: Distribution of principal component of Study by wave and sex

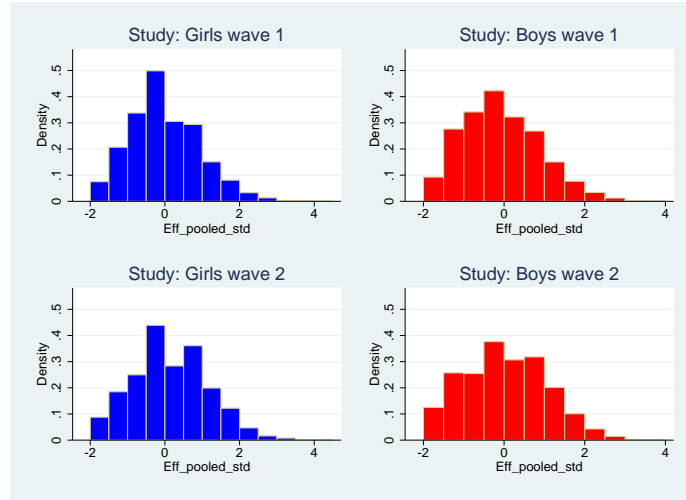


Figure A7: Distribution of principal component of Active Leisure by wave and sex

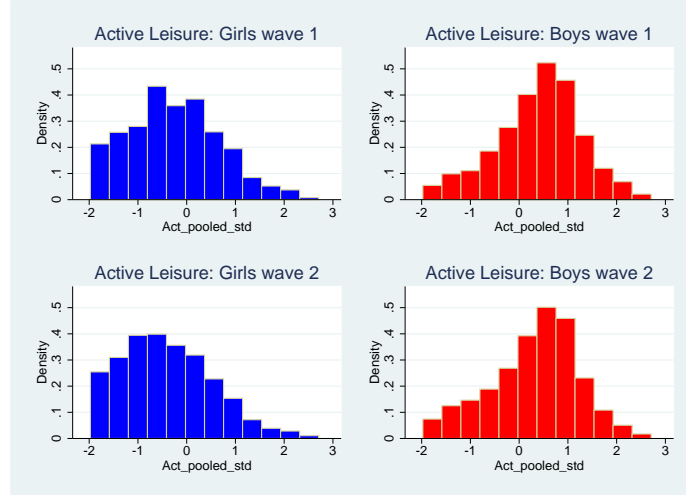


Figure A8: Distribution of principal component of Social Life by wave and sex

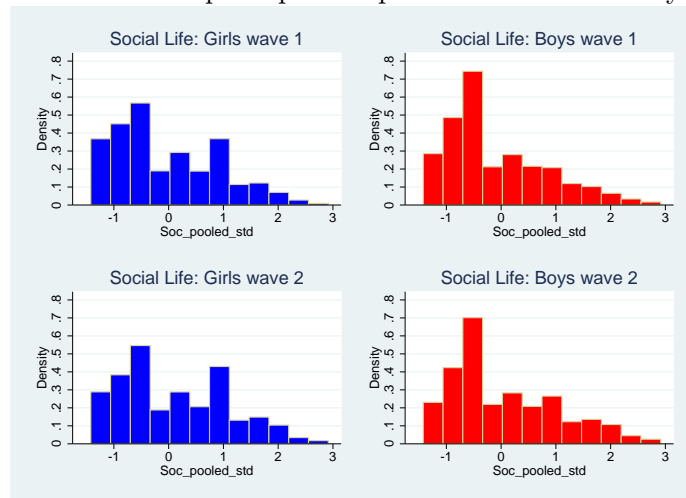


Figure A9: Distribution of principal component of Risky Behaviours by wave and sex

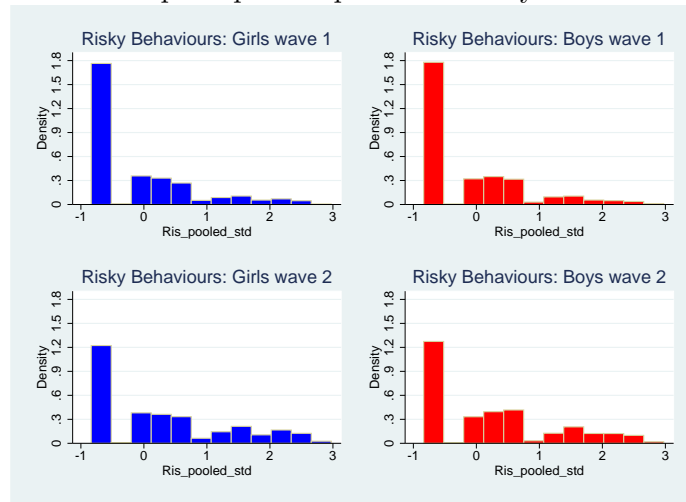
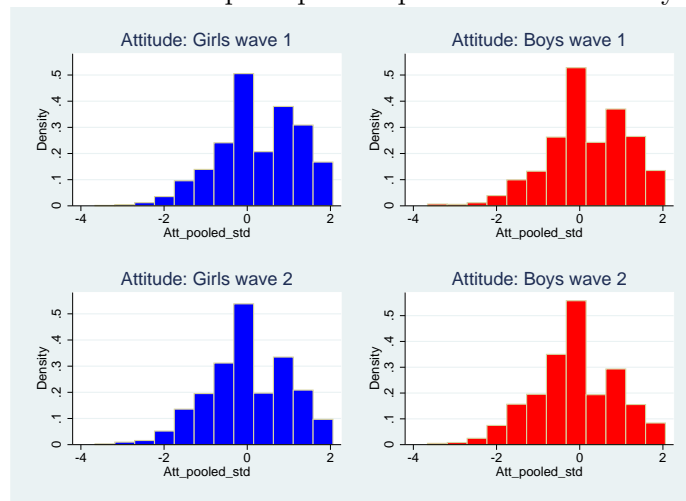


Figure A10: Distribution of principal component of Attitude by wave and sex



A.5 Time constraints and household responsibilities of boys and girls

A possible reason for the greater extent of crowd-out of study time by part-time work among girls might be that girls spend more time on ‘household production’ (caring for parents or completing household chores) than boys do. Reports on time spent on these activities are collected from the children at waves 1 and 2 of the LSYPE, and these data confirm that, in general, boys do spend less time on caring responsibilities and household chores than girls.

