

IZA DP No. 5564

**Returns to Compulsory Schooling in Britain:  
Evidence from a Bayesian Fuzzy Regression  
Discontinuity Analysis**

Siddhartha Chib  
Liana Jacobi

March 2011

# Returns to Compulsory Schooling in Britain: Evidence from a Bayesian Fuzzy Regression Discontinuity Analysis

**Siddhartha Chib**

*Washington University*

**Liana Jacobi**

*University of Melbourne  
and IZA*

Discussion Paper No. 5564  
March 2011

IZA

P.O. Box 7240  
53072 Bonn  
Germany

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

### **Returns to Compulsory Schooling in Britain: Evidence from a Bayesian Fuzzy Regression Discontinuity Analysis**

In this paper we reevaluate the returns to education based on the increase in the compulsory schooling age from 14 to 15 in the UK in 1947. We provide a Bayesian fuzzy regression discontinuity approach to infer the effect on earnings for a subset of subjects who turned 14 in a narrow window around the policy change and whose schooling was affected by the policy change. Our approach and our results are quite different from previous work that has focused on large sets of cohorts and 2SLS based approaches and has reported positive earnings and wage effects of 5% and above. Our empirical analysis, using data from the UK General Household Surveys, yields considerably lower earnings and wage effects for the additional year of compulsory schooling than previous work. These findings are consistent with the implementation of the policy change that affected students at the lower end of the schooling distribution and did not lead students to acquire additional qualifications. The results add further evidence to a number of recent studies that have found no effect from this policy change on socio-economic outcomes correlated with earnings.

JEL Classification: C11, C21, I21

Keywords: Bayesian inference, causal effects, imperfect compliance, natural experiment, principal stratification, regression discontinuity, returns to schooling

Corresponding author:

Liana Jacobi  
Department of Economics  
University of Melbourne  
5th floor Economics and Commerce Building  
Victoria 3010  
Australia  
E-mail: [ljacobi@unimelb.edu.au](mailto:ljacobi@unimelb.edu.au)

# 1 Introduction

Changes in compulsory schooling laws have provided an important backdrop for understanding the returns to education (Angrist and Krueger, 1991; Kane and Rouse, 1995; Harmon and Walker, 1995, 1999; Acemoglu and Angrist, 2001). One empirical finding from such natural experiments is that the labor market effects of education appear to be lower in Europe compared to those in the US (Devereux and Hart, 2010; Pischke and von Wachter, 2008; Meghir and Paelme, 2005). An experiment that has attracted considerable attention in the recent literature is the increase in the minimum school leaving age from 14 to 15 in the UK in April 1947. Harmon and Walker (1995), Oreopoulos (2006), Oreopoulos (2008) and Devereux and Hart (2010) have exploited the regression discontinuity in the schooling intake that resulted from the change in the compulsory schooling policy to estimate the returns to schooling based on 2SLS methods. The first two of these studies report returns to schooling of well above 10%, while the latter two report estimates of around 5%.

In this paper we propose an alternative approach for finding the returns to education from the 1947 policy change. We provide a flexible Bayesian inferential framework based on the fuzzy regression discontinuity approach (Trochim, 1984; Lee and Lemieux, 2010) to infer the effect on earnings for a subset of subjects who turned 14 in a narrow window around the policy change and whose schooling was affected by the policy change. The fuzzy RD perspective is required because compliance with the new policy was imperfect. Our approach is motivated by the specific features of the 1947 policy change that was implemented within a tripartite secondary school system with streaming of students based on an ability test at age 11. The increase in the school leaving minimum age thus affected students at the lower end of the education distribution (leaving before age 16) enrolled in non-academic types of secondary schools (Galindo-Rueda and Vignoles, 2007; Lindeboom et al., 2009). We therefore define the educational attainment variable as binary that takes the value of 1 if a student leaves school at age 15, and 0 if the subject leaves school at 14 and focus our attention on subjects who had turned 14 near the time of the policy change. We take the observation window to run from

1946 to 1948. This also improves control for omitted variables that change smoothly over time and ensures that the control and treated subjects had faced a comparable set of circumstances. Within this setup, we calculate the effect of an additional year of schooling, leaving school at age 15 rather than at age 14, on later earnings, for subjects that complied with the policy change.

Our analysis yields very little evidence for any positive returns to education from the additional year of compulsory schooling, with earnings effects estimates for compliers around 1.6% and wage effects estimates around -0.3%. The findings appear reasonable in the context of the secondary school system in place at the time with the extra year imposed on subjects at the low end of the schooling distribution and not leading to any special or additional qualifications or certificates. While our estimates are below those reported in Oreopoulos (2008) and Devereux and Hart (2010), they are consistent with the findings from a recent paper by Grenet (2009). The paper provides empirical evidence based on two different schooling reforms in France and England that a positive earnings effect of compulsory schooling emerge when there is an increase in the qualifications obtained by students. Further, our results are also in accordance with a number of studies into the effect of the increase in compulsory schooling on other outcomes such as children's health outcomes and mortality (Lindeboom et al., 2009; Clark and Royer, 2007; Galindo-Rueda and Vignoles, 2007) that have also found very little evidence for any effects of the 1947 reform.

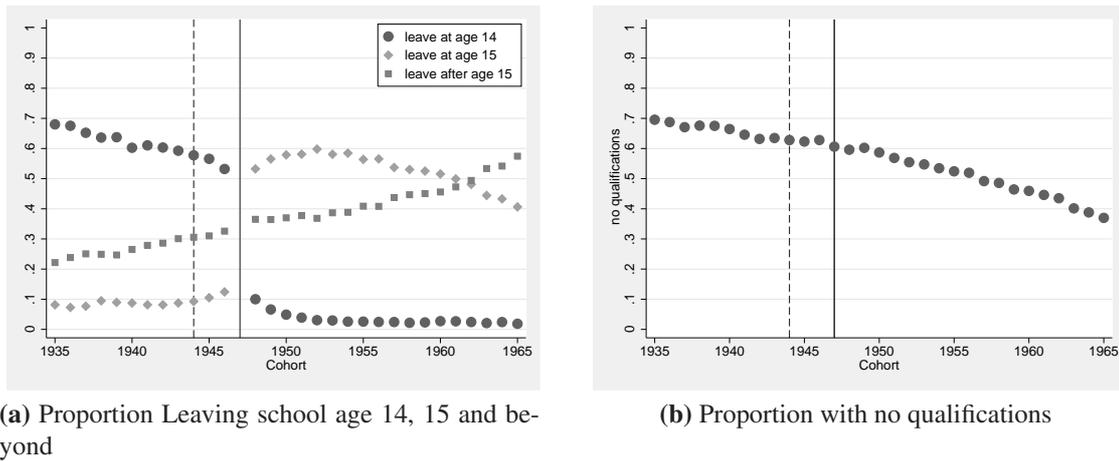
The rest of the paper is organized as follows. Section 2 provides more information about the effects of the compulsory schooling policy change and the schooling system. Section 3 introduces the identification strategy to isolate the returns to schooling based on the fuzzy RD approach, the flexible parametric framework for earnings and schooling and the Bayesian inferential framework to estimate the returns to schooling. In Section 4 we discuss the data and the construction of the sample for the RD analysis. The empirical results are presented in Section 5 and conclusions are given in Section 6.

