Title
School quality and the return to schooling in Britain: New evidence from a large-scale compulsory schooling reform

Author
Clark, Damon

Publication Date
2023-07-01

DOI
10.1016/j.jpubeco.2023.104902

Peer reviewed

School Quality and the Return to Schooling in Britain: New Evidence from a Large-Scale Compulsory Schooling Reform

Damon Clark *

March 19th, 2023

Abstract

What is the causal effect of schooling on subsequent labor market outcomes? In this paper I contribute evidence on this question by re-examining a British compulsory schooling reform that yields large-scale and quasi-experimental variation in schooling. First, I note that this reform was introduced in 1947, when British students attended higher-track (for the "top" 20%) or lower-track (for the rest) secondary schools. The reform increased the minimum school leaving age from 14 to 15 and I show that the vast majority (over 95%) of affected students attended lower-track schools. Second, I show that the additional schooling induced by the reform had close to zero impact on a range of labor market outcomes. Third, I attribute these low returns to the quality of these lower-track schools, which I argue was low along several dimensions.

JEL: I21, J24, J31.

*University of California, Irvine and NBER. Email: clarkd1@uci.edu. Much of this work was done at the UK Department for Work and Pensions (DWP). I thank them for providing work space and facilitating access to the relevant data. The views expressed here are mine, and not those of the DWP. I thank Derek Gillard for creating and maintaining the excellent `www.educationengland.org.uk` website, without which this paper would have been difficult to write. Finally, I am grateful for the many helpful and insightful comments that I have received from seminar participants and in private conversations.

1 Introduction

What is the causal effect of schooling on subsequent labor market outcomes? This is a vital question for economists. At the macroeconomic level the answer speaks to the impact of education on productivity and hence why some countries are richer than others (Krueger and Lindahl, 2001). At the microeconomic level the answer can inform models of education choices (e.g., the importance of various frictions) and policies designed to increase educational attainment. As discussed by Card (2001), economists have addressed this question by leveraging "quasi-experimental" variation in schooling driven by the supply side of the education market (e.g., education policy reforms); the resulting estimates are interpreted as the average returns to schooling among the affected students (Angrist et al., 1996).

Despite the appeal of this approach, convincing quasi-experiments have proven hard to find. At the college level, while some quasi-experiments suggest large returns (e.g., Zimmerman (2014)), Barrow and Malamud (2015) review the literature and conclude that "the scarcity of credible evidence regarding the causal effect of college on earnings is striking" (p.539). At the high school level, changes to compulsory schooling laws yield more quasi-experiments, but these often lack external validity. In some cases (e.g., the US) this is because the expansions affected only a small fraction of students.[1] In other cases it is because the expansions were part of broader education reform packages that are difficult to unpack.[2] It follows that much remains to be learned about the causal effect of schooling on labor market outcomes.

In this paper I contribute evidence on this question by re-examining a British compulsory schooling change that yields arguably the best approximation yet found to large-scale experimental variation in schooling. This change, introduced on 1 April 1947, increased the minimum school leaving age from 14 to 15 (roughly grade nine to grade ten). As a measure of its scale, more than one half of the cohorts exposed to the change completed an additional year of schooling. It approximates an experiment because the minimum leaving age depended on a student's date of birth relative to a sharp cutoff. Since it generated schooling variation among individuals born just days apart, regression discontinuity methods can provide highly credible estimates of its impacts.

I make three contributions that advance our understanding of this compulsory schooling change and its impact on later-life outcomes. First, I note that over the period covered by

[1] Angrist and Pischke (2014) show that in models that include state-specific trends, US compulsory schooling changes have small impacts on educational attainment. Angrist and Krueger (1991) exploit the interaction of US compulsory schooling laws and quarter of birth, but Buckles and Hungerman (2013) argue that some of the seasonality found in education and earnings outcomes is related to maternal characteristics.

[2] For example, reforms in Norway and Sweden changed compulsory schooling laws and de-tracked those school systems (see Aakvik et al. (2010) and Meghir and Palme (2005) respectively).

this change, Britain operated an elite school system in which the "top" 20% of students were assigned to higher-track schools and the remaining 80% to lower-track schools. Using new data, I estimate that over 95% of the students affected by this reform (i.e., "compliers") attended lower-track schools. The reason is simple. Since most lower-track schools *terminated* at the minimum leaving age, the reform added another year to the education of most lower-track students. In contrast, most higher-track schools continued until at least age 18 and few higher-track students left at the minimum leaving age. It follows that the compulsory schooling change can be viewed as mandating an extra year - and no more - of lower-track schooling.

Second, I estimate the impacts of this reform by leveraging new datasets with large samples and precise information on birth date. Using regression discontinuity methods, I find that the schooling induced by this reform had minimal impacts on subsequent labor market outcomes, at least on average. Indeed, I cannot reject zero impacts for any of the outcomes that I consider, including labor market participation, unemployment, earnings and home ownership (a proxy for wealth). These impacts are much smaller than those typically associated with a year of schooling - between 5% and 15% according to Card (2008).[3] Although my findings overturn prior research that found positive effects of this reform on earnings (e.g., Oreopoulos (2006); Devereux and Hart (2011)), I replicate key estimates reported in this prior research and can reconcile them with mine.[4]

Third, I argue that my findings are likely attributable to the low quality of the (lower-track) schooling induced by this reform. In particular, I focus on three features of this schooling: resources, curriculum and student effort. First, I show that lower-track students were taught in relatively larger classes, and by teachers with relatively little training. Second, I show that lower-track schools provided a curriculum that emphasized practical education at the expense of vocational or traditional academic education. Third, I argue that lower-track students likely faced weak incentives to exert effort. Although evidence on the effects of inputs such as resources and curriculum is still accumulating, I argue that all three mechanisms could help to explain my findings.

[3] Pischke and von Wachter (2008) estimate zero returns to a German compulsory schooling reform but a recent analysis finds returns of 6-8% for hourly wages (Cygan-Rehm, 2021). Oosterbeek and Webbink (2007) estimate zero returns to lower-track vocational schooling in Holland.

[4] My findings are consistent with - and provide a possible explanation for - this reform having little impact on subsequent health outcomes (Clark and Royer, 2013). Galama et al. (2018) discuss this result. Since tracking was abolished in the 1960s, it is less relevant to the other (1972) reform analyzed by Clark and Royer (2013).

2 Britain's Education System

I now describe Britain's education system and the compulsory schooling reform.

A: Education System

At the start of the twentieth century the British education system included distinct elementary and secondary sectors. Elementary schools were run by local school boards and voluntary agencies (e.g., churches) and they received grant funding from the central government for students of compulsory school age (between 5 and 13). Secondary schools offered a more advanced education that continued until at least age 16. These schools were private, and funded mostly by fees and endowments.[5]

The 1902 Education Act created a new system of Local Education Authorities (LEAs) that assumed responsibility for both sectors. Over the next 40 years, the sectors changed significantly but remained separate. On the elementary side, the school leaving age was increased to 14 (in 1921) and, following the report of Hadow (1926), elementary education was divided into two stages taught in separate schools: a primary stage up to age 11 taught in junior elementary schools and a post-primary stage from age 11 onwards taught in senior elementary schools. On the secondary side, starting in 1907, LEAs had to assign 25% of secondary seats to students entering from elementary schools (as opposed to private preparatory schools). These students were not charged fees but had to pass an exam to demonstrate they could benefit from an advanced academic education. By 1939 all secondary school students were admitted on the basis of an exam and then charged fees according to their means.

The 1944 Education Act integrated the elementary and secondary sectors into a single system that contained three types of secondary schools: Grammar schools (former secondary schools), Secondary Modern schools (former senior elementary schools) and Technical High schools (a new type of school). I refer to the first two types as higher-track and lower-track schools; I ignore the third type since few such schools were ever built. I examine the content of lower-track education in Section 5.

B: Education System and Compulsory Schooling Reform

The 1944 Act stated that the minimum school leaving age would eventually be raised from 14 (when most lower-track schools terminated) to 16 (the age to which higher-track students were expected to continue). As a first step, it was raised to 15 on 1 April 1947.

Figure 1 shows how the compulsory schooling reform affected the different tracks. Panel A uses administrative enrollment data (described in Appendix A.1). Figure 1i displays the

[5]Although some terms are specific to England and Wales, Scotland also had a tracked education system.

proportion of 13 year old students in lower-track schools, higher-track schools and other (mostly private) schools by year. The series are broadly stable, with about 75% attending lower-track schools, 15% attending higher-track schools and 10% attending other schools. Figures 1ii and 1iii display dropout rates by year and track. Before the reform (e.g., in 1946), the vast majority (around 90%) of lower-track students dropped out at age 14. After the reform (e.g., in 1948), only around 5% of lower-track students dropped out at age 14.[6] Throughout the period, less than 5% of higher-track students dropped out at age 14, with around 10% dropping out by age 15 and 50% dropping out by age 16.

Panel B of Figure 1 displays parallel graphs using data from the 1973 wave of the General Household Survey (GHS).[7] These data fill three gaps in the administrative enrollment data used in panel A. First, they include cohorts that turned 14 during the War. Second, they include respondents that attended other schools, hence Figure 1vi plots the fraction of higher-track *or* other students that left school at age 14.[8] Third, they include information on birth month, hence the horizontal axis refers to the *quarter* that a respondent turned 14.[9] This brings the compulsory schooling reform into sharper focus. Both panels suggest that the reform had a huge impact on lower-track students but little impact on other students.

To quantify that last point, I calculate the fraction of students affected by the compulsory schooling reform (i.e., the "compliers") that attended lower-track schools. The inputs into this calculation are the fraction of the relevant cohort that attended lower-track schools and the fraction of students in each type of school that were affected by the compulsory schooling reform (estimated via regression discontinuity models applied to the data shown in Figures 1v and 1vi). This calculation suggests that at least 95% of the students affected by the compulsory schooling reform attended lower-track schools (Appendix B).

This finding is not difficult to explain. First, since the vast majority of lower-track schools *terminated* at the minimum leaving age, it is not surprising that the vast majority of lower-track students dropped out at age 14 before the reform and at age 15 after the reform.[10] Second, since higher-track students were selected on the basis of their ability to benefit from an advanced academic education, it is not surprising that so few dropped out at age 14 before the reform. In addition, their families often had to commit to keeping them in school until at least age 16 (e.g., Lowndes (1947, p.116)).

[6] This could reflect non-compliance or measurement error.

[7] This is the only wave that collected information on type of school attended and birth month. These data are described in Appendix A.2.

[8] To avoid cluttering the graphs I do not display the fractions leaving by ages 15 or 16.

[9] The sample is relatively small, hence the patterns are clearer when data are presented at the quarter rather than the month level.

[10] This fact was established by Hadow (1926) for the late 1920s. There is little reason to think that provision changed over the next 20 years (e.g., the series in Figure 1 are remarkably stable).

C: Descriptive Statistics

Since the reform affected lower-track students, it helps to examine some descriptive statistics for this population. To that end, panel A of Table 1 uses the GHS data to describe the family background, educational attainment and labor market outcomes of lower-track students that turned 14 just prior to the reform (i.e., a type of control group). For context, I report equivalent statistics for higher-track and other students. As expected, lower-track students came from lower-SES families as measured by father's occupation (e.g., over 80% were employed in manual occupations). Also as expected, lower-track students rarely earned the School Certificate (which required continuing in school until age 16) and almost never attended college or university. At age 40, large fractions (around 70%) of these lower-track students were employed in manual occupations, and they earned around one third less than higher-track and private school students.

3 Data and Methods

We have seen that the compulsory schooling reform kept lower-track students in school for an additional year. I now ask whether this reformed improved subsequent labor market outcomes. Earnings are the main outcome of interest but I also examine labor market activity (i.e., labor force participation, unemployment and self-employment). Education could have important causal effects on labor market activity (Card, 1999). Moreover, since I study earnings among samples of employed workers, it is important to know whether the compulsory schooling reform affected the probability of being observed in these samples.

A: Data

I use 1991 Census data to measure labor force participation, unemployment and self-employment.[11] For the subsample of Census respondents born in England and Wales, these data consist of population counts of outcomes by month of birth and gender.

I use L2 data to measure earnings. The L2 is an administrative dataset that, for a 1% sample of the population, contains earnings information for each year from 1978 onwards.[12] The sample is drawn by extracting all workers with National Insurance numbers ending in a particular pair of digits (these numbers are randomly assigned). Since the same digits are used in each year, the L2 can contain multiple observations for a single person. I obtained records for individuals born 1887-1963 in years 1978-2009 (N=11,708,045). I dropped indi-

[11] I obtained these and other Census data from the Office for National Statistics (Crown Copyright). The Census is conducted every 10 years. There are no questions on earnings.

[12] See Appendix A.3 for additional details.

viduals with missing date of birth information (1,696), migrants (339,491), respondents living in Northern Ireland (most of whom would have been born in Northern Ireland - 260,898) and respondents aged 60 or older (3,881,888).

I use the L2 data to construct four outcomes. First, I construct "annual weeks worked" as the sum of total weeks employed in a given year.[13] Second, I construct "annual earnings" as the sum of earnings across these employments, deflated using the annual retail price index ("All Items" series with base period 2013).[14] These data do not contain self-employed earnings, hence I exclude observations with any weeks of self-employment. As shown below, Census data suggest that the compulsory schooling reform had no impact on self employment. Third, I construct "weekly earnings" as annual earnings divided by annual weeks worked.[15] Fourth, I construct a measure of "positive weekly earnings". It takes the value "0" if the observation is not associated with a valid measure of weekly earnings (e.g., if the person did not work or was self-employed) and "1" otherwise. If this value is "1", the observation is associated with non-zero values for weeks worked, annual earnings and weekly earnings.