There are no significant differences in propensity to have caring responsibilities at age 14 (about 5% each) but girls edge ahead at age 15 (6.2 to 5.3%), by which girls with caring responsibilities take about 100 minutes more per week on average. (Among those with positive hours, the modal time spent is 2 hours and median is 4 hours, but there is a long right tail). Girls are also about 10 percentage points more likely to spend time on domestic chores (85 v.75%), and the difference in time spent on average is about 45 minutes per week from an average of 165 minutes.

It is also the case that those with a job spend less time on caring responsibilities and household chores than those without. However, Table A3, which presents some tobit regressions of time spent on just gender, employment status, and their interaction, shows that among those with a job, there is no significant gender difference in time spent on either of these activities. In some cases the sign is even reversed. This means we find no support for the idea that the difference in the rate of crowd-out of study is due to parents being less accommodating of part-time work (in terms of their expectations of contributions to the household) among girls than boys.

Nevertheless, these results are indeed consistent with boys being less time-constrained than girls: That taking on an additional activity, part-time paid work, has little or no bearing on time spent by boys on household caring or chores indicates that they enjoy enough leisure time for that to be their margin of adjustment. For girls, these results suggest that a portion of their time spent on household production is discretionary, and as with their study time, ripe for being crowded out.

Table A3: Time spent on household production by gender and employment status.

	Hours per week on caring responsibilities		Hours per week on domestic chores	
	Wave 1 (Age 14)	Wave 2 (Age 15)	Wave 1 (Age 14)	Wave 2 (Age 15)
Male	-0.710 (0.571)	-2.261*** (0.611)	-0.922*** (0.057)	-0.952*** (0.062)
Has a part-time job	-0.713 (1.006)	-2.872*** (0.938)	-0.244** (0.099)	-0.350*** (0.090)
Male × Has a part-time job	0.330 (1.345)	3.019** (1.292)	0.382*** (0.132)	0.528*** (0.125)
Observations	15,337	13,187	15252	13132

Notes: Tobit regression. No additional covariates.

A.6 Employment as a gateway to future employment

Table A4 shows the state dependence of current employment in waves 2 and 3 of the LSYPE on employment in the previous wave, or equivalently academic year. We show random-effects linear regression coefficients for hours of employment per week. These results show strong positive state dependence in employment. The best linear prediction is that each hour of employment per week in one year adds about 0.4 hours per week in the following year.

Table A4: State dependence in hours of employment

	Boys		Girls	
	No initial conditions	Wooldridge treatment	No initial conditions	Wooldridge treatment
Lagged Work Hours	0.400*** (0.027)	0.398*** (0.027)	0.365*** (0.024)	0.360*** (0.024)
Month of birth	-0.053*** (0.014)	-0.053*** (0.014)	-0.055*** (0.013)	-0.055*** (0.013)
Contemporaneous: LAD age 18-24 unemployment rate	-6.872*** (1.836)	20.843** (8.180)	-4.642** (1.905)	27.777*** (9.710)
Mean over time: LAD age 18-24 unemployment rate	.	-30.609*** (8.578)	.	-35.735*** (10.270)
Individuals	4772	4772	4695	4695
Observations	9359	9359	9261	9261

Notes: Longitudinal weights applied. *: $p < 0.1$; **: $p < 0.05$; ***: $p < 0.01$. Month-of-birth within academic year: Sept (oldest in year) = 1, Aug (youngest in year) = 12. **Additional controls:** Parent's income, parent's socio-economic status, parents' employment (plus time-varying mean), parent receives disability benefit (plus time-varying mean), housing situation (plus time-varying mean), child's prior educational performance, resident and non-resident siblings, lone parent family, child's special educational needs (SEN) classification, urban-rural classification, child's ethnicity, timing of interview.

In the main body of this paper, a static specification, accounting for employment at only one point in time, would ignore this dynamic dependence. The coefficient on hours of employment in wave 1 could not be interpreted as the direct causal effect of labour supply at that time, for example, because we would not hold constant the hours of employment undertaken in wave 2. Not only is this highly correlated, but, Table A4 shows, it is also causally dependent on employment at wave 1. In addition, in confirming the strong serial correlation in employment, these results give credence to our interpretation that the large negative coefficient on wave 2 employment for girls in our instrumental variables specifications is biased by the effects of activities crowded out by employment in close proximity to final high-stakes examinations.

These results are presented using specifications both with and without the Wooldridge (2005) treatment of initial conditions, which entails including the mean of time-varying regressors in the specification. This makes no difference to the inference relating to the state dependence of employment. This is to be expected, since the first wave of the LSYPE corresponds to the first full school year in which teenagers may enter the labour market. Hence, we observe the initial

condition. However, the high serial correlation of the local youth unemployment rate over the short period considered here leads to large and opposite-signed coefficients on the time-specific level and on the mean. The effect of the mean dominates, being larger in magnitude and negative. In the main body of this paper, for clarity about the source of identifying variation in employment, we control only for the contemporaneous level of the unemployment rate.