## 2 Background

The increase in the minimum school leaving age from 14 to 15 in April 1947 was part of the 1944 Education Act that reorganized the secondary school system with the aim to make secondary education available to all students. The Education act introduced free secondary education for all students and a tripartite system with grammar schools, secondary modern schools and technical schools. Admission to the academically oriented grammar schools was based on a test administered at age 11. Students not admitted to a grammar school attended a secondary modern school with a small portion transferring to a technical school at ages 12 or 13 (Halsey and Ridge, 1980). These schools provided lower level academic education, including for those students leaving school at the minimum school leaving age (Lindeboom et al., 2009).

Introduced in this setting, the 1947 increase in the school leaving minimum age resulted mainly in an immediate increase in the proportion of students that stayed an extra year in non-academic type of secondary schools (Galindo-Rueda and Vignoles, 2007). Figure 1 shows the school leaving behavior of cohorts of students turning 14 between 1932 and 1965. Graph (a), based on subjects in the UK General Household surveys, shows the proportion of students in a cohort that left school at age 14, 15 or beyond. The solid vertical line in 1947 refers to the increase in the school leaving age and the dashed line to the introduction of the tripartite schooling system and access to free secondary education. We observe a jump in the proportion of students leaving school at age 15 from the 1946 to 1948 cohorts of 14 year old subjects, that is mirrored by a sharp decline in the proportion of students leaving at age 14. For example, in the cohort of students who turned 14 in 1946, over 50% dropped out of school at age 14, and only around 15% at age 15. In comparison, over 50% of the students that were 14 in 1948 dropped out at age 15, while roughly 10% dropped out at age 14. In contrast, the graph indicates a smooth trend for the proportion of subjects leaving school after the age of 15.

At the time of the policy change students could take two external examinations, the school certificate and the higher school certificate. These were usually taken by grammar schools



**Figure 1:** Schooling and qualifications of students leaving school at age 14, age 15 and beyond in the 1935 to 1965 cohorts.

students at age 16 and age 18 and replaced by the general certificate for education (0-level, A-level) in 1951. Students completing the full course (4 years before the policy change and 5 years afterwards) could obtain a school leaving certificate. However, secondary modern students often left school at the earliest possible moment without taking a school certificate (Halsey and Ridge, 1980). Thus it is not surprising that the increase in the school leaving minimum age which affected students at the low end of the schooling distribution did not lead to an increase in qualifications. This is confirmed by the graph in panel (b) in Figure 1 which shows that the proportion of students without any qualifications declined smoothly over time, showing no discontinuity around the policy change.

### 3 Framework

Our aim is to identify the effect of an additional year of schooling, for those leaving school at age 15 rather than at age 14, on later earnings by exploiting the change in the schooling intake of students induced by the policy change. In this section we set up a Bayesian framework to estimate the effect of staying in secondary school for one additional year, focussing on students who turned 14 in the window between 1946 and 1948. Let  $T$  be a continuous variable that represents the year and quarter a student turned 14 in the window between 1946 and 1948

and let  $\tau$  denote the 2nd quarter of 1947 when the new policy was introduced. Further, define  $z^* = (T - \tau)$  as the distance measured in quarters between when a student turned 14 and when the new policy was introduced. The indicator variable  $z = I[z^* > 0]$  assigns a subject in the window to a policy regime. It is one if a student who turned 14 faced the new policy and 0 if a student who turned 14 faced the old policy. Specifically,

$$z = \begin{cases} 1 & \text{if min schooling leaving age is 15 when subject turned 14} \\ 0 & \text{if min schooling leaving age is 14 when subject turned 14} \end{cases}$$

Now define the binary schooling variable  $x$  for subjects that turned 14 between 1946 and 1948 as

$$x = \begin{cases} 1 & \text{if subject dropped out of school at age 15} \\ 0 & \text{if subject dropped out of school at age 14} \end{cases}.$$

As discussed in the introduction, the policy change altered the probability distribution of  $x$  around the time of the change in policy. If we let  $z^* = 0^+$  denote the value of  $z^*$  just to the right of  $\tau$  and  $z^* = 0^-$  denote the value of  $z^*$  to the left of  $\tau$ , then we observe a discontinuity in the probability distribution of  $x$  around the threshold, ie.,

$$\Pr(x = 1|z^* = 0^+) \neq \Pr(x = 1|z^* = 0^-) \quad (3.1)$$

as shown in panel (a) in Figure 1. The graph also shows that not all students, however, complied with the policy change. In particular, some students under the old policy regime left school at age 15, while some students under the new policy regime left school at age 14.

In such a fuzzy RD design (Lee and Lemieux, 2010), where  $\Pr(x = 1|z^* = 0^-) < 1$ , we suppose that  $x$  is determined by  $z$  and an additional unobserved variable that is correlated with earnings. Borrowing from the statistical literature on principal stratification and the econometrics literature on LATE (Sommer and Zeger, 1991; Imbens and Angrist, 1994), we model this unobserved confounder as an exogenous discrete variable that reflects compliance behavior and is defined as follows: compliers (C) for subjects who comply with the policy in place, always drop-out 14 (AD14) for subjects who always leave school at age 14 regardless of the policy in place, and always drop-out 15 (AD15) for subjects who always leave at age 15 regardless of the policy in place. In what follows, subject type is denoted by  $s = k$ ,

where  $k = \{C, AD14, AD15\}$ . The unknown probabilities of subject types are denoted by  $\Pr(s = k) = p_k$ .

Under the definition of the types, the schooling intake is determined by policy regime and subject type as

$$x = \begin{cases} 0 & \text{if } z = 0 \text{ and } s = AD14 \text{ or } s = c \\ 0 & \text{if } z = 1 \text{ and } s = AD14 \\ 1 & \text{if } z = 0 \text{ and } s = AD15 \\ 1 & \text{if } z = 1 \text{ and } s = AD15 \text{ or } s = C \end{cases}$$

It follows directly that compliers are observed under the two different schooling levels, whereas subjects of the other two types are only ever observed under one schooling level. It is easily seen that the discontinuity in expression (3.1) holds if some subjects comply with the policy change. By simple calculation it follows that  $\Pr(x = 1|z^* = 0^+) = q_{AD15} + q_C$  and  $\Pr(x = 1|z^* = 0^-) = q_{AD15}$ . Thus, condition (3.1) holds if  $q_C > 0$ .

The type definition also implies that a subject's observed schooling intake given the minimum school leaving policy in place provides information about the subject's type, although type is not known for all subjects. As shown in Table 1, those who leave school at age 15

Policy Indicator	Schooling Intake	
	$x = 0$ (leave age 14)	$x = 1$ (leave age 15)
$z = 0$ (old policy)	C, AD14	AD15
$z = 1$ (new policy)	AD14	C, AD15

**Table 1:** Distribution of a subject's type by observed policy regime and schooling intake

under the old policy regime are always drop-out 14 and those who leave at age 14 under the new policy regime are always drop-out 15. Both types also appear in a diagonal cell where we also observe compliers.

Now let  $\{y_{0,C}, y_{0,AD14}, y_{1,C}, y_{1,AD15}\}$  be the four potential outcomes that refer to the earnings of a subject of a given type under the possible schooling levels. Also let  $\mathbf{w}$  denote any observed control variables that affect the distribution of these potential outcomes. Then, under the condition that  $y_{0,C}$  and  $y_{1,C}$  are smooth functions around the threshold  $\tau$ , it can be shown

that the average treatment effect for compliers, with  $T$  close to  $T^*$ , is identified and given by

$$E[y_{1,C} - y_{0,C} | \mathbf{w}, s = c] = \frac{E[y | \mathbf{w}, z^* = 0^+] - E[y | \mathbf{w}, z^* = 0^-]}{\Pr(x = 1 | z^* = 0^+) - \Pr(x = 1 | z^* = 0^-)} \quad (3.2)$$

In the next Section we specify a parametric framework to estimate this effect.