B: Methods

I use regression discontinuity methods to estimate the impacts of exposure to the compulsory schooling reform. Later, when comparing my findings to the related literature, I convert these "reduced-form" estimates into treatment effect estimates of the returns to schooling. I now describe these regression discontinuity methods in more detail.

The Census data contain one observation per cohort defined by month turned 14 (e.g., January 1947). For outcomes derived from these data I estimate versions of the following model:

$$Y_c = \pi_0 + \pi_1 T_c + \pi_2 R_c + \pi_3 R_c T_c + \varepsilon_c \tag{1}$$

where c indexes cohort, Y_c is a cohort outcome (e.g., fraction unemployed), T_c is an exposure indicator that is one for cohorts exposed to the reform (i.e., turned 14 after 1 April 1947) and zero otherwise. The running variable R_c is a cohort index defined relative to 1 April 1947 and $R_c T_c$ is the interaction of the running variable and the exposure indicator.[16] I estimate this equation by ordinary least squares: $\widehat{\pi_1}$ is my estimate of the impact of reform exposure. With possible seasonality in mind, I check robustness to adding month-of-birth fixed effects (i.e., for being born in January, February, etc).

[13] I reset these values to missing in the small number of cases (708) in which they are negative.

[14] I multiply these values by "-1" in the small number of cases (155) in which they are negative.

[15] I set weekly earnings to missing if they exceed 10,000.

[16] Instead of coding the running variable to be 0 for March 1933, 1 for April 1933 etc, I code it to be -0.5 for March 1933, 0.5 for April 1933 etc. Following Dong (2015), this lessens any bias associated with using local linear regression with a running variable that is a rounded version of a continuous variable (i.e., birth month versus date of birth). In practice, the estimates are very similar across the two coding schemes.

The L2 data contain individual-level observations. For outcomes derived from these data I use a two-step approach. First, I construct mean outcomes by cohort by averaging outcomes across individuals and years within cohorts. Second, I estimate versions of equation (1) with these cohort-level outcomes as the dependent variable (i.e., using the methods described above). A cohort-level approach applied to both sets of outcomes simplifies the exposition at little cost. That is because estimates obtained from these second-step cohort-level models are very similar to those obtained when equation (1) is estimated using individual-level data.[17] In addition to checking robustness to adding month-of-birth fixed effects I check robustness to an alternative first-step procedure in which I extract the estimated cohort fixed effects from an individual-level regression that relates the relevant outcome (e.g., earnings) to these cohort fixed effects, year dummies and a polynomial in age.

Equation (1) can be viewed as a linear approximation to the unknown relationship between outcome and cohort; the OLS estimator of $\widehat{\pi}_1$ can be viewed as a local linear regression (LLR) estimator with a uniform kernel function. Rather than focus on estimates at particular bandwidths believed to be optimal, I construct graphs that display estimates for bandwidths ranging from one to seven years. This illustrates the extent to which estimates are sensitive to bandwidth choice and preserves the simplicity and transparency of the LLR approach (Cattaneo et al., 2019).[18] I report robust standard errors but also construct "honest confidence intervals" that account for the potential bias implied by the nonparametric regression approach (Kolesár and Rothe, 2018).[19] These bias-aware confidence intervals are wider than standard confidence intervals but ensure correct coverage provided the true relationship between Y_c and R_c belongs to a class of smooth functions with curvature less than some value K. I choose K using the rule-of-thumb suggested by Noack and Rothe (2019).[20]

C: Validity

To validate the regression discontinuity approach, I check there is no discontinuity in the size of the underlying population as measured by the number of 1991 Census respondents (Figure F1).

[17] The estimates are numerically identical when the cohort-level models are weighted by cohort size. Although I report unweighted estimates, cohort size exhibits little variation (Figure F1) and weighted estimates (available on request) are very similar.

[18] In addition, standard approaches for finding MSE-optimal bandwidths assume that the running variable is continuous (Cattaneo et al., 2019).

[19] I base these honest confidence intervals on robust standard errors. While this is more conservative than necessary (see footnote 16 of Kolesár and Rothe (2018)), the alternative nearest-neighbor standard error estimators are not valid given one observation per value of the running variable (i.e., the case here).

[20] This involves estimating a global second-order polynomial and choosing K to be twice as large as the maximal curvature found in the data.

4 Results

I now examine how exposure to the compulsory school reform impacted labor market outcomes. I then relate my findings to the wider literature.

A: Labor Market Activity

Figure 2 plots labor force participation, unemployment and self-employment rates by cohort. These graphs - which are based on population data - suggest that exposure to the compulsory schooling reform had little impact on these outcomes. This impression is confirmed by regression discontinuity estimates of exposure impacts. Across a range of bandwidths, these estimates are small and statistically indistinguishable from zero. For example, using a bandwidth of one year, and for both men and women, I can rule out labor force participation increases and unemployment decreases of one percentage point.

For another estimate of the impact of reform exposure on labor market activity, I examine the probability of being observed with positive earnings in the L2 data. To this end I pool observations across individuals and years within a given cohort (e.g., turned 14 in April 1947) and calculate the fraction of these observations with positive weekly earnings. I then estimate effects on this fraction. Consistent with the Census-based analysis I obtain small point estimates and can never reject zero impacts (Figure F2).

B: Earnings

Since I find no evidence to suggest that the compulsory school reform impacted labor market activity, I restrict the L2 sample to observations with positive weekly earnings. I then proceed in two steps. First, I calculate cohort means for three earnings measures: log weekly earnings, log annual earnings and log annual earnings among full-time workers (i.e., the subsample of annual earnings records for which "weeks worked" exceed 40). Second, I estimate impacts on these outcomes of exposure to the compulsory schooling reform.

In the top row of Figure 3 I plot log weekly earnings by cohort and gender.[21] As expected, later-born cohorts earn more than earlier-born cohorts. This is because these cohorts are observed in the labor market between ages 38 and 60 (i.e., on the flat part of the earnings-experience profile), when age effects (a given cohort earns more as they age) are dominated by cohort effects (later-born cohorts earn more at a given age). We do not see any obvious discontinuity in earnings at the April 1947 cutoff.

In the bottom row of Figure 3 I plot estimated exposure impacts at every bandwidth from one through seven years. I also plot conventional 95% confidence intervals (based on robust

[21] The outcomes for the pooled sample ("Both") are estimated cohort fixed effects from first-step regression models that include a gender dummy and these fixed effects. These models use cohorts born 1925-1940.

standard errors) and "honest" (bias-aware) 95% confidence intervals (Kolesár and Rothe, 2018). For men, across all bandwidths, the estimates are close to zero and conventional confidence intervals reject effects on the order of 2%. Honest confidence intervals yield slightly larger upper bounds at small bandwidths and much larger upper bounds at larger bandwidths. For women the estimates are slightly larger but less precise, presumably because hours and weeks worked are more variable (because of part-time work). The pooled estimates are again close to zero, and effects on the order of 2-3% can be ruled out.

Table 2 examines the robustness of these estimates to specification choices. In column (1) I report a selection of the estimates displayed in Figure 3 (at bandwidths of two, four and six years). To avoid cluttering the Table I report robust standard errors but not honest confidence intervals.[22] In column (2) I report the estimates obtained after the outcome has been regression adjusted for age. In column (3) I report the estimates obtained from a specification that includes month-of-birth fixed effects. In column (4) I report the estimates obtained after removing the first cohort exposed to the reform.[23] As discussed in the next section, this addresses the possibility that schools adjusted to the reform slowly. The estimates are broadly robust to all of these choices. For example, focusing on the pooled estimates, and looking across all three bandwidths, estimates range from close to zero to around 1%. The same patterns emerge when I study annual earnings rather than weekly earnings (Figure F4 and Table F1) and when I restrict the sample to full-time workers (Figure F5 and Table F2).

C: Wealth

Finally, since the preceding analyses do not cover outcomes realized before age 38, I estimate impacts on wealth as measured by home ownership. If exposure to the compulsory schooling reform impacted early-career outcomes and earnings, we might expect to find impacts on various proxies for wealth. Instead, I estimate precise zero impacts on this outcome (Figure F6).[24] These estimates are therefore consistent with those for labor market activity and earnings. Taken together, they suggest that exposure to the compulsory schooling reform had minimal impacts on subsequent labor market outcomes.

D: Relation to literature

This paper builds on a large literature that estimates the economic return to a year of schooling. As noted by Card (2008), these estimates range between 5% and 15%. In contrast,

[22] Figure 1 shows the relationship between robust standard errors and honest confidence intervals at different bandwidths; those relationships are robust to other specifications used in Table 2.

[23] At a two-year bandwidth these estimates are imprecise and hence not reported.

[24] Clark and Cummins (2020) estimate zero impacts on housing values as measured by average prices in the areas in which individuals lived in 1999.

after converting the reduced-form estimates reported above into estimates of the return to schooling, I can never reject a zero return to schooling and can reject returns above 5%.[25]

My estimates are also smaller than those reported in prior studies of this compulsory schooling reform (e.g., Oreopoulos (2006, 2008); Devereux and Hart (2011)).[26] For example, Devereux and Hart (2011) estimate returns for men on the order of 4-7%. These estimates are based on a "global polynomial" approach to regression discontinuity estimation in which a fourth-order polynomial in birth cohort is fit to a large range of data (birth cohorts 1921-1951). Gelman and Imbens (2019) criticize this type of approach and argue that it yields estimates that are sensitive to polynomial specification and confidence intervals that do not achieve nominal coverage.

In Appendix C I consider this prior work in more detail. First, I show that key estimates are indeed sensitive to alternative sample and specification decisions. For example, using GHS data and focusing on hourly earnings, Devereux and Hart (2011) estimate a return to men close to 5%. I replicate this estimate and then split their sample into two groups: those that left school at age 15 or below and those that left school at age 16 or above. Since the reform did not impact the probability of staying in school beyond 15, we would expect any impacts of the reform to be concentrated among the first group, and we would not expect to see any impacts among the second group. Instead, I estimate a close to zero impact among the first group and a much larger impact among the second group. I also follow Dolton and Sandi (2017) and show that these estimates are sensitive to polynomial choices.[27]

Second, as another illustration of the Gelman and Imbens (2019) critique, I show that the data used here can also yield positive and statistically significant returns when a global polynomial approach is applied.[28] Third, I show that local linear methods applied to data used in prior studies generate imprecise estimates – a reflection of the small samples involved.

[25] To estimate this return (denoted β_1), I divide the reduced-form estimate (π_1) by the first-stage estimate (γ_1). By the delta method, $Var(\beta_1) = Var(\pi_1)\frac{1}{\gamma_1^2} - 2Cov(\beta_1, \pi_1)\frac{\pi_1}{\gamma_1^3} + Var(\gamma_1)\frac{\pi_1^2}{\gamma_1^4}$. This simplifies to $Var(\pi_1)\frac{1}{\gamma_1^2}$ if the reduced-form and first-stage estimates are obtained from independent samples and if the reduced-form estimates are small relative to the first-stage estimates (as in this case). Table B1 reports first-stage effects around 0.5, but these will be biased down by measurement error in self-reported "age left schooling" (Bingley and Martinello, 2017; Kane et al., 1999). Since Figure 1 suggests compliance of around 85% among the 70% of students that attended lower-track schools, a first-stage effect around 0.6 is more plausible. Combined with a reduced-form estimate (standard error) of 0.006 (0.014) - see Table 2, panel C, column 1 - this implies an estimated return (standard error) of 0.01 (0.023) and rules out returns larger than 5.5%.

[26] These papers build on Harmon and Walker (1995).

[27] For example, as noted above, the 5% return estimated for men is based on a fourth-order polynomial in birth cohort. If a second- or third-order polynomial is used instead, the estimate is only half as large.

[28] For example, if I aggregate the L2 data to the year of birth level, use data covering 30 years and estimate RD models that include fourth-order polynomials, I obtain estimates (standard errors) of 0.037 (0.015) and 0.012 (0.006) when I do and do not interact the polyomial with the reform dummy.

In summary, a key feature of the analyses presented here is that they leverage new datasets that contain large samples and information on birth month. This allows me to sidestep problems with the global polynomial approach and estimate local linear models using data much closer to the cutoff.

5 Interpretation

In principle, the additional schooling have improved students' labor market productivity and yet failed to improve the main outcomes studied here. As discussed in Appendix D.2 however, there is little support for mechanisms that could underpin this story (e.g., rigid wages or general equilibrium effects). Instead, the more plausible explanation is that the additional schooling did little to improve students' labor market productivity. This is consistent with the results presented above. It is also consistent with evidence on skill use. In particular, using occupation recorded on the 1991 Census and measures of skill use employed by Autor et al. (2003), I find little evidence that exposure to the reform altered skill use in ways that might be associated with higher productivity (see Appendix D.1).[29] Turning to why the additional schooling failed to improve productivity, I conclude this is likely attributable to the low quality of the (lower-track) schooling induced by this reform. I draw this conclusion after considering three dimensions of this schooling: resources, curriculum and incentives.

A: Resources

As seen in panel B of Table 1, lower-track students were taught in relatively large classes. For example, on the eve of reform (January 1947), 32% of lower-track classes contained over 35 students (versus 5% for higher-track classes). Lower-track classes were also taught by relatively inexperienced teachers. For example, during this period, fewer than 10% of lower-track teachers had a university degree (versus 80% in higher-track schools).[30] These differences in class size and teacher characteristics likely mirror other resource differences (e.g., in buildings, books and equipment). For example, over the period 1946-1955, expenditure per year on the average lower-track student was estimated to be around 37% of that on the average higher-track student. (Vaizey, 1958, p.102)

Note that my results cannot be explained by a drop in resources driven by the reform itself. First, qualitative and quantitative evidence suggests that resources quickly adjusted

[29] For example, while Autor et al. (2003) argue that more-educated workers will hold comparative advantage in nonroutine versus routine tasks, I find little evidence that reform exposure led workers to perform more nonroutine tasks.