### 3.1 Parametric Specification

Our approach is to estimate the Complier Average Earnings Effect from equation (3.2) in a Bayesian parametric RD framework of flexible log earnings models from a sample of males. For each subject  $i$  in the sample, let  $y_i$  denote the subject's observed log earnings. Following the approach in Chib and Jacobi (2008) we model the observed earnings and schooling intake in terms of the potential earnings models  $p_j(y_i | z_i = l, s_i = k)$  and the type probability  $\Pr(s_i = k)$  as

$$p(y_i, x_i = j | z_i = l) = \sum_{k \in K_{lj}} \Pr(s_i = k) p_j(y_i | z_i = l, s_i = k), \quad j = 0, 1; \quad l = 0, 1$$

where  $s_i$  is the subject's type and  $K_{lj}$  the set of possible values that  $s_i$  can take given the observed values of the schooling and policy indicators. From the previous discussion (Table 1) it follows that  $K_{00} = \{C, AD14\}$ ,  $K_{01} = \{AD15\}$ ,  $K_{10} = \{AD14\}$ ,  $K_{11} = \{C, AD15\}$ .

To allow for outliers common in earnings data we allow for student-t errors in the log earnings models rather than using the standard but more restrictive normality assumption. We formulate two linear regression models for the potential earnings of compliers under the two different schooling intakes, and one model each for always drop-outs 14 under  $x_i = 0$  and always drop-outs 15 under  $x_i = 1$ :

$$\begin{aligned} p_j(y_{ji,C} | \mathbf{w}_i, s_i = C, z = l) &= t_\nu(\mathbf{w}'_i \boldsymbol{\beta}_{j,C}, \sigma_{j,C}^2), \quad j = 0, 1 \\ p_0(y_{0i,AD14} | \mathbf{w}_i, s_i = AD14, z = l) &= t_\nu(\mathbf{w}'_i \boldsymbol{\beta}_{0,AD14}, \sigma_{0,AD14}^2) \\ p_1(y_{1i,AD15} | \mathbf{w}_i, s_i = AD15, z = l) &= t_\nu(\mathbf{w}'_i \boldsymbol{\beta}_{1,AD15}, \sigma_{1,AD15}^2) \end{aligned} \quad (3.3)$$

where vector  $\mathbf{w}_i$  refers to a set of control variables and  $\nu$  to the degrees of freedom of the student-t distribution.

We note that in this model specification the effect of schooling on earnings is modeled flexibly through an effect on the intercept as well as the slope coefficients and the variances. The causal effect of interest is the average causal earnings effect for a subject who is a complier and is given by

$$E[y_{1C}|s = C, \mathbf{w}] - E[y_{0C}|s = C, \mathbf{w}] = \mathbf{w}'\boldsymbol{\beta}_{1,C} - \mathbf{w}'\boldsymbol{\beta}_{0,C} \quad (3.4)$$

Under the assumptions of the model, this effect is identified.

### 3.2 Estimation and Model Comparison

Our parametric assumptions from the previous section lead to a likelihood of the observed earnings and schooling intake vectors,  $\mathbf{y} = \{y_1, \dots, y_n\}$  and  $\mathbf{x} = \{x_1, \dots, x_n\}$ , given the observed assignments and covariate vectors  $\mathbf{z} = \{z_1, \dots, z_n\}$  and  $\mathbf{W} = \{\mathbf{w}_1, \dots, \mathbf{w}_n\}$ , that takes the form of a mixture of flexible student-t distributions:

$$\begin{aligned} f(\mathbf{y}, \mathbf{x}|\boldsymbol{\theta}, \mathbf{W}, \mathbf{z}) &= \prod_{i \in I_{00}} [p_C t_\nu(y_i|\mathbf{w}'_i\boldsymbol{\beta}_{0,C}, \sigma_{0,C}^2) + p_{AD14} t_\nu(y_i|\mathbf{w}'_i\boldsymbol{\beta}_{AD14}, \sigma_{AD14}^2)] \\ &\times \prod_{i \in I_{01}} p_{AD15} t_\nu(y_i|\mathbf{w}'_i\boldsymbol{\beta}_{AD15}, \sigma_{AD15}^2) \\ &\times \prod_{i \in I_{10}} p_{AD14} t_\nu(y_i|\mathbf{w}'_i\boldsymbol{\beta}_{AD14}, \sigma_{AD14}^2) \\ &\times \prod_{i \in I_{11}} [p_{AD15} t_\nu(y_i|\mathbf{w}'_i\boldsymbol{\beta}_{AD15}, \sigma_{AD15}^2) + p_C t_\nu(y_i|\mathbf{w}'_i\boldsymbol{\beta}_{1,C}, \sigma_{1,C}^2)] \end{aligned}$$

where  $\boldsymbol{\theta} = \{\boldsymbol{\beta}, \boldsymbol{\sigma}^2, \mathbf{p}\}$  denotes the parameter vector, where  $\boldsymbol{\beta} = (\boldsymbol{\beta}_{C0}, \boldsymbol{\beta}_{C1}, \boldsymbol{\beta}_{AD14}, \boldsymbol{\beta}_{AD14})$ ,  $\boldsymbol{\sigma}^2 = (\sigma_{C0}^2, \sigma_{C1}^2, \sigma_{AD14}^2, \sigma_{AD15}^2)$  and  $\mathbf{p} = \{p_C, p_{AD14}, p_{AD15}\}$  and  $I_{lj}$  refers to the set of subjects with  $z_i = l$  and  $x_i = j$ .

Here we proceed under the Bayesian inferential approach to estimate the model parameters and the complier average treatment effect. The posterior distribution of the model parameters is proportional to the product of the likelihood and the prior. We assume flexible proper prior distributions of the following form

$$\pi(\boldsymbol{\theta}) = Dir(\mathbf{p}|\alpha_{C,0}, \alpha_{AD14,0}, \alpha_{AD15,0}) \prod_{j=0}^1 \prod_{k \in K_j} \mathcal{N}_p(\boldsymbol{\beta}_{kj,0}|\boldsymbol{\beta}_{kj,0}, B_{kj,0}) \mathcal{IG}(\boldsymbol{\sigma}_{kj,0}^2|\nu_{kj,0}, \delta_{kj,0}) \quad (3.5)$$

where  $K_0 = \{C, AD14\}$  and  $K_1 = \{C, AD15\}$ . The prior parameters are specified with the help of an additional sample of male cohorts from the 1979 to 1998 UKGHS surveys just outside the 1946-1948 window, those turning 14 either in the 1944-45 period or the 1949-1950 period. For a sensitivity check we repeat the empirical analysis under alternative prior assumptions.

To simulate the posterior distribution by efficient Markov chain Monte Carlo (MCMC) methods we include the unobserved compliance type indicators into the parameter space, but not the unobserved counterfactuals, following Chib (2007). From the previous discussion it follows that for subjects with  $(x_i = z_i = 0)$  the type variable can take the value of either complier or always drop-out 14, while subjects with  $(x_i = z_i = 1)$  can be either compliers or always drop-out 15. We define the vectors  $\mathbf{s}_{00} = \{s_i : x_i = z_i = 0\}$  and  $\mathbf{s}_{11} = \{s_i : x_i = z_i = 1\}$ , where  $s_i = 1$  denotes that the subject is a complier. The posterior distribution of the parameters is then simulated by MCMC methods as proposed in Chib and Jacobi (2008) and described in detail in the appendix.