[30] These statistics were not collected after the War and hence these numbers refer to 1938. But there is no reason to expect the pattern to differ after the War. The typical lower-track teacher attended a higher-track school and then completed a two-year teacher training course (McNair, 1944).

to the increased enrollment. For example, class size increased only slightly.[31] Second, if the adjustment was slow, we would expect the drop in resources to be felt most strongly by the first cohort of students exposed to the reform. In fact, estimates are robust to dropping this cohort (see column (4) of Table 2).

B: Curriculum

Curriculum decisions were made by schools but constrained by government guidance and regulations. This guidance encouraged lower-track schools to provide education with a "practical bias" and discouraged them from providing traditional academic courses (i.e., oriented to specific subjects and potentially leading to exams).[32] All schools were prohibited from providing narrowly vocational courses. The available evidence is consistent with lower-track schools providing this type of practical education. First, I estimate that exposure to the compulsory schooling reform had no impact on the probability of receiving any academic or vocational qualifications (Appendix E.1). Second, official statistics suggest that the vast majority of lower-track schools provided courses in practical subjects (Appendix E.2). Third, newspaper articles and official reports provide numerous examples of lower-track school curricula including a significant practical component (Appendix E.2).

C: Incentives

Theoretical models suggest that when firms lack productivity information, students in tracked school systems will face weak incentives to exert effort (MacLeod and Urquiola, 2013). These incentives will be further weakened when, as in the setting studied here, the lower-track curriculum yields no external measures of student performance (e.g., test scores or exam results). These incentives could be yet further weakened if assignment to lower-track education demotivates students by lowering their self-esteem and reducing their perceived returns to study effort (Gamoran and Berends, 1987). Evidence on these points is hard to obtain, but we know that lower-track students completed little if any homework (Appendix E.3). To the extent that this reflects a weak demand for homework, it is consistent with weak incentives for effort.

D: Summary

With this discussion in mind, it seems plausible to suppose that the low quality of this additional schooling can explain why it failed to improve labor market productivity and later-

[31] Average class size in lower-track schools is 31.2 in 1947 (as shown in Table 1), and then 33, 32.6, 32.1 and 31.7 in 1948-1951 (figures from the relevant Ministry of Education Annual Reports).

[32] For example, lower-track schools were expected to provide "a good all-around secondary education, not focused primarily on the traditional subjects of the school curriculum, but developing out of the interests of the children." (HMSO (1947), p.29).

life outcomes. This is consistent with evidence on the relationship between school quality and these various inputs.[33] It is also consistent with key differences between this schooling and schooling that appears to have generated larger labor market returns. In particular, while compulsory schooling reforms in Sweden and Norway were associated with returns of around 5% and 10% (see Meghir and Palme (2005) and Aakvik et al. (2010) respectively), these reforms also detracked these school systems. Since Aakvik et al. (2010) suggest that the additional schooling likely improved on the lower-track schooling that preceded it, the additional schooling likely differed from that studied here.[34]

6 Conclusion

I study a compulsory schooling reform that yields arguably the best approximation yet found to large-scale experimental variation in schooling. I find that the schooling induced by this reform did little to advance students' labor market prospects. Based on a characterization of this schooling in terms of resources, curriculum and incentives, I conclude that these low returns are likely attributable to the low quality of this schooling. One caveat to this conclusion is that I cannot rule out that these low returns reflect some particular characteristics of the affected students and would be invariant to the type of schooling received. However, unless 80% of the British population could not have benefited from any schooling beyond age 14, at a time when around 60% of the US population graduated from high school (Snyder, 1993, Table 19), these findings must reveal more about the schooling provided than the students that received it.

To the extent that my findings are attributable to the content of the additional schooling involved, they illustrate the connection between the return to schooling and its content. Card and Krueger (1996) model this connection and it is explicit in many studies of the relationship between schooling returns and educational inputs (e.g., Jackson et al. (2016); Deming et al. (2016)). This connection - and the fact that education systems differ greatly across countries - implies that we would not necessarily expect the returns to schooling estimated in one setting to be the same as those estimated in another.[35] Ideally, future research will contribute to an evidence base of estimated returns that can be linked to specific features of a country's education system. This can complement research that is

[33]See Jackson (2018) on resources, Altonji et al. (2012) on curricula and Deming et al. (2016) on incentives.

[34]Ideally, higher-track schools would also provide a point of comparison. However, since very few higher-track students dropped out at age 14, returns cannot be calculated.

[35]Card and Krueger (1992) document heterogeneous returns across US states at a single point in time and for a specific group (white males - see Table 2). We would expect even more variation across countries and time periods.

focused on specific settings and intended to shed light on the impacts of specific educational inputs (e.g., resources and accountability).

With this agenda in mind, an important hypothesis that emerges from this analysis is that returns will be low when schooling is characterized by features associated with lower-track education. These are said to include low teacher expectations leading to less rigorous curricula, low student self-esteem leading to reduced effort and an inequitable distribution of resources that exposes lower-track students to large classes and less-effective teachers (Oakes, 2005; Gamoran and Berends, 1987; Gamoran, 2010). This hypothesis connects the analysis presented here to debates about tracking, specifically the claim that lower-track schooling will inevitably be low-quality schooling. The analysis presented here is consistent with that claim, but evidence from other settings would help to determine whether this is a broader phenomenon.

References

Aakvik, A., Salvanes, K. G., and Vaage, K. (2010). Measuring heterogeneity in the returns to education using an education reform. *European Economic Review*, 54:483–500.

Altonji, J. G., Blom, E., and Meghir, C. (2012). Heterogeneity in human capital investments: High school curriculum, college major, and careers. *Annual Review of Economics*, 4:185–223.

Angrist, J., Imbens, G. W., and Rubin, D. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, 91(434):444–472.

Angrist, J. and Pischke, S. (2014). *Mastering 'Metrics: The Path from Cause to Effect*. Princeton University Press, Princeton, NJ.

Angrist, J. D. and Krueger, A. B. (1991). Does compulsory school attendance affect schooling and earnings? *Quarterly Journal of Economics*, 106(4):979–1014.

Autor, D. H., Levy, F., and Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly journal of economics*, 118(4):1279–1333.

Barrow, L. and Malamud, O. (2015). Is college a worthwhile investment? *Annual Review of Economics*, 7:519–555.

Bingley, P. and Martinello, A. (2017). Measurement error in income and schooling and the bias of linear estimators. *Journal of labor economics*, 35(4):1117–1148.

Buckles, K. and Hungerman, D. (2013). Season of birth and later outcomes: Old questions, new answers. *The Review of Economics and Statistics*, 95(3):711–724.

Card, D. (2001). Estimating the returns to schooling: Progress on some persistent econometric problems. *Econometrica*, 69(5):1127–1160.

Card, D. (2008). Returns to schooling. In Durlauf, S. N. and Blume, L. E., editors, *The New Palgrave Dictionary of Economics*, pages 1–11. Palgrave Macmillan UK, London.

Card, D. and Krueger, A. B. (1992). School quality and black-white relative earnings: A direct assessment. *The quarterly journal of Economics*, 107(1):151–200.

Card, D. and Krueger, A. B. (1996). The labor market effects of school quality: Theory and evidence. In Burtless, G., editor, *Does Money Matter? The Effect of School Resources on Adult Success*. Brookings Institution, Washington, DC.

Card, D. E. (1999). The causal effect of education on earnings. In Ashenfelter, O. and Card, D. E., editors, *The Handbook of Labor Economics Vol 3a*. Elsevier/North Holland, Amsterdam.

Cattaneo, M. D., Idrobo, N., and Titiunik, R. (2019). *A practical introduction to regression discontinuity designs: Foundations*. Cambridge University Press.

Clark, D. and Royer, H. (2013). The effect of education on adult mortality and health: Evidence from Britain. *American Economic Review*, 103(6):2087 – 2120.

Clark, G. and Cummins, N. (2020). Does education matter? tests from extensions of compulsory schooling in england and wales 1919-22, 1947, and 1972.

Cygan-Rehm, K. (2021). Are there no wage returns to compulsory schooling in germany? a reassessment. Journal of Applied Econometrics, forthcoming.

Deming, D., Cohodes, S., Jennings, J., and Jencks, C. (2016). School accountability, postsecondary attainment, and earnings. *The Review of Economics and Statistics*, 98(5):848–862.

Devereux, P. and Hart, R. A. (2011). Forced to be rich? Returns to compulsory schooling in Britain. *Economic Journal*, 120:1345–1364.

Dolton, P. and Sandi, M. (2017). Returning to returns: Revisiting the British education evidence. *Labour Economics*, 48:87–104.

Galama, T. J., Lleras-Muney, A., and Van Kippersluis, H. (2018). The effect of education on health and mortality: A review of experimental and quasi-experimental evidence.

Gamoran, A. (2010). Tracking and inequality: New directions for research and practice. In Apple, M., Ball, S., and Gandin, L., editors, *The Routledge International Handbook of the Sociology of Education*, pages 213–228. Routledge, New York, NY.

Gamoran, A. and Berends, M. (1987). The effects of stratification in secondary schools: Synthesis of survey and ethnographic research. *Review of educational research*, 57(4):415–435.

Gelman, A. and Imbens, G. (2019). Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business & Economic Statistics*, 37(3):447–456.

Harmon, C. and Walker, I. (1995). Estimates of the economic return to schooling for the UK. *American Economic Review*, 85(5):1278–1296.

HMSO (1947). The new secondary education: Ministry of Education Pamphlet no. 9. London, UK.

Jackson, C. K. (2018). Does school spending matter? the new literature on an old question. NBER Working Paper 25368, Cambridge, MA.

Jackson, C. K., Johnson, R. C., and Persico, C. (2016). The effects of school spending on educational and economic outcomes. *Quarterly Journal of Economics*, 131(1):157–218.

Kane, T. J., Rouse, C. E., and Staiger, D. O. (1999). Estimating returns to schooling when schooling is misreported.

Kolesár, M. and Rothe, C. (2018). Inference in regression discontinuity designs with a discrete running variable. *American Economic Review*, 108(8):2277–2304.

Krueger, A. B. and Lindahl, M. (2001). Education for growth: Why and for whom? *Journal of Economic Literature*, 34:1101–1136.

Lowndes, G. A. (1947). *The Silent Social Revolution: An Account of the Expansion of Public Education in England and Wales, 1895-1965*. Oxford University Press.

MacLeod, B. and Urquiola, M. (2013). Anti-lemons: Reputation and educational quality. Mimeo, version March 13, 2013.

McNair, A. (1944). Teachers and youth leaders. a report of the board of education.

Meghir, C. and Palme, M. (2005). Educational reform, ability and family background. *American Economic Review*, 95(1):414–424.

Noack, C. and Rothe, C. (2019). Bias-aware inference in fuzzy regression discontinuity designs. *arXiv preprint arXiv:1906.04631*.

Oakes, J. (2005). *Keeping Track: How Schools Structure Inequality*. Yale University Press, New Haven and London, 2nd edition.

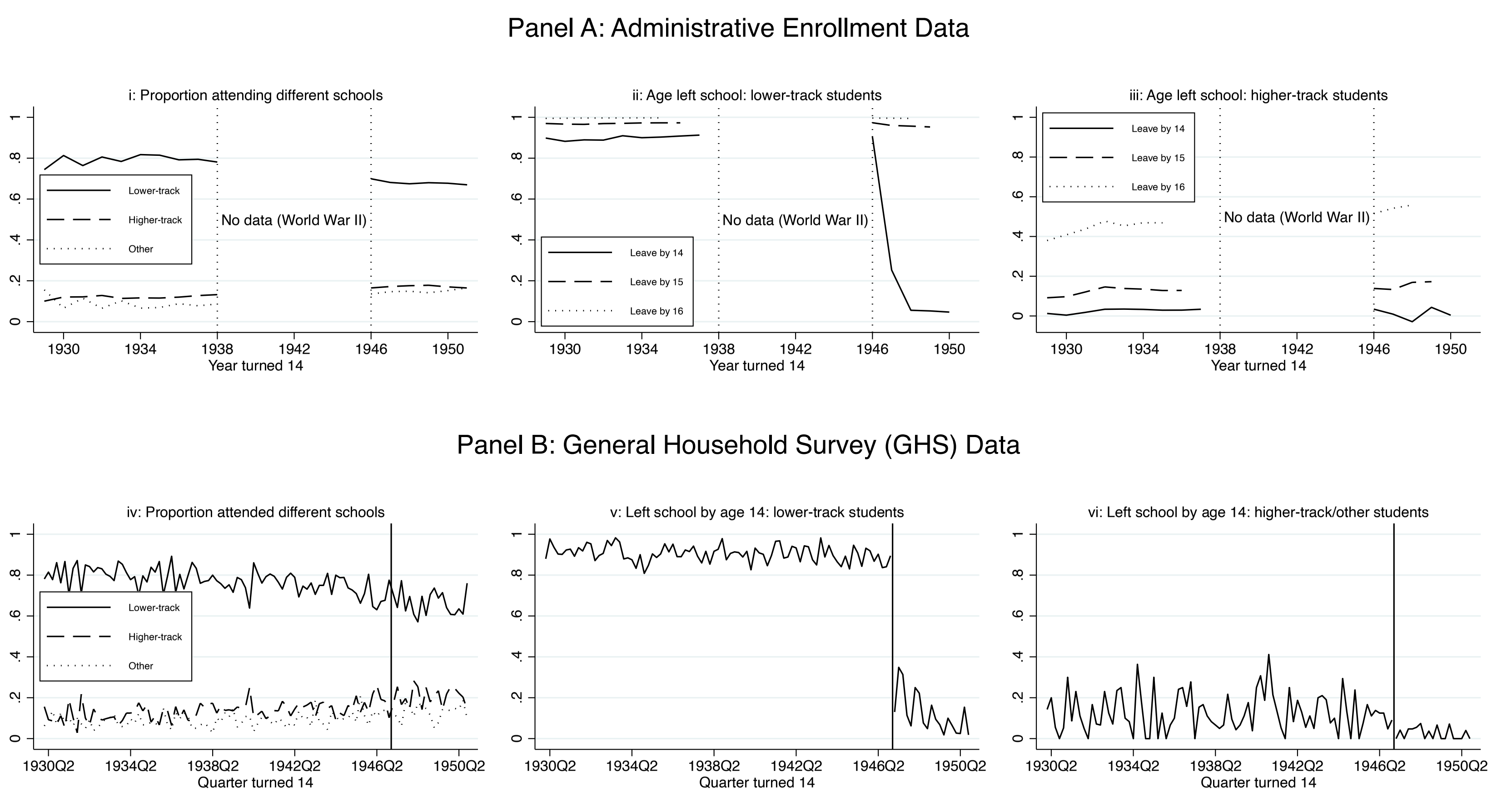
Oosterbeek, H. and Webbink, D. (2007). Wage effects of an extra year of basic vocational education. *Economics of Education Review*, 26:408–419.

Oreopoulos, P. (2006). Estimating average and local average treatment effects of education when compulsory school laws really matter. *American Economic Review*, 96(1):152–175.

Oreopoulos, P. (2008). Estimating average and local average treatment effects of education when compulsory schooling laws really matter: Corrigendum. *American Economic Review*. https://assets.aeaweb.org/asset-server/articles-attachments/

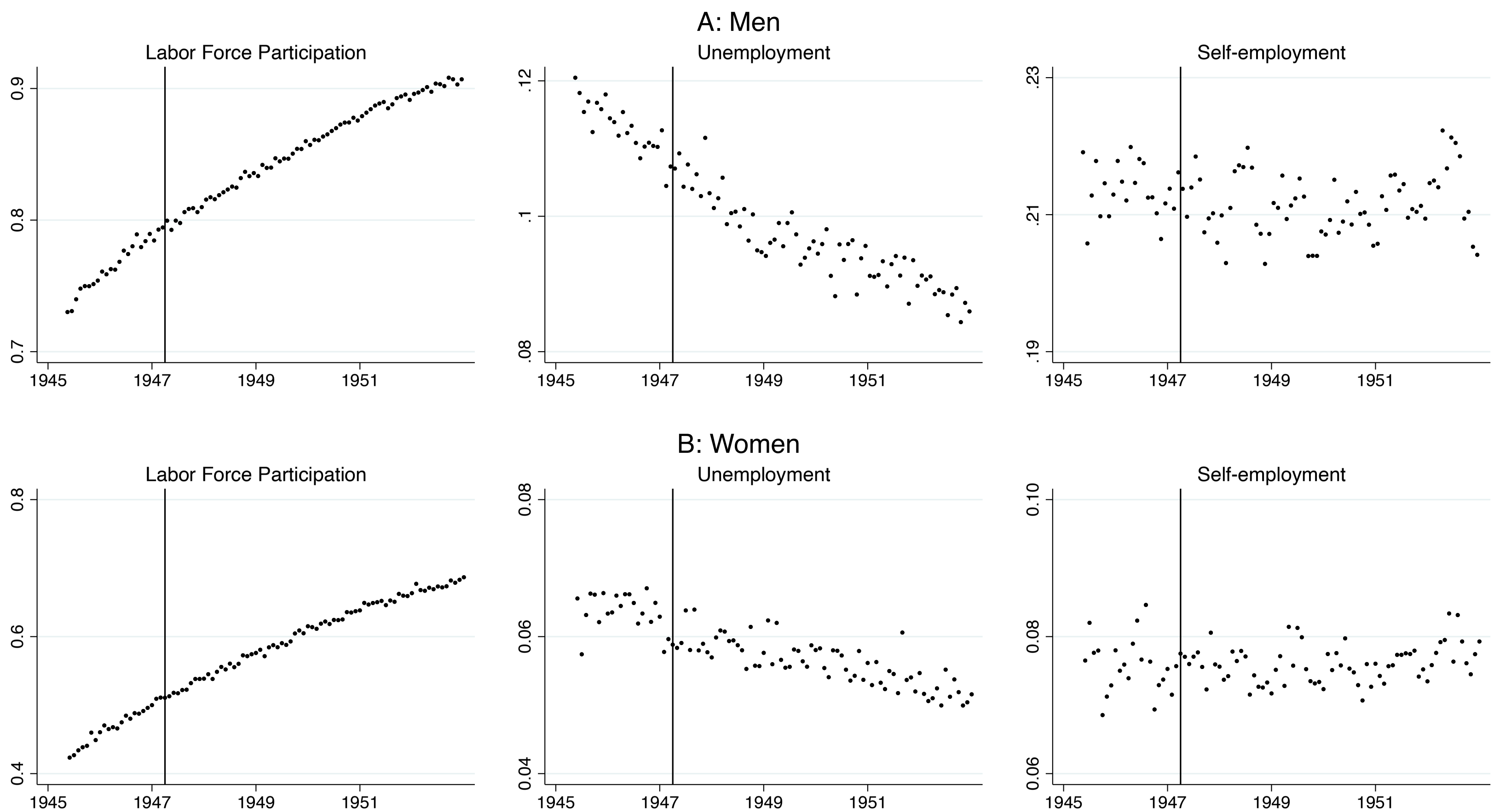
aer/contents/corrigenda/corr_aer.96.1.152.pdf.
Pischke, J.-S. and von Wachter, T. (2008). Zero returns to compulsory schooling in Germany: Evidence and interpretation. *Review of Economics and Statistics*, 90(3):592–598.
Snyder, T. D. (1993). *120 years of American education: A statistical portrait*. US Department of Education, Office of Educational Research and Improvement ....
Vaizey, J. (1958). *The costs of education*. George Allen and Unwin.
Zimmerman, S. D. (2014). The returns to college admission for academically marginal students. *Journal of Labor Economics*, 32(4):711 – 754.

Figure 1 Tracked school system and compulsory schooling reform



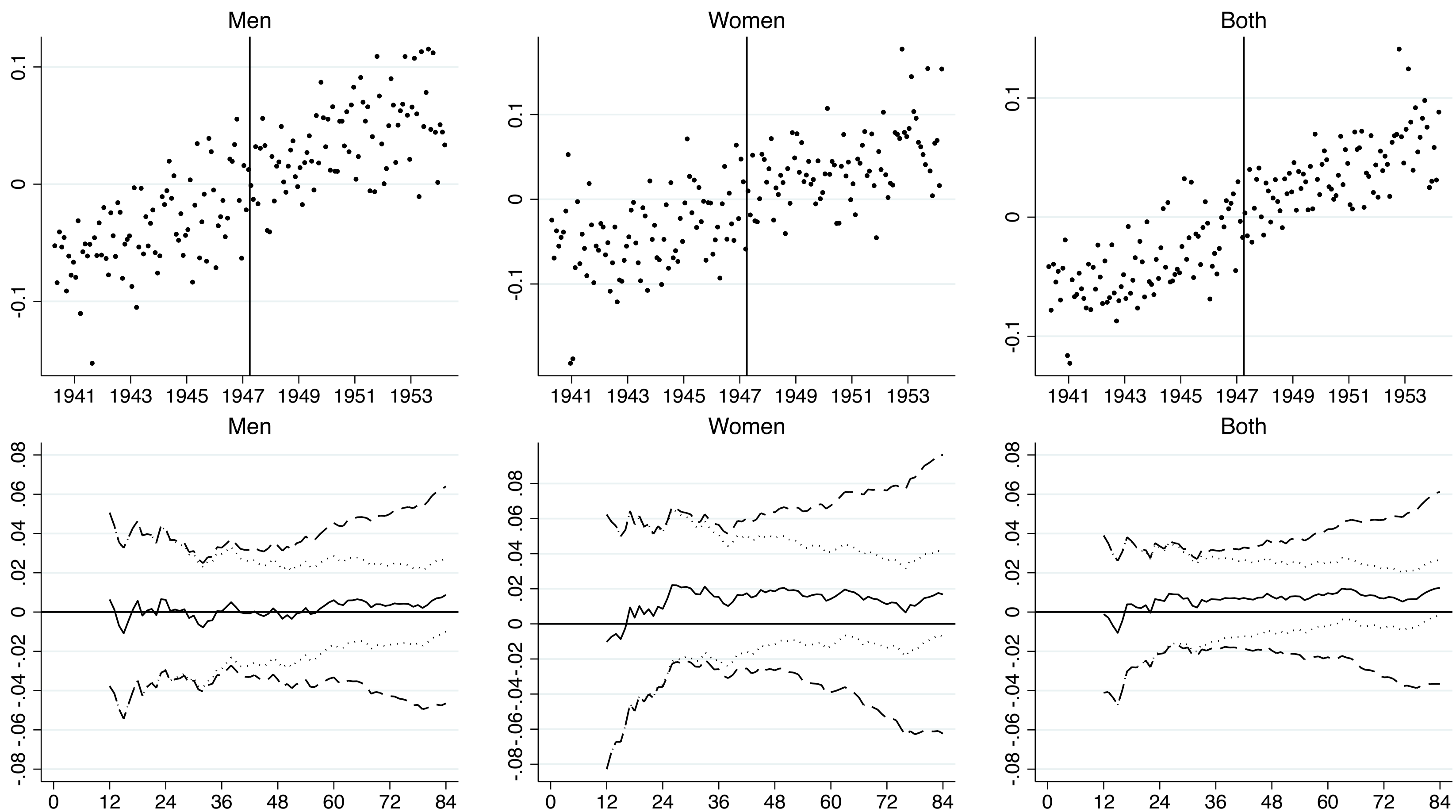
Notes: The graphs in Panel A use administrative enrollment data (see Appendix A.1). Figure 1i displays the proportion of 13 year olds in different types of school by year. Figures 1ii and 1iii display the proportion of 13 year olds in each year that will leave school by ages 14, 15 and 16 (by the types of school attended). The graphs in Panel B use GHS data (see Appendix A.2). Figure 1iv displays the proportion of respondents that attended different types of school by the quarter in which they turned 14. Figures 1v and 1vi display the fraction of respondents that left school by age 14 (by the types of school attended). See Appendix A.1 and A.2 for definitions of school type.

Figure 2 Impacts of compulsory schooling reform on labor market activity



Notes: The graphs display the cohort-specific fraction of individuals associated with the relevant outcome as reported on the 1991 Census. The sample includes only those born in England and Wales. Cohorts are indexed by the month in which they turned 14. The vertical line corresponds to 1 April 1947, when the compulsory schooling reform took effect.

Figure 3 Impacts of compulsory schooling reform on log weekly earnings



Notes: All of these graphs use L2 data (see text and Appendix A.3). The graphs in the top panel display cohort-specific average log weekly earnings (de-meaned to average zero in this sample). Cohorts are indexed by the month in which they turned 14. The vertical line corresponds to 1 April 1947, when the compulsory schooling reform took effect. The graphs in the bottom panel display regression discontinuity estimates of the impacts of this reform on that outcome. These estimates are based on least-squares estimation of equation (1). Estimates are displayed for every bandwidth from 12 months to 84 months. In addition to these estimates (solid lines), the graphs in the bottom panel display 95% confidence intervals based on robust standard errors (dotted lines) and Kolesár and Rothe (2018) "honest" confidence intervals (dashed lines).

Table 1 Descriptive statistics for cohorts born 1930-1933

Panel A: Family background, educational attainment and labor market outcomes

	Lower-track	Higher-track	Other	All
N	1034	284	188	1506
Family background				
Male	0.50	0.40	0.44	0.47
Professional	0.17	0.50	0.60	0.28
Skilled manual	0.52	0.39	0.28	0.46
Semi-skilled manual	0.32	0.10	0.13	0.25
Educational attainment				
Earned School Certificate	0.01	0.52	0.39	0.17
Apprenticeship	0.10	0.05	0.06	0.09
Attended College	0.04	0.14	0.24	0.09
Attended University	0.01	0.19	0.17	0.06
Labor market outcomes				
Labor force participation	0.82	0.78	0.77	0.80
Unemployment	0.03	0.02	0.01	0.02
Self-employment	0.10	0.14	0.14	0.11
Annual earnings (GBP 2020)	17,434	23,801	27,212	19,704
Occupation: Professional	0.10	0.27	0.32	0.16
Occupation: Intermediate	0.22	0.52	0.48	0.31
Occupation: Skilled manual	0.33	0.13	0.12	0.27
Occupation: Semi/unskilled manual	0.35	0.07	0.09	0.26

Panel B: Class size and teacher characteristics

	Lower-track	Higher-track	Other	All
Average class size	31.2	27.4	n/a	n/a
% classes with >30 students	56.8	40	n/a	n/a
% classes with >35 students	31.7	5.5	n/a	n/a
% teachers with university degree	7.1	78.4	n/a	n/a

Notes: Panel A uses GHS data (1972 and 1973 waves). See Appendix A.2 for more details on these data and the definitions used to construct the table. Panel B uses class size data from the Ministry of Education Annual Report 1947 (Table 10). These statistics refer to class size in January 1947. Panel B uses teacher characteristics from the Board of Education Annual Report 1938 (Tables 8 and 10). These statistics refer to teacher characteristics in March 1938 (this series was not continued after the War).