We estimate the complier average treatment effect by calculating the posterior distribution of  $\mathbf{w}'\beta_{1,C} - \mathbf{w}'\beta_{0,C}$  from the sampled draws of the parameters from our MCMC procedure. At each MCMC iteration, we compute and store the average difference in the potential earnings for all subjects who are classified as compliers.

In our empirical analysis we consider a range of specifications with different covariate vectors and different degrees of freedom. We use the model marginal likelihood to pick the model that best fits the data. The marginal likelihood of each model is estimated via the Chib method (Chib, 1995) by evaluating the following expression

$$\ln m(\mathbf{y}, \mathbf{x}) = \ln f(\mathbf{y}, \mathbf{x}, \mathbf{W}, \mathbf{z} | \boldsymbol{\theta}^*) + \ln \pi(\boldsymbol{\theta}^*) - \ln \pi(\boldsymbol{\theta}^* | \mathbf{y}, \mathbf{x}, \mathbf{W}, \mathbf{z})$$

where  $\boldsymbol{\theta}^*$  refers to the vector of the posterior means of the parameters. The likelihood and the prior ordinate are calculated directly and the posterior ordinate is estimated using the output of the full and reduced MCMC runs of the MCMC algorithm (see the appendix for details).

## 4 Data

The data for our analysis comes from the UK General Household Surveys (UKGHS) which is a panel of cross-section surveys. The UKGHS is a continuous survey that has been carried out by the Social Survey Division of the Office for National Statistics (ONS) since 1971 to collect information on a range of topics from people living in private households in Great Britain. Following Devereux and Hart (2010) we exclude the pre-1979 surveys to avoid problems due to a different reporting scheme of earnings in the early surveys. The collected data includes information on birth-dates, education, income and employment.

### *A. Construction of Sample and Key Variables*

Following our previous discussion, we focus on subjects from the 1979 to 1998 surveys that turned 14 between 1946 and 1948, e.g. born between 1932 and 1934. The 1986 to 1995 and 1998 surveys report both birth year and birth months for each respondent and we simply assign a value of zero to  $z_i$  for subjects who turned 14 between January 1946 and March 1947 (born between January 1932 and March 1933), and a value of one for subjects who turned 14 between April 1947 and December 1948 (born between April 1933 and December 1934). Subjects in the 1986 to 1995 surveys and the 1998 survey that turned 14 before 1946 and after 1948 are excluded. For subjects in the remaining surveys information on birth year and month are not available. We utilize the reported age together with the survey year and month and construct a one year interval for when the subject turned 14. If the upper bound of this interval is at or below March 1947, then the policy regime indicator of the subject is set at zero. If the lower bound of the interval is at or above April 1947, the subject is assigned the value one for the policy indicator. Subjects for whom the interval includes April 1947 are dropped from the sample as we are not able to determine the policy regime from the available data. Subjects outside the window, those with the upper bound of the interval below January 1946, and those with the lower bound above December 1948 are also excluded.

To define the binary schooling intake variable  $x_i$  for each subject we again follow previous

work and use information on the reported age when a subject left school. We assign a value of zero to the schooling intake if a subject reported leaving school at age 14, and a value of one if the subject reported leaving school at age 15. We exclude subjects from the sample that reported a school leaving age above 15 and also the few subjects that reported leaving before age 14. We further exclude the few subjects who report an age above the school leaving age for when they left full-time education.

For the definition of the earnings variable  $y_i$  we follow Devereux and Hart (2010) and construct a variable for real log weekly earnings based on the reported gross weekly earnings (including earnings from self employment), deflated by the UK retail price index with base year 1998. We also construct a hourly wage variable based on the weekly earnings and the reported work hours. We exclude subjects from the pre-1979 surveys due to different earnings measures based on the year prior to the survey, rather than the earnings based on the week before the survey. Further, to avoid retirement related issues, we omit subjects age 60 and older. As shown in Banks and Blundell (2005) many low-skilled men quit working before the age of 65 in this time period, leading to an employment rate of low-skilled men age 60-64 of about 40%. Following Devereux and Hart (2010) we also drop subjects whose weekly working hours is missing or above 84 hours.

### *B. Sample Features*

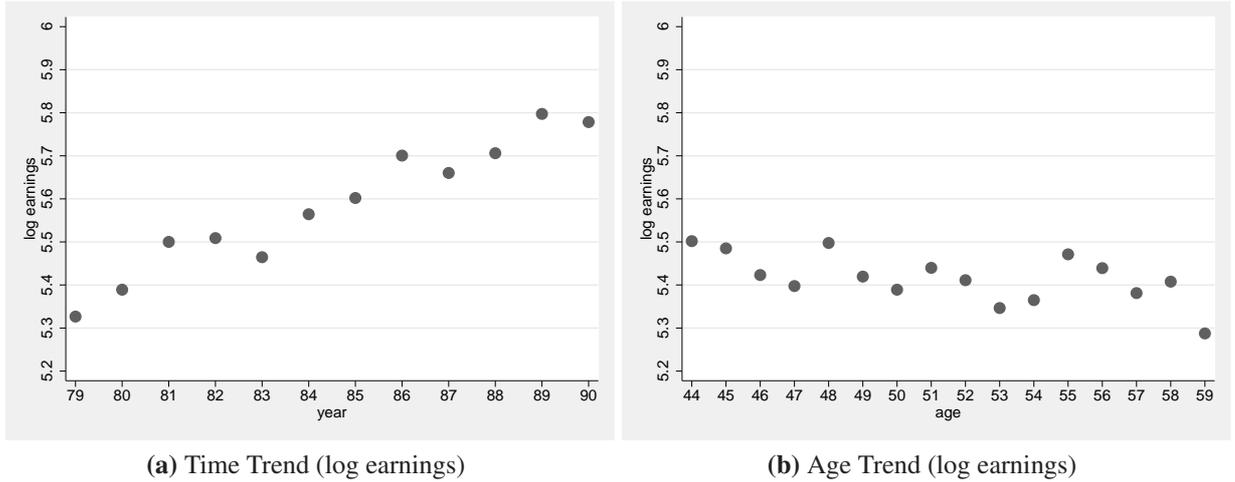
Restricting our attention to males we obtain a sample of 1,935 subjects. Of these, 987 turn 14 under the old policy regime with 807 subjects leaving school at age 14 and 180 at age 15. From the 848 subjects that turned 14 under the new school leaving minimum age of 15, 163 subjects dropped out at age 14 and the remaining 787 at age 15. Table 2 provides some descriptive statistics for the sample. The subjects are between the ages of 44 and 59 and come from the 1979 to 1994 surveys.