Table 2 Impacts of compulsory schooling reform on log weekly earnings

	Bandwidth: 2 years			Bandwidth: 4 years				Bandwidth: 6 years			
	(1)	(2)	(3)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Panel A: Male											
Reform	0.0063	0.0058	-0.0045	-0.00029	-0.00051	-0.0045	-0.0035	0.0035	0.0041	0.0035	0.0057
	(0.018)	(0.018)	(0.018)	(0.013)	(0.013)	(0.012)	(0.016)	(0.010)	(0.010)	(0.010)	(0.011)
Observations	48	48	48	96	96	96	84	144	144	144	132
Panel B: Female											
Reform	0.0086	0.0081	0.014	0.019	0.018	0.022	0.029	0.011	0.012	0.011	0.0046
	(0.022)	(0.022)	(0.023)	(0.015)	(0.015)	(0.016)	(0.020)	(0.013)	(0.013)	(0.013)	(0.015)
Observations	48	48	48	96	96	96	84	144	144	144	132
Panel C: Both											
Reform	0.0061	0.0058	0.0020	0.0079	0.0079	0.0073	0.011	0.0069	0.0077	0.0069	0.0054
	(0.014)	(0.014)	(0.014)	(0.0093)	(0.0093)	(0.0095)	(0.011)	(0.0076)	(0.0077)	(0.0076)	(0.0089)
Observations	48	48	48	96	96	96	84	144	144	144	132

Notes: For selected bandwidths (2 years, 4 years and 6 years), the table displays regression discontinuity estimates of the impacts of the compulsory schooling reform on log weekly earnings. The estimates (robust standard errors) in column (1) correspond to those displayed in the bottom panel of Figure 3. The estimates in column (2) are from models that include month-of-birth fixed effects. The estimates in column (3) are from models applied to age-adjusted outcomes (see text). The estimates in column (4) are based on samples that exclude individuals that turned 14 in the first 12 months after the reform. All of these estimates use L2 data (see text and Appendix A.3).

School Quality and the Return to Schooling in Britain: New Evidence from a Large-Scale Compulsory Schooling Reform

ONLINE APPENDICES

Contents

Appendix A. Additional Data Details

A.1: Administrative Enrollment Data

Panel A of Figure 1 uses data from Ministry of Education Annual Reports (1946-1951) and Board of Education Annual Reports (1928-1938). The tables in these Reports provide information on enrollment by age and school type, and estimates of the relevant population. I now describe how Figures 1i-1iii are constructed.

Figure 1i

To construct Figure 1i, I classify the population of 13 year olds into three groups:

- *Lower-track students*: in the period 1946-1951 these are students attending All-Age and Secondary Modern Schools; in the period 1928-1938 they are students attending Elementary Schools maintained by Local Education Authorities in England and Wales.
- *Higher-track students*: in the period 1946-1951 these are students attending Secondary Grammar Schools; in the period 1928-1938 they are students attending Grant-Aided Secondary Schools.
- *Other*: these are students attending neither lower- nor higher-track schools as defined above (i.e., the estimated 13 year old population less lower-track and higher-track students). The majority of this group will attend private schools.[1]

Figures 1ii and 1iii

To construct Figures 1ii and 1iii I calculate track-specific dropout rates separately for the periods 1946-1951 and 1928-1938. I now describe these two sets of calculations.

Ministry of Education Annual Reports (1946-1951)

The tables in these reports provide information on students enrolled in January of each year. I use this information to construct the fraction that drop out at age 14 in 1946 (for example) as follows:

$$D^{14}_{1946} = 1 - \frac{\#14_{Jan1947}}{\#13_{Jan1946}}$$

[1] In the period 1928-1938, the public (i.e., state) schools in this group include Non-Local Public Elementary Schools, Certified Efficient Public Elementary Schools, Certified Special Public Elementary Schools, Junior Technical Schools (including Junior Technical and Commercial Schools, Junior Housewifery Schools, Junior Departments in Art Schools and Schools of Nautical Training - see footnote to Table 2, Board of Education Annual Report 1938), Pupil-Teachers in Centres, Rural Pupil-Teachers and Nursery Schools. Combined, these schools likely enrolled less than 2% of the relevant age group. I include this group in the "other" category since I do not have information for them on the number of students enrolled by ages 14-14.25, 14.25-15, etc (e.g., see Table 6, Board of Education Annual Report 1938). In the period 1946-1951, the public schools in the "other" group include Secondary Technical Schools (the successors to Junior Technical Schools) and various types of special schools. Direct Grant schools are included in the "other" category throughout the period (these were private schools before 1946 and public schools after 1946).

The denominator in the second term on the right-hand side is the number of students aged 13 enrolled in January 1946. These students turned 14 in 1946 and could have left school at the end of the term in which they turned 14 (i.e., Easter 1946, summer 1946 or Christmas 1946). The numerator in the second term on the right-hand side is the number of students aged 14 in January 1947. Hence under weak assumptions the second-term is the proportion of students that enrolled in school beyond the minimum leaving age and D^{14}_{1946} is the fraction that dropped out at 14 (i.e., at the minimum leaving age) in 1946.[2] I calculate D^{15}_{1946} and D^{16}_{1946} in an analogous way: as the fraction of students aged 13 in January 1946 that left school by age 15 or by age 16.

Board of Education Annual Reports (1928-1938)

The tables in these reports provide information on students enrolled on 31st March each year. This presents a slight complication, since a subset of the students aged 13 on March 31st 1937 (for example) will not have reached the minimum school leaving age by March 31st 1938. Specifically, whereas students that turned 13 between April and December 1936 will have reached the minimum school leaving age by 31st March 1938, students that turned 13 between January and March 1937 may not have done.[3]

To deal with this problem I calculate the dropout rate at age 14 in 1937 (for example) for the students that were aged 13 on March 31st 1937 but that turned 13 between 1st April and 31st December 1936. Specifically, I calculate it as:

$$D^{14}_{1937} = 1 - \frac{\#[14.25 - 15]_{March1938}}{0.75 * \#13_{March1937}}$$

The numerator of the second term on the right-side is exactly the one I want: the set of students enrolled in March 31st 1938 aged between 14.25 and 15 (i.e., turned 13 between 1st April 1936 and 31st December 1936). The denominator in that term is an estimate of the total number that turned 13 between 1st April 1936 and 31st December 1936.[4] The assumption here is that the dropout rate among students that turned 13 between April 1936 and December 1936 is the same as the overall dropout rate among students that turned 13 in 1936.

Note that the number of 14 year olds in higher-track schools is not broken into the number aged between 14 and 14.25 and so on. I therefore calculate dropout for these schools using the full March enrollment counts. This will understate the dropout rate if the relevant denominator includes students yet to reach the minimum schooling age, but there is no such difficulty for Ministry of Education or GHS data and higher-track dropout rates are consistently low. Note that this problem does not affect the calculation of dropout by ages 15 and 16, since students at these ages have passed the minimum leaving age and can drop out at any time. I therefore calculate D^{15}_{1937} and D^{16}_{1937} using the full March enrollment counts.

[2] The key assumption is that 13 year old students did not continue their education in different types of school. This is consistent with the data. For example, of the lower-track students that left school in 1938 aged between 14 and 14.25, only 287 (less than 0.1%) switched to a higher-track school (Table 7, Board of Education Annual Report 1938).

[3] They will have reached it at the end of the term at which they turn 14, likely April 1938.

[4] Enrollment in these tables was reported for ages 14-14.25 etc but not ages 13-13.25 etc.

A.2: GHS Data

The General Household Survey (GHS) was conducted every year between 1971 and 2007 (except 1997-1998 and 1999-2000). It was based on a random sample of around 13,000 households in Great Britain. I use GHS data to construct Table 1 panel A, Figure 1 panel B and Table B1.

Table 1 (panel A)

I use the 1972 and 1973 waves (the only ones that contain information on school type). I keep respondents that were born in England and Wales and turned 14 in the three years before the 1947 reform.[5] I classify schools as follows (the % refers to this sample):

- *Lower-track schools*: Elementary including All-Age (52.35%) Central/Intermediate/Higher-Grade (3.72%), Secondary Modern (11.61%).
- *Higher-track schools*: Grammar (18.84%).
- *Other*: Comprehensive (0.16%), Private including Direct Grant (7.45%), Technical (2.25%), Other (3.61%).

Figure 1 (lower panel) and Table B1

The samples used here differ in two ways from those described above. First, I relax the requirement that individuals turned 14 in the three years before the reform, Second, I restrict to individuals surveyed in the 1973 wave (since the 1972 wave lacks information on birth month).

A.3: L2 Data

I obtained these data from the Department for Work and Pensions. The L2 data are drawn from the National Insurance Reporting System (NIRS). This is used to track National Insurance contributions and calculate entitlement to various National Insurance benefits. A person enters NIRS when they receive a National Insurance number. This usually occurs automatically, when they turn 16. If the number is assigned later (e.g., because the person is a migrant), the data are back-filled to age 16. Weeks employed are defined as employments for which an employer paid National Insurance Contributions ("Class 1 employments") hence they exclude those that generated pay below the lower earnings limit. Weeks employed are capped at 52 in the L2 data (although they could exceed 52 if employments run concurrently). The earnings measure excludes employer pension contributions.

[5]Since the 1972 wave does not contain birth month, I include respondents surveyed before April and aged 39-41 or surveyed after March and aged 40-41.

Appendix B: Track Composition of Affected Students

In Table B1 (below) I report regression discontinuity estimates of the impacts of the compulsory schooling reform on the probability that respondents left school aged 14 or below. For completeness I also examine total years of schooling measured as age left school. I report estimates for the full sample and two subgroups defined by the type of school attended: lower-track schools and other types of schools (i.e., higher-track or "other" schools as defined in Appendix A.2). The estimates are obtained from least-squares estimation of equation (1) in the paper, where Y is the relevant outcome (e.g., left school at age 14 or below) and R is cohort defined by month turned 14. I use the sample described in Appendix A.2. I report estimates for bandwidths of 2 years, 4 years and 6 years.

The Table reveals three main findings. First, the reform is estimated to have reduced the overall fraction that left school by age 14 by around 45-50%, with no evidence that the reform "spilled up" the education distribution. This mirrors prior findings from this literature (e.g., see Table 2 of Clark and Royer (2013)). Second, among students that attended lower-track schools, the reform is estimated to have reduced the fraction that left school at age 14 by around 65-70%, again with this effect commensurate with the effect on total years of schooling (i.e., no spillovers). Third, among students that attended other types of schools, the reform is estimated to have reduced the fraction that left school at age 14 by around 5-8% (on the margins of statistical significance).[6] These last two findings are consistent with the patterns seen in Figure 1.

I use the estimates reported in Table B1 to calculate the proportion of respondents affected by the reform that attended lower-track schools. That is, the proportion of "compliers" that attended lower-track schools. Let $C_i = \{0, 1\}$ and $L_i = \{0, 1\}$ be indicators for whether a GHS respondent was a complier with respect to the compulsory schooling reform and whether a respondent attended a lower-track school respectively. I then calculate this proportion as $P(L_i = 1|C_i = 1) = \frac{P(C_i=1|L_i=1)P(L_i=1)}{P(C_i=1)} = \frac{P(C_i=1|L_i=1)P(L_i=1)}{P(C_i=1|L_i=1)P(L_i=1)+P(C_i=1|L_i=0)P(L_i=0)}$. Using $\frac{N_L}{N}$, $|\widehat{\beta}_{L=1}|$ and $|\widehat{\beta}_{L=0}|$ to estimate $P(L_i = 1)$, $P(C_i = 1|L_i = 1)$ and $P(C_i = 1|L_i = 0)$, the estimated proportion is 0.959 and the associated standard error is 0.0199. This calculation is based on the estimates associated with a bandwidth of 6 years (i.e., column (5) of Table B1); I obtain the standard error by bootstrapping the whole procedure (using 5000 replications).

[6] Separate estimates for higher-track and "other" schools (available on request) are similar, but imprecisely estimated on account of the small sample sizes.

Table B1 Impacts of compulsory schooling reform on age left school

	Bandwidth: 2 Years		Bandwidth: 4 Years		Bandwidth: 6 Years	
	Left<= 14	Years of Sch.	Left<=14	Years of Sch.	Left<=14	Years of Sch.
Panel A: Full sample						
Reform	-0.450***	0.443**	-0.441***	0.451***	-0.482***	0.502***
	(0.055)	(0.163)	(0.037)	(0.114)	(0.029)	(0.090)
Observations	1003	1003	2106	2106	3267	3267
Panel B: Attended lower-track schools						
Reform	-0.627***	0.592***	-0.637***	0.654***	-0.665***	0.665***
	(0.057)	(0.088)	(0.039)	(0.066)	(0.030)	(0.048)
Observations	679	679	1461	1461	2287	2287
Panel C: Attended higher-track or other schools						
Reform	-0.051	0.160	-0.057	0.143	-0.060	0.048
	(0.057)	(0.360)	(0.040)	(0.252)	(0.032)	(0.195)
Observations	324	324	645	645	980	980

Notes: For selected bandwidths (2 years, 4 years and 6 years), the table displays regression discontinuity estimates (robust standard errors) of the impacts of the compulsory schooling reform on the probability of leaving school by ages 14 and on years of schooling. The estimates come from least-squares estimation of equation (1) in the paper. These estimates use GHS data (1973 wave) - see Appendix A.2.

Appendix C. Prior Literature

In this Appendix I survey the prior literature and compare my findings to those reported in prior studies. Oreopoulos (2006) was the first paper to use regression discontinuity methods to estimate the impacts of the compulsory schooling reform. Using a sample of General Household Survey (GHS) data that pooled men and women, his preferred specifications implied that an additional year of schooling induced by this reform increased earnings by 8-15% (Oreopoulos, 2008).[7] Devereux and Hart (2011) used GHS data to replicate the Oreopoulos (2008) estimates. They argued that these estimates masked important differences between the impacts on men and women. For women, Devereux and Hart (2011) found no statistically significant impacts in earnings. For men, they found statistically significant impacts on the order of 4-7%. They also reported similar estimates based on the New Earnings Survey Panel Dataset (NESPD), a 1% sample of the British population.