Given our focus on subjects within a narrow window of the policy change, a respondent's age and the survey year when earning are reported are highly correlated. By construction, we only observe subjects from three different cohorts in each survey year with subjects under the

Summary Statistics for Male Sample						
	Variable					
Description		[short]	Mean	St.Dev.	Min	Max
Leaving School at Age 15	[x]		0.50		0	1
Under new Policy	[z]		0.49		0	1
Age	[age]		52.22	4.17	44	59
Married	[married]		0.59		0	1
survey year*	[year]		85.35	4.05	79	94
Log Weekly Earnings**	[learn]		5.54	0.51	1.87	7.56
Log hourly Wage**	[lwage]		1.85	0.49	0.03	3.56
work hours			41.41	9.10	3	81

**Table 2:** \*Calender Years are reported as (year - 1900).\*\* Real earnings in 1998 pounds.

new policy being slightly younger. The average age of subjects with  $x = 1$  is 51, which is 1.5 year lower than that of subjects with  $x = 0$ . Hence we can only include controls for either age or survey year in our earnings models, and not controls for both as it is commonly done in the empirical literature. Since the subjects in our sample are mature employees of age 44 (and holding mainly blue collar jobs) we would expect a flat earnings profile and include controls for survey year. As earnings are observed over a 16 year time period from 1979 to 1994, we would expect a strong time effects so we include controls for survey year. The graphs for time and age trends in the general UKGHS sample of males from the 1935 to 1965 in Figure 2 provide support for these covariate adjustments. We observe strong systematic upward trends in the earnings, but only a weak negative age trend in earnings. The same is true for log wages. Since in a given survey year subjects that comply with the minimum school leaving age and leave school at age 15 (under the new policy) are younger than those leaving school with age 14 (under the old policy), we cannot rule out a small upward bias in our earning effects estimates due to a small negative age effect.



**Figure 2:** Average weekly earnings for males leaving school age 14 and 15 for age 50 in different survey years (a), and average earning by age in the 1980 survey (b) for males in 1935 to 1965 cohorts.

## 5 Results

In this section we present the results from the analysis of earnings for the sample of males from the 1946 to 1948 cohorts described in the previous section. Following previous work we implement the analysis for both log weekly earnings and log hourly wages. We consider two different specifications for the covariate vector in the earnings model, one with a control for survey year (S1) and one with an additional control for marital status (S2). The former is comparable to previous work although we need to exclude age as discussed in Section 4. Instead of proceeding with the analysis by setting the degrees of freedom for the student-t log earnings distributions at one particular level, we fit the models under a range of values for the degrees of freedom and as well under the Normality assumption. The models are then compared based on their (log) marginal likelihoods.

The upper panel in Table 3 gives the log marginal likelihoods for the various model fit with the two dependent variables, log weekly earnings and log hourly wages, respectively. According to the marginal likelihood criterion, the data exhibit thick tails as the models with the lowest degrees of freedom are preferred.

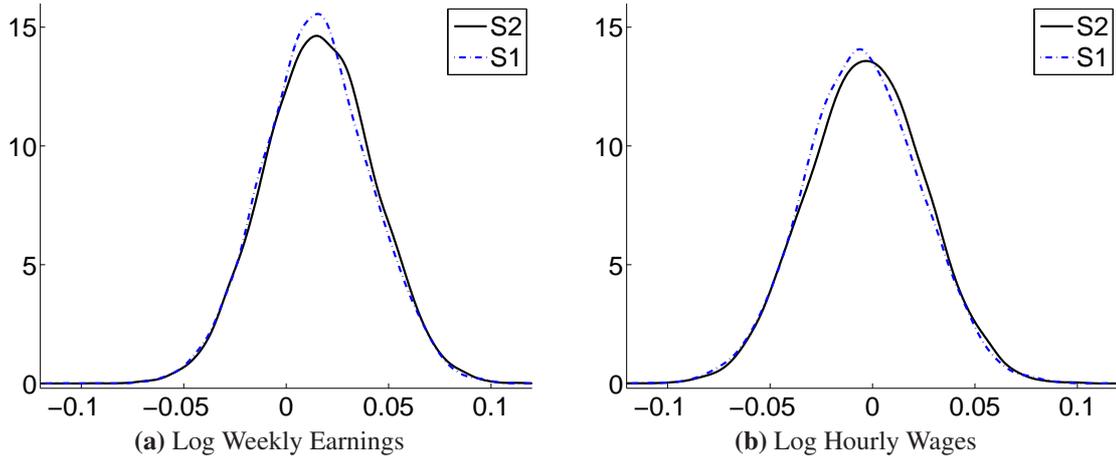
In the lower panel of Table 3 we present the estimates of the parameters for the earnings

Model Comparison and Model Parameter Estimates								
<i>Log Marginal Likelihoods</i>								
dof /spec.	5/1	5/2	10/1	10/2	20/1	20/1	$\infty/1$	$\infty/2$
Earn	-2004.42	<b>-1990.96*</b>	-2031.60	-2018.39	-2044.12	-2030.96	-2061.49	-2052.38
Wage	-2066.26	<b>-2056.06*</b>	-2074.33	-2064.80	-2083.75	-2075.16	-2091.78	-2085.09
<i>Posterior Means (Stdv) for Earnings Model (5/2)</i>								
	$C(x = 0)$	$C(x = 1)$	AD14	AD15				
$\beta$	4.390 (0.464)	4.929 (0.471)	6.133 (0.848)	5.424 (0.913)				
	0.013 (0.006)	0.007 (0.006)	-0.008 (0.010)	0.002 (0.011)				
	0.140 (0.045)	0.120 (0.049)	0.167 (.077)	0.159 (0.084)				
$\sigma^2$	0.094 (0.010)	0.115 (0.019)	0.106 (0.010)	0.143 (0.019)				
$p_k$		0.644 (0.017)	0.173 (0.012)	0.183 (0.012)				

**Table 3:** Estimated log marginal likelihoods for various degrees of freedom and covariate specifications (\* Model preferred based on marginal likelihood criterion) and model parameter estimates for preferred earnings model. Results are based on the draws from the MCMC sampler with 10,000 iterations and 1000 burn-in iterations.

models defined in equation (3.3) for the preferred model that has student-t errors, 5 degrees of freedom and covariate specification 2. Columns (2) and (3) provide the means and standard deviations for the parameters in the earnings models for compliers leaving school at ages 14 and 15. We observe small differences in the posterior means of the intercept and slope coefficients. In comparison, we observe larger differences in the coefficient estimates between compliers and the AD14 and AD15 subjects. The estimated type probabilities are 0.64 for compliers and 0.17 and 0.18 for AD14 and AD15, respectively.

To evaluate the causal effect of an extra year of schooling for compliers we estimate the complier average causal earnings and wage effects. Figure 3 shows the kernel-smoothed graphs of the posterior densities of the complier average returns to schooling for log weekly earnings in Panel (a) and hourly wages in Panel (b) for the model specifications with the highest marginal likelihoods. For log weekly earnings the densities are centered slightly to the



**Figure 3:** Kernel-smoothed densities of the posterior density of CATE for specifications 1 and 2 under 5 degrees of freedom for earnings and wages.

right of zero while that those for hourly wages are centered slightly to the left of zero.

The estimates of the returns to schooling in terms of the posterior means of the CATE are presented in the upper part of Table 4. For the winning specification with 5 degrees of freedom the average returns to schooling for compliers is 1.6% in terms of earnings and -0.3% for hourly wages. Across the various specifications the estimates for the complier average

Estimates of Average Returns to Schooling					
D.o.F		CATE: Mean and Stand.Dev.			
		5	10	20	$\infty$
Earnings	S1	0.015 (0.026)	0.008 (0.025)	0.015 (0.023)	0.019 (0.022)
	S2	0.016 (0.027)	0.010 (0.025)	0.017 (0.023)	0.019 (0.022)
Wages	S1	-0.005 (0.028)	0.003 (0.030)	-0.009 (0.029)	0.001 (0.027)
	S2	-0.003 (0.028)	0.000 (0.030)	-0.008 (0.028)	0.002 (0.026)

**Table 4:** Posterior means and standard deviations for complier average treatment effect based on the fittings of the various model specifications.

earnings effects vary between 0.8% and 1.9% and the complier average wage effects between -0.9% and 0.3%. Overall these estimates point to very small positive returns to schooling in terms of earnings and even smaller and slightly negative effects for wages. To further explore

the robustness of the findings we have redone the above analysis using a set of uninformed priors. The results (see Appendix) confirm the conclusions from our main analysis.