To help reconcile these findings with mine, I start by replicating the Devereux and Hart (2011) GHS estimates. I focus on the "reduced-form" estimates and report these in Table C1 (below).[8] The different columns ("Left<=15" and "Left>=16") correspond to two subsets of the full sample: those that left school at age 15 or below and those that left at age 16 or above. The different panels correspond to different polynomial specifications. For the two outcomes considered (Log Hourly Earnings and Log Weekly Earnings), the estimates reported by Devereux and Hart (2011) are those corresponding to the "Full sample" and "Panel C: 4th-order polynomial". Looking across the columns, we see that these estimates are larger among students that left at age 16 or above than among student that left at 15 or below. Since the reform did not impact the probability of staying in school beyond 15 (see Table B1), this is the opposite of what we would expect. Looking down the rows, we see that estimates are sensitive to polynomial choices (e.g., compare the "Full Sample" estimates in panels B and C), as Dolton and Sandi (2017) have already shown. Finally, note that the reported standard errors are clustered by year of birth - a common practice but one criticized by Kolesár and Rothe (2018) - and that robust standard errors are around 30% larger than these. With all of this in mind, I conclude that these estimates do not provide robust evidence that the compulsory schooling reform improved earnings among men.

The NESPD is not publicly available and hence it is not easy to replicate the NESPD estimates reported by Devereux and Hart (2011). However, these estimates are smaller than their GHS estimates, albeit still statistically significant. They are also based on the same specification and hence subject to the same concerns about the global polynomial approach (Gelman and Imbens, 2019). As another illustration of the limitations of this approach - and an attempt to reconcile my findings with these NESPD estimates - I fit these models to my L2 data (i.e., I aggregate to year of birth and fit high-order polynomials to data covering 30 birth years). Despite the impression given by Figure 3, these estimates are often positive and sometimes statistically significant. For example, focusing on log weekly earnings, excluding the partially-treated cohort born in 1933 and using second- through fifth-order polynomials

[7]An earlier paper by Harmon and Walker (1995) documented the existence of the 1947 reform and a later 1972 reform that increased the minimum leaving age from 15 to 16. Using instrumental variables methods to exploit both reforms, they estimated earnings returns on the order of 15%. Card (1999) argued that these estimates were based on strong assumptions about the age profile of earnings.

[8]Their code can be found on the Economic Journal website.

Table C1 Devereux and Hart (2011): Sensitivity to sample and specification

	Log Hourly Earnings			Log Weekly Earnings		
	Full sample	Left<=15	Left>=16	Full Sample	Left<=15	Left>=16
Panel A: 2nd-order polynomial						
Reform	0.010	0.001	0.015	0.021	0.019	0.001
	(0.011)	(0.013)	(0.025)	(0.012)	(0.015)	(0.027)
Observations	46995	29828	17167	46995	29828	17167
Panel B: 3rd-order polynomial						
Reform	0.011	-0.010	0.019	0.013	0.002	-0.001
	(0.012)	(0.014)	(0.024)	(0.015)	(0.017)	(0.026)
Observations	46995	29828	17167	46995	29828	17167
Panel C: 4th-order polynomial						
Reform	0.022	0.004	0.050	0.032*	0.022	0.039
	(0.013)	(0.015)	(0.032)	(0.014)	(0.018)	(0.030)
Observations	46995	29828	17167	46995	29828	17167
Panel D: 5th-order polynomial						
Reform	0.022	0.008	0.058	0.032*	0.026	0.047
	(0.015)	(0.016)	(0.034)	(0.015)	(0.020)	(0.030)
Observations	46995	29828	17167	46995	29828	17167

Notes: The table reports estimates of the impacts of the compulsory schooling reform on log hourly earnings and log weekly earnings. The estimates in Panel C under "Full Sample" are based on the same sample and specification used by Devereux and Hart (2011) and hence are identical to those reported in that paper. Starting from this sample and specification, the estimates in other panels come from changing the order of the polynomial used to control for cohort. The estimates in the other columns come from splitting the sample into respondents that left school at age 15 or below and respondents that left school at age 16 or above.

in birth year, I obtain estimates (robust standard errors) of 0.015 (0.006), -0.001 (0.006), 0.012 (0.006) and 0.013 (0.007). If I interact these polynomials with the reform dummy, I obtain estimates (robust standard errors) of 0.003 (0.007), 0.013 (0.013), 0.037 (0.015) and -0.020 (0.019).

Finally, in Appendix C of their paper, Clark and Royer (2013) also used GHS data to estimate the earnings impacts of the compulsory school reform. They used only GHS waves that contained information on birth month and implemented a "local linear" approach. They reported positive and statistically significant earnings estimates for men, although these were noisy and sensitive to the outcome and the sample used. For example, as seen in panel A of Appendix Table C1 of their paper, the estimate (robust standard error) is 0.028 (0.03) when the outcome is log hourly earnings and the sample is restricted to those working normal hours (35-50 hours per week). This sensitivity is a possible reflection of the fact that this sample covers the years 1986-1995, when the cohorts born around 1933 are aged 55-65 and gradually withdrawing from the labor force.

In another attempt to apply a local linear approach to GHS data, I obtained data from the 1973-1976 waves. These do not include information on weekly or hourly earnings and have not been used in any prior studies of the reform. They do however include birth month and information on annual earnings (from the prior year). This facilitates an analysis that is based on local linear methods and that uses earnings measured when the relevant cohorts were in their 40s. The estimates are noisy (the sample is relatively small), but are consistent with the schooling induced by the 1947 reform having close to zero impacts on earnings. For example, using a bandwidth of 40 months on either side of the April 1933 threshold (the same bandwidth used by Clark and Royer (2013)), my reduced-form estimate (robust standard error) is 0.008 (0.027). More details are available on request.

Appendix D. Additional Discussion

I interpret my findings as suggesting that the additional schooling induced by the reform did little to improve students' labor market productivity. Consistent with this, I show in this Appendix that the additional schooling had little impact on the types of skills actually used on the job. I then consider and reject some alternative interpretatations.

D.1: Evidence on Skill Use

To investigate impacts on skill use, I use the occupational skill use measures first used by Autor et al. (2003). The definitions and associated task interpretations are as follows:[9]

- DCP (Direction, Control and Planning): Measure of nonroutine analytic tasks.
- EYEHAND (Eye Hand Foot Coordination): Measure of nonroutine interactive tasks.
- FINGDEX (Finger Dexterity): Measure of routine manual tasks.
- MATH (GED Math): Measure of nonroutine analytic tasks.
- STS (Set Limits, Tolerances or Standards): Measure of routine cognitive tasks.

These measure various dimensions of skill use on a continuous scale from 0 to 10, where higher numbers refer to more intensive use of the task. These measures are coded for the US 1980 Classification of 3-digit Occupations. I construct measures for the UK 1990 Classification of 3-digit Occupations using the crosswalks provided by Lambert and Prandy (2018). I then construct averages of these measures by 2-digit occupation (available in the 1991 Census data).

In Figure D.1 I plot these measures separately for men and women and by birth month cohort. In Figure D.2 I plot RD estimates (and robust and honest confidence intervals) of reform impacts on these measures of occupational skill. As in the main analysis, I use bandwidths between 0 and 24 months. For both men and women, point estimates are small (relative to the across-occupation variation in these measures reported on the graphs in Figure D.2). Moreover, for none of these measures and for neither men nor women do we see obvious impacts on these outcomes.

D.2: Alternative Interpretations

I now consider - and reject - two alternative interpretations.

Undetected impacts on productivity?

An alternative interpretation is that lower-track schooling could have increased productivity without increasing earnings. This could be the case if wages were rigid. It could also be the case if the compulsory schooling reform reduced the earnings return to additional schooling

[9]These are taken from Appendix 1 of Autor et al. (2003). More detailed definitions can be found there.

(i.e., general equilibrium effects). Against rigid wages, I found no impacts on margins that should be unaffected by wage rigidities (e.g., unemployment). Against general equilibrium effects, different types of non-college workers are typically thought to be perfect substitutes in production (e.g., Card (2009)), such that the wage return to lower-track schooling should be unrelated to the fraction of students that leave at age 15 rather than age 14.

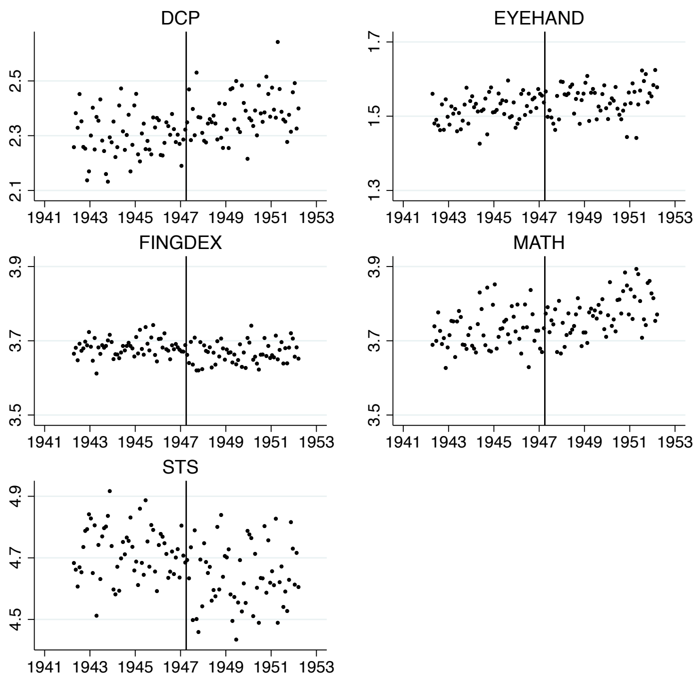
Undetected impacts on earnings?

It is also possible that the additional schooling had important earnings impacts that my analysis failed to detect. This could be because positive returns to schooling were offset by positive returns to experience (i.e., among workers of the same age, more schooling implies less experience) or because I cannot observe workers before age 40 and hence could have missed early-career impacts.[10] Against the experience hypothesis, I look at workers in their forties and fifties, likely on the flat part of their earnings-experience profile (Heckman et al., 2008). Early-career impacts are hard to reconcile with the absence of any effects of the reform on proxies for wealth such as home ownership.

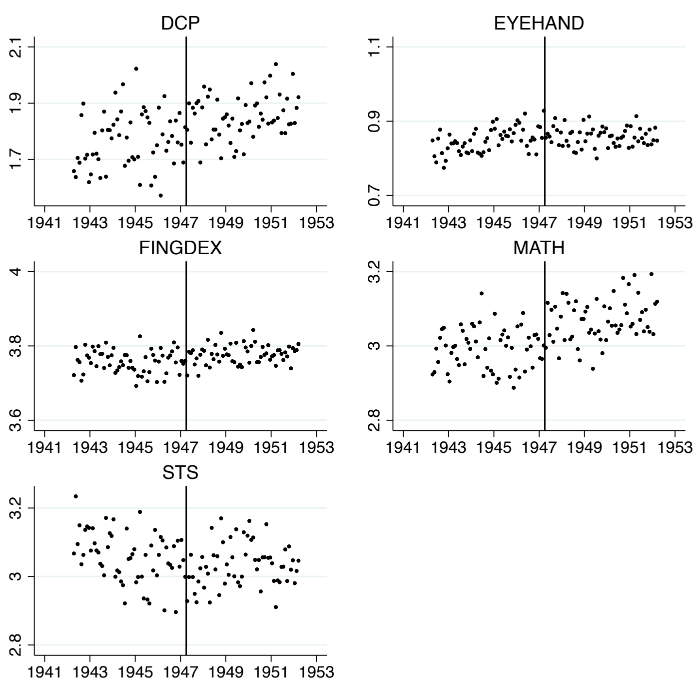
[10]For example, in a different context, Buscha and Dickson (2015) find that the returns to schooling vary over the life-cycle.