In summary, our analysis suggests returns to schooling from the increase in the compulsory school leaving age from 14 to 15 in the UK in 1947 below those reported in Oreopoulos (2008) and Devereux and Hart (2010). However, these low estimates are to be expected given the secondary education system at the time of the raise in school leaving minimum age that is exploited here to estimate the earnings and wage effects of an additional year of education. The 1947 increase in the school leaving minimum age resulted mainly in an immediate increase in the proportion of students that stayed an extra year in non-academic type of secondary schools (Galindo-Rueda and Vignoles, 2007) and did not raise the level of qualifications held by students.

## **6 Concluding Remarks**

In this paper we reevaluate the returns to education based on the increase in the compulsory schooling age from 14 to 15 in the UK in 1947. Our approach and our results are quite different from previous work that has focused on large sets of cohorts and 2SLS based approaches and has reported positive earnings and wage effects of 5% and above. Our estimates point to at most small earnings effects of 1.5% and below and suggest no positive wage effects from the additional year of schooling resulting from the policy change. These findings are consistent with the implementation of the policy change that affected students at the lower end of the schooling distribution and did not lead students to acquire additional qualifications given the school system in place. Further, our results add further evidence to a number of recent studies that have found no effect from this policy change on socio-economic outcomes that are considered to be correlated with earnings.

## 7 Appendix

### 7.1 MCMC Algorithm for Prior-Posterior Analysis

We simulate the joint posterior distribution for the augmented parameter space  $\theta, s_{00}, s_{11}, \lambda$  which takes the form

$$\begin{aligned} \pi(\theta) & \prod_{i=1}^n \mathcal{G}(\lambda_i | \frac{\nu}{2}, \frac{\nu}{2}) \\ & \prod_{i \in I_{00}} [I[s_i = C] p_C \mathcal{N}(y_i | \mathbf{w}'_i \boldsymbol{\beta}_{C0}, \lambda_i^{-1} \sigma_{C0}^2) + I[s_i = AD14] p_{AD14} \mathcal{N}(y_i | \mathbf{w}'_i \boldsymbol{\beta}_{AD14}, \lambda_i^{-1} \sigma_{AD14}^2)] \\ & \prod_{i \in I_{01}} [I[s_i = AD15] p_{AD15} \mathcal{N}(y_i | \mathbf{w}'_i \boldsymbol{\beta}_{AD15}, \lambda_i^{-1} \sigma_{AD15}^2)] \\ & \prod_{i \in I_{10}} [I[s_i = AD14] p_{AD14} \mathcal{N}(y_i | \mathbf{w}'_i \boldsymbol{\beta}_{AD14}, \lambda_i^{-1} \sigma_{AD14}^2)] \\ & \prod_{i \in I_{11}} [I[s_i = C] p_C \mathcal{N}(y_i | \mathbf{w}'_i \boldsymbol{\beta}_{C1}, \lambda_i^{-1} \sigma_{C1}^2)] + \prod_{i \in I_{11}} [I[s_i = AD15] p_{AD15} \mathcal{N}(y_i | \mathbf{w}'_i \boldsymbol{\beta}_{AD15}, \lambda_i^{-1} \sigma_{AD15}^2)] \end{aligned}$$

where  $\theta = \{\boldsymbol{\beta}, \sigma^2, \mathbf{p}\}$  and the sub-samples  $I_{lj}$  of individuals are defined as  $I_{jk} = \{i : z_i = l \text{ and } x_i = j\}$ . By introducing the latent type variables  $s_{00}$  and  $s_{11}$  and the scale parameters  $\lambda$  the posterior distribution can be estimated by a Gibbs algorithm based the conditional distributions of the parameters. The unobserved types  $s_{00}$  and  $s_{11}$  can be sampled directly from their conditional posterior distributions, as in Chib and Jacobi (2008). For example, for a subject  $i \in I_{00}$ , who can either be a complier or always drop-out 14,

$$\Pr(s_i = C | \boldsymbol{\beta}, \mathbf{y}, \mathbf{x}, \mathbf{z}) \propto p_C t_\nu(y_i | \mathbf{w}'_i \boldsymbol{\beta}_{C0}, \sigma_{C0}^2).$$

with normalizing constant

$$c = \frac{1}{p_C t_\nu(y_i | \mathbf{w}'_i \boldsymbol{\beta}_{C0}, \sigma_{C0}^2) + p_{AD14} t_\nu(y_i | \mathbf{w}'_i \boldsymbol{\beta}_{AD14}, \sigma_{AD14}^2)}$$

Similarly, for a subject  $i \in I_{11}$ , who can be of type always drop-out 15 or complier,

$$\Pr(s_i = C | \boldsymbol{\beta}, \mathbf{y}, \mathbf{x}, \mathbf{z}) = \frac{p_C t_\nu(y_i | \mathbf{w}'_i \boldsymbol{\beta}_{C1}, \sigma_{C1}^2)}{p_C t_\nu(y_i | \mathbf{w}'_i \boldsymbol{\beta}_{C1}, \sigma_{C1}^2) + p_{AD15} t_\nu(y_i | \mathbf{w}'_i \boldsymbol{\beta}_{AD15}, \sigma_{AD15}^2)}$$

Based on the above observations, our MCMC algorithm is defined as follows.

1. Sample  $(\mathbf{s}_{00}, \mathbf{s}_{11} | \mathbf{y}, \mathbf{x}, \boldsymbol{\beta}, \boldsymbol{\sigma}^2)$  by sampling  $s_i$  for  $i \in I_{00}$  with  $\Pr(s_i = c | y_i, x_i, \boldsymbol{\beta}_{0c}, \boldsymbol{\beta}_{0n}, \boldsymbol{\sigma}_{0c}^2, \boldsymbol{\sigma}_{0n}^2)$  and  $s_i$  for  $i \in I_{11}$  with  $\Pr(s_i = c | y_i, x_i, \boldsymbol{\beta}_{1c}, \boldsymbol{\beta}_{1a}, \boldsymbol{\sigma}_{1c}^2, \boldsymbol{\sigma}_{1a}^2)$

2. Sample  $\lambda_i$  for  $i = 1, \dots, n$  from the Gamma density

$$\mathcal{G} \left( \lambda_i \mid \frac{\nu + 1}{2}, \frac{\nu + (y_i - \mathbf{w}'_i \boldsymbol{\beta}_{kj})^2}{2} \right)$$

3. Sample  $\mathbf{p} | \mathbf{s}_{00}, \mathbf{s}_{11}$  from the Dirichlet density  $\pi(\mathbf{p} | \mathbf{s}_{00}, \mathbf{s}_{11}) = \text{Dir}(\alpha_C, \alpha_{AD14}, \alpha_{AD15})$  where  $\alpha_k = a_{k,0} + \sum_{i=1}^n I[s_i = k]$  with  $k = C, AD14, AD15$

4. Sample  $\boldsymbol{\beta}_{kj} | \mathbf{y}, \mathbf{x}, \boldsymbol{\sigma}^2, \mathbf{s}_{00}, \mathbf{s}_{11}, \boldsymbol{\lambda}$  from the normal density

$$\mathcal{N} \left( \boldsymbol{\beta}_{kj} \mid \mathbf{B}_{kj} \left\{ \mathbf{B}_{kj,0}^{-1} \boldsymbol{\beta}_{kj,0} + \sum_{i \in I_{kj}} \lambda_i \mathbf{w}_i \sigma_{kj}^{-2} y_i \right\}, \left\{ \mathbf{B}_{kj,0}^{-1} + \sum_{i \in I_{kj}} \lambda_i \mathbf{w}_i \sigma_{kj}^{-2} \mathbf{w}'_i \right\}^{-1} \right)$$

where  $n_{kj}$  denotes the number of individuals in the set  $I_{kj}$

5. Sample  $\sigma_{kj}^2 | \mathbf{y}, \mathbf{x}, \boldsymbol{\beta}, \mathbf{s}_{00}, \mathbf{s}_{11}, \boldsymbol{\lambda}$  from inverse gamma density

$$\text{IG} \left( \sigma_{kj}^2 \mid \frac{n_{kj,0} + \nu_{kj}}{2}, \frac{\delta_{kj,0} + \sum_{i \in I_{kj}} \lambda_i (y_i - \mathbf{w}'_i \boldsymbol{\beta}_{kj})^2}{2} \right)$$

The algorithm updates the regression parameters  $\boldsymbol{\beta}_{kj}$  and the variance parameters  $\sigma_{kj}^2$  conditional on the subject types  $\mathbf{s} = \{s_i : i \leq n\}$  and the sub-samples of individuals  $i \in I_{jk}$  where  $I_{jk} = \{i : x_i = j \text{ and } s_i = k\}$  from normal and inverse gamma distributions respectively. The scale parameters  $\lambda$  can be directly sampled from gamma distributions conditional on the types and the other parameters and the type probabilities are drawn from Dirichlet distributions conditional on the type updates. The above steps are repeated a large number of times. After the first 1000 burn-in draws, the subsequent 10,000 draws are used to derive the posterior quantities of interest.

## 7.2 Prior Specification

To help specify suitable parameters for the prior distribution (equation 3.5) for the empirical analysis we fit the model to the an auxiliary sample of 2,518 males that turned 14 in 1944-1945 or 1949-1950 period using flexible and uninformed priors: the means for the normal

priors of the regression coefficients are set at zero and standard deviations at 5; the prior means in the inverse gamma distributions for the variance parameters are fixed at 0.5 and the standard deviations at 1; the parameters of the Dirichlet prior for the type probabilities are set to 50, 30 and 20. The posterior distributions of the model parameters for the various model specifications for earnings and wages based on the auxiliary sample are obtained using the algorithm described in the previous section. For the analysis of the main sample the prior means of the regression coefficients and the variances are set to the corresponding estimated posterior means. The prior standard deviations are set at five times the values of estimated posterior standard deviations. The values of the Dirichlet priors are according the estimated posterior means obtained from the analysis of the auxiliary data set. Following this approach the values for the prior means and standard deviations from the two models with 5 degrees of freedom are set at the values shown in Table 5:

	Prior Distribution Parameter Values							
	C0	Specification 1			C0	Specification 2		
		C1	AD14	AD15		C1	AD14	AD15
Log Weekly Earnings								
$\beta_{kj,0}/B_{kj,0}^{0.5}$	3.479/1.674 0.024/0.020	4.591/1.373 0.012/0.016	4.720/3.872 0.009/0.046	3.191/4.689 0.029/0.057	4.198/2.275 0.015/0.028 0.078/0.192	4.923/2.062 0.007/0.025 0.115/0.174	5.741/5.230 -0.003/0.064 0.171/0.556	5.854/8.411 -0.004/0.104 0.220/0.511
$\nu_{kj,0}/\delta_{kj,0}$	0.092/0.033	0.094/0.033	0.167/0.131	0.125/0.091	0.091/0.032	0.091/0.035	0.168/0.128	0.130/0.100
$\alpha_{k,0}$	0.747		0.107	0.146	0.746		0.107	0.147
Log Hourly Wages								
$\beta_{kj,0}/B_{kj,0}^{0.5}$	-0.371/1.732 0.026/0.021	0.887/1.325 0.012/0.016	1.206/4.186 0.007/0.049	-0.687/4.189 0.031/0.051	0.421/2.341 0.016/0.029 0.086/0.200	1.293/1.835 0.006/0.023 0.102/0.179	2.373/5.525 -0.008/0.067 0.195/0.571	1.170/6.699 0.008/0.083 0.195/0.474
$\nu_{kj,0}/\delta_{kj,0}$	0.099/0.036	0.099/0.034	0.189/0.144	0.135/0.096	0.099/0.035	0.098/0.488	0.185/0.926	0.138/0.689
$\alpha_{k,0}$	0.747		0.107	0.146	0.747		0.107	0.146

**Table 5:** Parameter values for prior distribution used in the estimation of the model with student-t errors with 5 degrees of freedom.

### 7.3 Marginal Likelihood Estimation

We compute the marginal likelihood for the model comparison via Bayes factors marginalized over the latent type variables  $s_{00}$  and  $s_{11}$  by evaluating the following expression

$$\ln m(\mathbf{y}, \mathbf{x}) = \ln f(\mathbf{y}, \mathbf{x}, \mathbf{W}, \mathbf{z} | \boldsymbol{\theta}^*) + \ln \pi(\boldsymbol{\theta}^*) - \ln \pi(\boldsymbol{\theta}^* | \mathbf{y}, \mathbf{x}, \mathbf{W}, \mathbf{z})$$

where  $\boldsymbol{\theta}^*$  refers to the vector of the posterior means of the parameters. The first expression of the likelihood can be evaluated directly at the posterior means of the parameters using the expression ... from the previous section. The second expressions  $\pi(\boldsymbol{\theta}^*)$ , the prior distribution evaluated at the posterior means, can also be computed directly from

$$\pi(\boldsymbol{\theta}^*) = Dir(\mathbf{p}^* | \alpha_{c,0}, \alpha_{ad0,0}, \alpha_{ad1,0}) \prod_{j=0}^1 \prod_{k \in K_j} \mathcal{N}_p(\boldsymbol{\beta}_{kj}^* | \boldsymbol{\beta}_{kj,0}, B_{kj,0}) \mathcal{IG}(\boldsymbol{\sigma}_{kj}^{2*} | \nu_{kj,0}, \delta_{kj,0})$$

The last expression, the posterior ordinate, can be evaluated using the following decomposition

$$\pi(\boldsymbol{\theta}^* | \mathbf{y}, \mathbf{x}, \mathbf{W}, \mathbf{z}) = \pi(\boldsymbol{\sigma}^{2*} | \mathbf{y}, \mathbf{x}, \mathbf{W}, \mathbf{z}) \pi(\boldsymbol{\beta}^* | \boldsymbol{\sigma}^{2*}, \mathbf{y}, \mathbf{x}, \mathbf{W}, \mathbf{z}) \pi(\mathbf{p}^* | \boldsymbol{\sigma}^{2*}, \boldsymbol{\beta}^*, \mathbf{y}, \mathbf{x}, \mathbf{W}, \mathbf{z})$$