Figure D1 Compulsory schooling reform and skill use

(a) Men



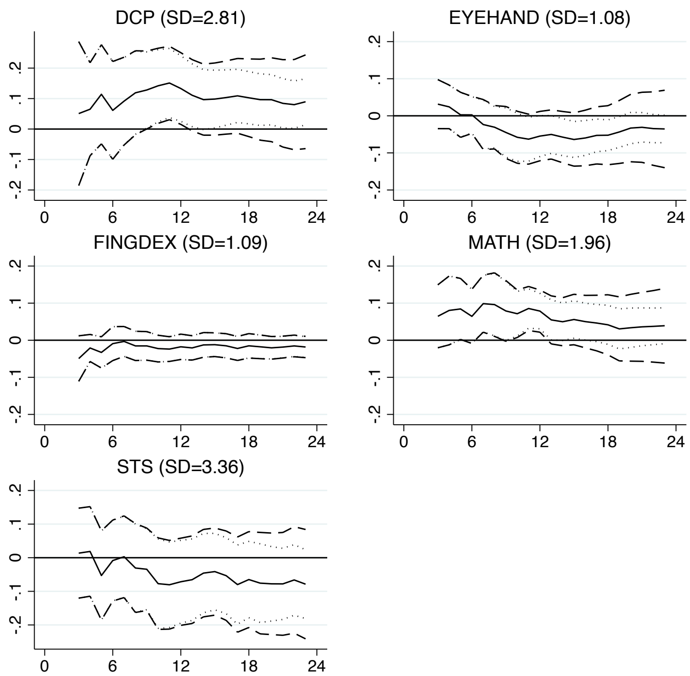
(b) Women



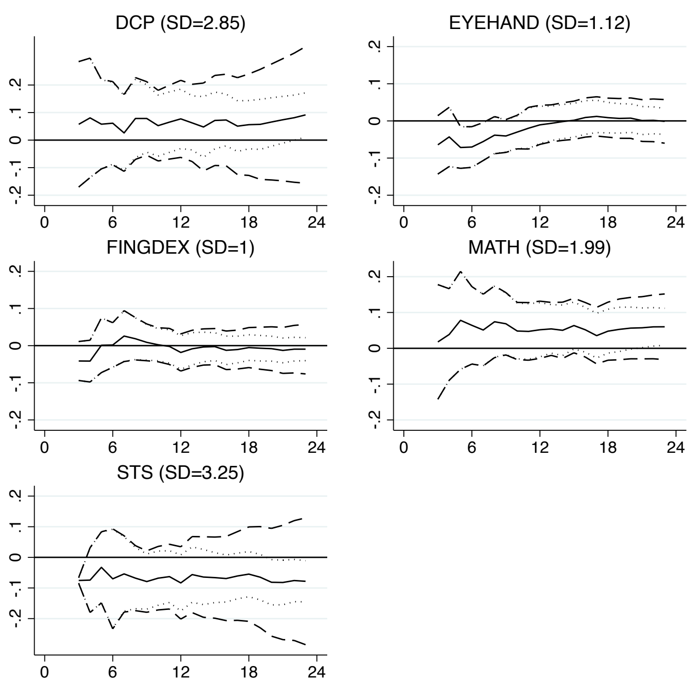
Notes: The graphs display cohort-specific average skill use as computed from the occupation recorded at the 1991 Census. The sample includes only those born in England and Wales. See text for definitions of skill use.

Figure D2 Impacts of compulsory schooling reform on skill use

(a) Men



(b) Women



Notes: See notes to Figure 3 (analogous to this Figure) for details of the estimates and standard errors represented on this graph.

Appendix E. The Content of Lower-Track Schooling

In this Appendix I provide additional details on the content of lower-track schooling.

E.1: Qualifications

I argue in the paper that lower-track schools provided education that was neither academic (in the sense of being organized around specific subjects and oriented to external examinations) nor vocational (in the sense of training students for particular types of jobs). In support of this claim, I now show that the compulsory schooling reform had little impact on qualification receipt.

I use data from the 2011 Census. The obvious disadvantage of these data is that the relevant cohorts are in their late 70s by 2011. But the reform had no obvious impact on mortality (Clark and Royer, 2013), and these data include detailed qualification information for the entire population. I define academic and vocational qualifications as follows:

- *Academic*: I code respondents as having an academic qualification if they report having a qualification in either categories 1 or 2 below.[11] Since the only academic qualifications available to students that left school in the late 1940s were School Certificates (O levels were not introduced until the 1950s), I could have used category 2 only. I add category 1 in case respondents conflate School Certificate exams with O Levels.

    1. 1 - 4 O levels / CSEs / GCSEs (any grades), Entry Level, Foundation Diploma.
    2. 5+ O levels (passes)/CSEs (grade 1)/GCSEs (grades A*-C), School Certificate, 1 A level/ 2-3 AS levels/VCEs, Higher Diploma.

- *Vocational*: I code respondents as having a vocational qualification if they report having a qualification in any of the following categories. Many of the qualifications included in these categories were not established until well after the relevant cohorts left school, but these categories should include any vocational qualifications that respondents might have received as part of their studies.

    1. NVQ Level 1, Foundational GNVQ, Basic Skills.
    2. NVQ Level 2, Intermediate GNVQ, City and Guilds Craft, BTEC First/General Diploma, RSA Diploma.
    3. NVQ Level 3, Advanced GNVQ, City and Guilds Advanced Craft, ONC, OND, BTEC National, RSA Advanced Diploma
    4. Other vocational/work-related qualifications

Respondents were also asked whether they had received an Apprenticeship and whether they had any higher-level academic qualifications (e.g., A levels, Degree) or professional qualifications (e.g., teaching, nursing). To ensure that my definitions of academic and vocational qualifications do not miss any impacts on these other types of qualifications I examine four qualifications outcomes: having an academic qualification (as defined above), having a

[11] These are the actual categories used on the Census form.

vocational qualification (as defined above), having an apprenticeship and having no qualifications.

In Figure E1 (below) I plot these outcomes by cohort. These graphs suggest that the compulsory schooling reform had little impact on qualification receipt. Regression discontinuity estimates confirm this impression.[12] It is striking that large fractions of the relevant cohorts received no qualifications, especially since these numbers will understate the fraction that received no qualifications in school (since some respondents will report qualifications received later). Assuming that most higher-track students received some qualifications, it follows that the vast majority of lower-track students earned no academic or vocational qualifications in secondary school.

E.2: Official Statistics, Newspaper Reports and Official Reports

I note in the paper that lower-track schools were discouraged from providing traditional academic courses and encouraged to provide education that was not vocational but that had a "practical bias". I now describe evidence consistent with these schools providing education with a "practical bias". This evidence comes from official statistics, newspaper reports and official reports.

Official statistics

Before the War, Board of Educational Annual Reports contained tables documenting the types of practical instruction provided by schools. Although these statistics were not continued after the War, it seems to reasonable to assume that the lower-track curriculum was broadly stable over this period. According to the last table published before the War (Board of Education Annual Report 1938, Table 18), only 10% of lower-track schools made no provision for practical instruction. Among the remainder that did, subjects included domestic subjects, woodwork and metalwork, gardening, rural science, dairywork and arts and crafts (described as "Bookbinding, Leatherwork and Weaving, etc").

Newspaper reports

Evidence on the curriculum also comes from newspaper reports of the plans made for the "extra year" of schooling resulting from the compulsory schooling reform. In particular, in the months before the reform was introduced, the Times Educational Supplement reported on the plans for the extra year being made by 26 local authorities around the country.[13] The curriculum was mentioned in 16 of these 26 cases. The exact references are reproduced below (in the order they appear in the newspaper):

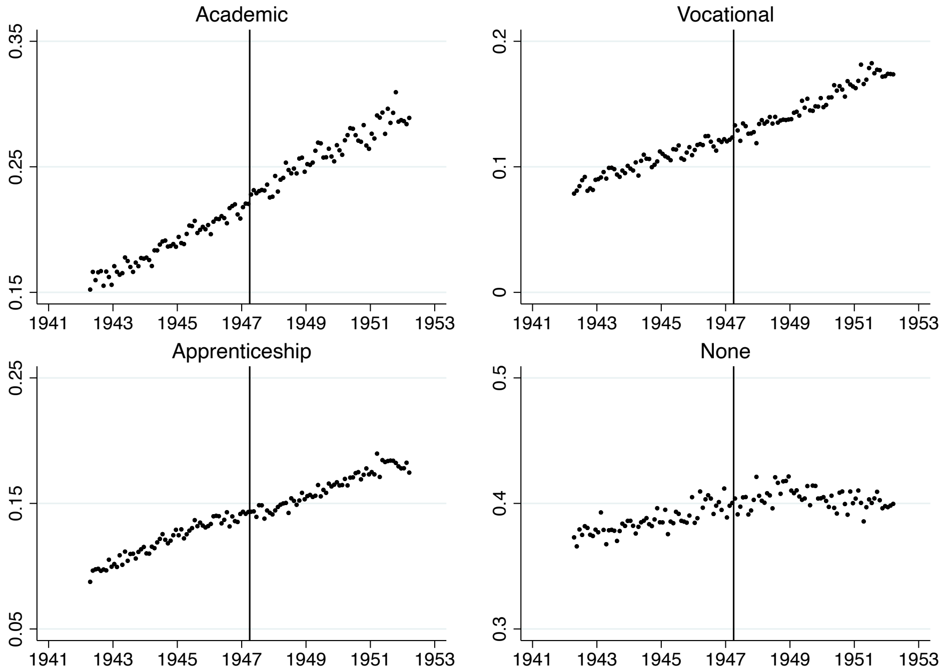
- West Riding: "It has been decided that the curriculum during the extra year must be related to the previous three years work and that all pupils shall have some experience in practical subjects, though courses in the arts, social studies, English, commerce, and Mathematics will be arranged as required."

[12] For some bandwidths these estimates reveal marginally significant positive effects on the receipt of academic qualifications. However, the magnitudes are small (on the order of 0.005) and these estimates are sensitive to bandwidth and specification choices (e.g., the inclusion of month-of-birth fixed effects). These estimates are available on request.

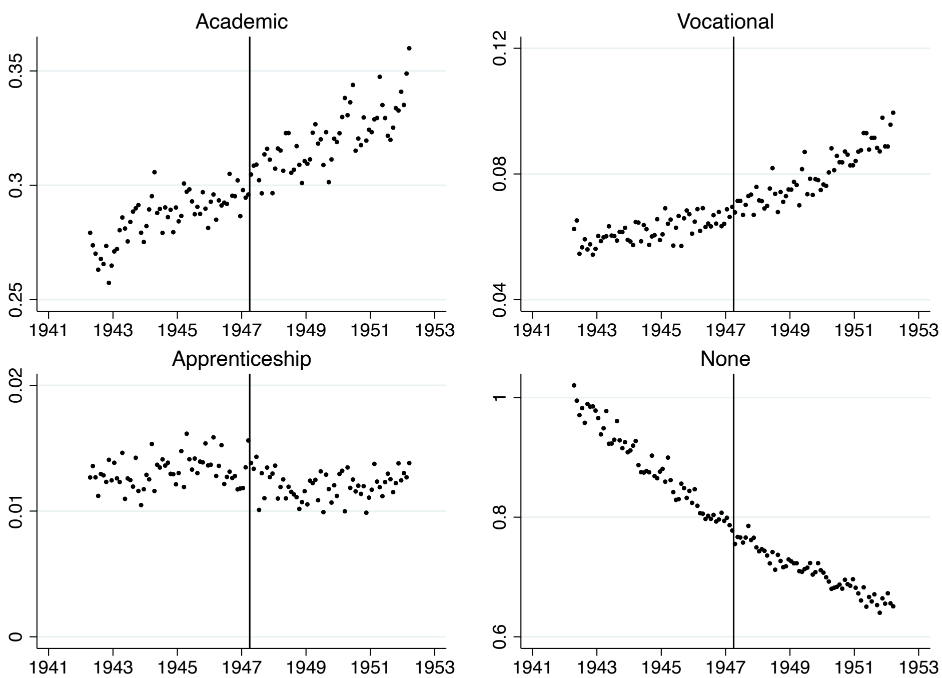
[13] The Times Educational Supplement was a widely-circulated weekly newspaper aimed at teachers. These reports can be found in the 27 March 1947 and 5 April 1947 issues.

Figure E1 Compulsory schooling reform and qualifications received

(a) Men



(b) Women



Notes: The graphs display the cohort-specific fraction of individuals holding different types of qualifications as reported on the 2011 Census. The sample includes only those born in England and Wales. See Appendix E.1 for qualification definitions.

- Holland: “The curriculum will be left to the decision of the head masters concerned, but emphasis will be laid on practical subjects. Preparations are being made in Secondary Modern schools for the development of technical subjects based on agriculture and the building industries.”
- Kevesten: “A special sub-committee has recommended that the curriculum during the extra year should be flexible with no complete break from previous years work”
- Derbyshire: “The instruction will include domestic subjects and the science of the countryside as well as normal school subjects”.
- Shropshire: “The curriculum is to be regarded as an extension of the three-year course with a desirable minimum of 3.5 days practical work a week, including project work and outdoor activities.”
- Birmingham: “The secondary education sub-committee of the education committee suggests that the conception of the “young citizen” might well inspire the modern school course, particularly in its last year, with particular reference to civic, practical subjects, and the extended use of project methods.”
- Sheffield: “While it is not intended to give specific vocational training, the development of fundamental skills applicable to future careers is covered.”
- Carnaervonshire: “Modern schools..will lay emphasis during the extra year on practical work closely integrated with the industrial life of the particular school district.”
- Pembrokeshire: “The authority is now considering interim curriculum requirements, which will include practical and cultural subjects with a “modern” approach.”
- Devon: “There will be no attempt to produce strictly vocational courses in the extended curriculum, but schools will be encouraged to provide children with the broad general background they will need in their future work.”
- Cornwall: “..More teachers have been engaged to undertake the practical work planned for the extra curriculum.”
- Southampton: “..Head masters are considering the establishment of a course on such subjects as shop display and dressmaking.”
- Portsmouth: “..it is proposed to convert a number of air-raid shelters into workshops for practical work.”
- Kent: “No attempt will be made at vocational training for the extra year, but greater emphasis will be given to the practical side, with an appreciation of craftsmanship and its importance to the community.”
- West Suffolk: “In rural schools practical courses in animal and plant husbandry are being arranged.”
- Norfolk: “The authority is investigating the possibility of starting one-month residential courses with an emphasis on practical subjects.”