The first component of the posterior ordinate of can be evaluated directly via the Rao-Blackwell method as

$$\pi(\boldsymbol{\sigma}^{2*} | \mathbf{y}, \mathbf{x}, \mathbf{W}, \mathbf{z}) = \frac{1}{M} \sum_{g=1}^M \mathcal{IG} \left( \sigma_{kj}^{2*} \left| \frac{n_{kj,0} + \nu_{kj}}{2}, \frac{\delta_{kj,0} + \sum_{i \in I_{kj}^{(g)}} \lambda_i^{(g)} (y_i - \mathbf{w}_i' \boldsymbol{\beta}_{kj}^{(g)})^2}{2} \right. \right)$$

using the draws on  $\lambda_i$ 's and  $\boldsymbol{\beta}_{kj}$  from the main run of the MCMC algorithm. The first reduced run of the MCMC fitting algorithm with  $\sigma_{kj}^2$  fixed at  $\sigma_{kj}^{2*}$  to estimate  $\pi(\boldsymbol{\beta}^* | \boldsymbol{\sigma}^{2*}, \mathbf{y}, \mathbf{x}, \mathbf{W}, \mathbf{z})$  as

$$\frac{1}{M} \sum_{g=1}^M \mathcal{N} \left( \boldsymbol{\beta}_{kj}^* | \mathbf{B}_{kj} \left\{ \mathbf{B}_{kj,0}^{-1} \boldsymbol{\beta}_{kj,0} + \sum_{i \in I_{kj}^{(g)}} \mathbf{w}_i \sigma_{kj}^{-2*} y_i \right\}, \left\{ \mathbf{B}_{kj,0}^{-1} + \sum_{i \in I_{kj}^{(g)}} \mathbf{w}_i \sigma_{kj}^{-2*} \mathbf{w}_i' \right\}^{-1} \right)$$

Finally,  $\pi(\mathbf{p}^* | \boldsymbol{\sigma}^{2*}, \boldsymbol{\beta}^*, \mathbf{y}, \mathbf{x}, \mathbf{W}, \mathbf{z})$  is estimated from

$$\frac{1}{M} \sum_{g=1}^M Dir(\mathbf{p}^* | \alpha_{c,0} + \sum_{i=1}^n I[s_i = c], \alpha_{ad0,0} + \sum_{i=1}^n I[s_i = ad0], \alpha_{ad1,0} + \sum_{i=1}^n I[s_i = ad1])$$

using a second reduced run of the MCMC algorithm with  $\sigma_{kj}^2$  fixed at  $\sigma_{kj}^{2*}$  and  $\boldsymbol{\beta}_{kj}$  fixed at  $\boldsymbol{\beta}_{kj}^*$ .

## 7.4 Appendix: Additional Results

Model Comparison								
<i>Log Marginal Likelihoods</i>								
dof /spec.	5/1	5/2	10/1	10/2	20/1	20/1	$\infty$ /1	$\infty$ /2
Earn	-2029.13	<b>-2023.80*</b>	-2049.93	-2054.71	-2071.13	-2067.62	-2095.32	-2096.73
Wage	-2090.95	<b>-2089.27*</b>	-2091.59	-2098.80	-2107.75	-2107.31	-2121.77	-2125.24

**Table 6:** Estimated log marginal likelihoods for various degrees of freedom and covariate specifications (\* Model preferred based on marginal likelihood criterion) using uninformed prior. Results are based on the draws from the MCMC sampler with 10,000 iterations and 1000 burn-in iterations.

Estimates of Average Returns to Schooling					
D.o.F		5	CATE: Mean and Stand.Dev.		
			10	20	$\infty$
Earnings	S1	-0.013 (0.045)	0.009 (0.035)	0.011 (0.028)	0.021 (0.010)
	S2	-0.004 (0.043)	0.011 (0.033)	0.011 (0.027)	0.018 (0.022)
Wages	S1	-0.037 (0.049)	-0.010 (0.037)	-0.004 (0.034)	0.000 (0.027)
	S2	-0.027 (0.047)	-0.009 (0.036)	-0.004 (0.034)	0.000 (0.027)

**Table 7:** Posterior means and standard deviations for complier average treatment effect based on the fittings of the various model specifications with uninformed priors.

## References

- Daron Acemoglu and Joshua Angrist. How large are human-capital externalities? evidence from compulsory-schooling laws. In *NBER Macroeconomics Annual 2000*, volume 15 of *NBER Chapters*, pages 9–74. MIT Press, January-D 2001. URL <http://ideas.repec.org/h/nbr/nberch/11054.html>.
- Joshua D. Angrist and Alan B. Krueger. Does compulsory school attendance affect schooling and earnings? *Quarterly Journal of Economics*, 106(4):979–1014, November 1991.
- James Banks and Richard Blundell. Private pension arrangements and retirement in Britain. *Fiscal Studies*, March 2005.
- Siddhartha Chib. Analysis of treatment response data without the joint distribution of potential outcomes. *Journal Of Econometrics*, October 2007.
- Siddhartha Chib. Marginal likelihood from the gibbs output. *Journal Of The American Statistical Association*, 90:1313–1321, December 1995.
- Siddhartha Chib and Liana Jacobi. Analysis of treatment response data from eligibility designs. *Journal Of Econometrics*, 144:465–478, June 2008.
- Damon Clark and Heather Royer. The effect of education on adult mortality and health: evidence from the United Kingdom. *Centre for Economic Performance*, mimeo, July 2007.
- Paul J. Devereux and Robert A. Hart. Forced to be rich? returns to compulsory schooling in Britain. *Economic Journal*, 120:1345–1364, December 2010.
- Fernando Galindo-Rueda and Anna Vignoles. The heterogeneous impact of selection in secondary schools. In Ludger Woessman and Paul Peterson, editors, *Schools and the Equal Opportunity Problem*. MIT Press, Cambridge, 2007.
- Julien Grenet. Is it enough to increase compulsory education to raise earnings? evidence from French and British compulsory schooling laws. *CEP Working Paper*, 2009.

- Heath Anthony .F. Halsey, Albert H. and John M. Ridge. *Origins and Destinations: Family, Class and Destinations in Modern Britain*. Oxford University Press, New York, 1980.
- Colm Harmon and Ian Walker. Estimates of the economic return to schooling for the united kingdom. *American Economic Review*, December 1995.
- Colm Harmon and Ian Walker. The marginal and average returns to schooling in the uk. April 1999.
- Guido W. Imbens and Joshua D. Angrist. Identification and estimation of local average treatment effects. *Econometrica*, 62(2):467–475, March 1994.
- Thomas J. Kane and Cecilia E. Rouse. Labor-market returns of two-and four-year college. *American Economic Review*, 85(3):600–614, June 1995.
- David S. Lee and Thomas Lemieux. Regression discontinuity in economics. *Journal of Economic Literature*, 48:281–355, June 2010.
- Maarten Lindeboom, Ana Llana-Nozal, and Bas van der Klaauw. Parental education and child health: Evidence from a schooling reform. *Journal Of Health Economics*, 28:109–131, January 2009.
- Costas Meghir and Marten Paelme. Educational reform, ability and parental background. *American Economic Review*, March 2005.
- Philip Oreopoulos. Estimating average and local average treatment effects of education when compulsory schooling laws really matter. *American Economic Review*, March 2006.
- Philip Oreopoulos. Estimating average and local average treatment effects of education when compulsory schooling laws really matter: Corrigendum. 2008.
- Jorn-Steffen Pischke and Till von Wachter. Zero returns to compulsory schooling in germany: Evidence and interpretation. *Review of Economics and Statistics*, June 2008.

Alfred Sommer and Scott Zeger. On estimating efficacy in clinical trials. *Statistics in Medicine*, January 1991.

William M. K. Trochim. In *Research Design for Program Evaluation: The Regression-Discontinuity Approach*. Sage Publication, Beverly Hills, 1984.