Official Reports

Official reports provide additional clues as to the content of education provided by lower-track schools. First, these reports consistently recommended that traditional subjects be abolished and instead emphasized the importance of “practical bias”. Although they rarely provided recommendations as to how practical bias should be incorporated into the curriculum, HMSO (1947) described the “project method” as “one of the most promising lines of solution” (p.38). Second, some official reports summarized schools inspectors’ descriptions of the curriculum actually taught. For example, looking back at the curriculum followed during the “extra year” of education introduced in 1947, HMSO (1948) provides the following summary:

> Some schools sought to make the year profitable either by consolidating work already done or by teaching new subjects and approaching old subjects in a new way. Where schooling had been interrupted by the war, it was probably wise to take the opportunity of improving the children’s proficiency in the basic school subjects. But this method of consolidation, particularly where it was not specially justified or where it was overdone, often tended to produce boredom and frustration. The livelier schools planned new courses and introduced a great many experiments. Children were given a choice of subjects and were taught to undertake pieces of work involving research and the collation of material, either on their own or in small groups without continuous supervision. Where premises and staff were available the amount and range of practical work were increased, especially in its more interesting and exacting aspects, such as metalwork, and fruit culture and bee-keeping in rural areas. (HMSO, 1948, paragraph 5)

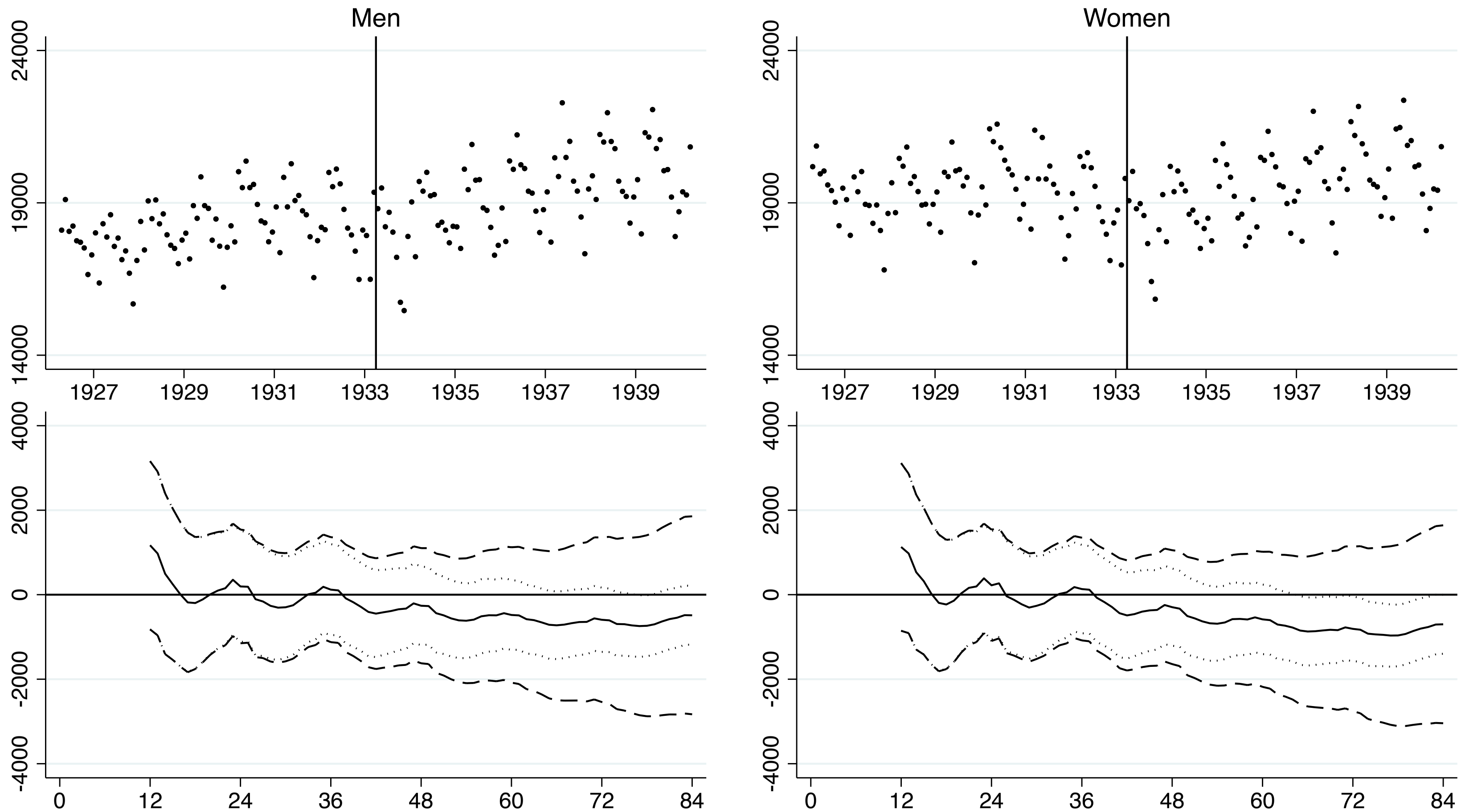
E.3: Homework Assigned to Students

I argue in the paper that lower-track students likely lacked incentives to exert effort in school. Consistent with that hypothesis, the evidence suggests that lower-track schools assigned little or no homework. In particular, in the late 1930s the Board of Education asked School Inspectors to report on the homework being set by schools. Drawing on these reports, Board of Education (1937) discussed various issues related to homework. Referring to lower-track students, they noted that "Their normal school course does not make homework necessary; nor in fact is homework set, or asked for, except in the case of a small minority, estimated at between 5 per cent and 10 per cent, who are working for some examination".(Board of Education, 1937, p.20) In contrast, they noted that higher-track students were expected to devote 1-2 hours per day to homework.(Board of Education, 1937, p.30) There is little reason to think that more homework was assigned to lower-track students over the following decade. Indeed, as Board of Education (1937) noted, the practical nature of the lower-track curriculum reduced the scope for assigning homework.

Appendix F. Robustness Checks

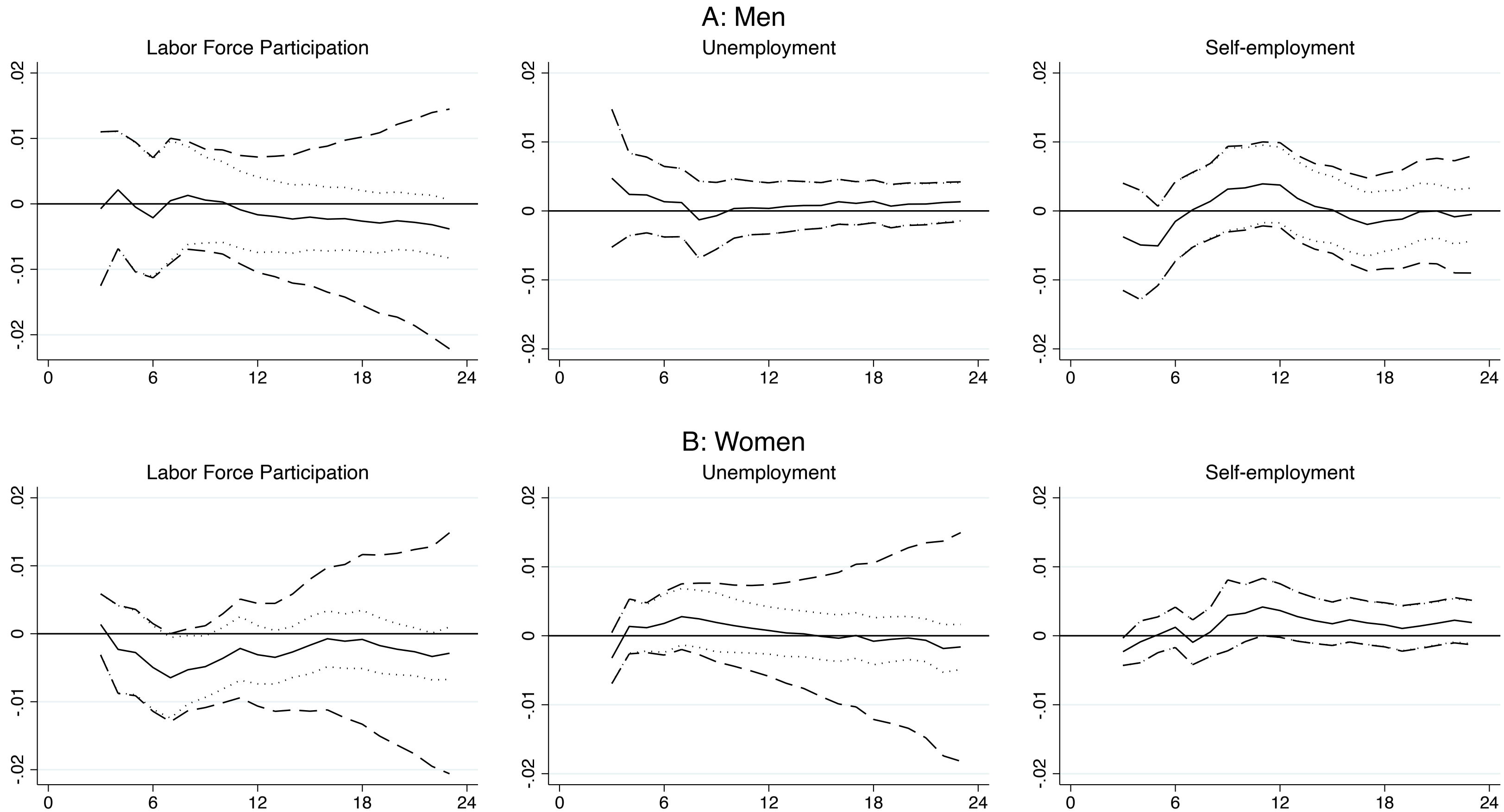
This Appendix contains the following Figures and Tables:

Figure F1 Impacts of compulsory schooling reform on resident population



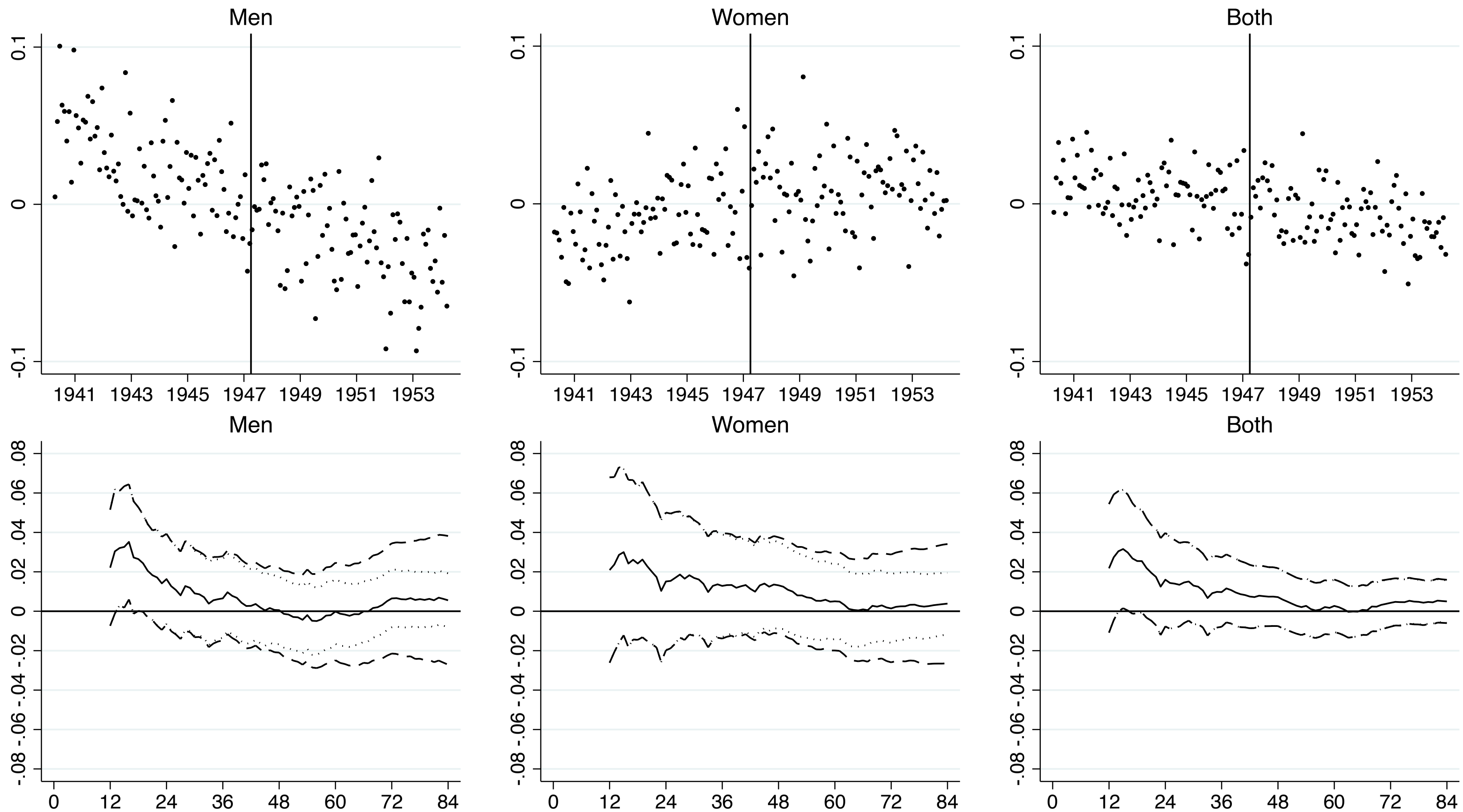
Notes: The graphs in the top panel display the cohort-specific resident population born in England and Wales as enumerated at the 1991 Census. Cohorts are indexed by the month in which they turned 14. The vertical line corresponds to 1 April 1947, when the compulsory schooling reform took effect. The graphs in the bottom panel display regression discontinuity estimates of the impacts of this reform on the outcome displayed in the top panel. These estimates are based on least-squares estimation of equation (1). Estimates are displayed for every bandwidth from 12 months to 84 months. In addition to these estimates (solid lines), the graphs in the bottom panel display 95% confidence intervals based on robust standard errors (dotted lines) and Kolesár and Rothe (2018) "honest" confidence intervals (dashed lines).

Figure F2 Impacts of compulsory schooling reform on labor market activity



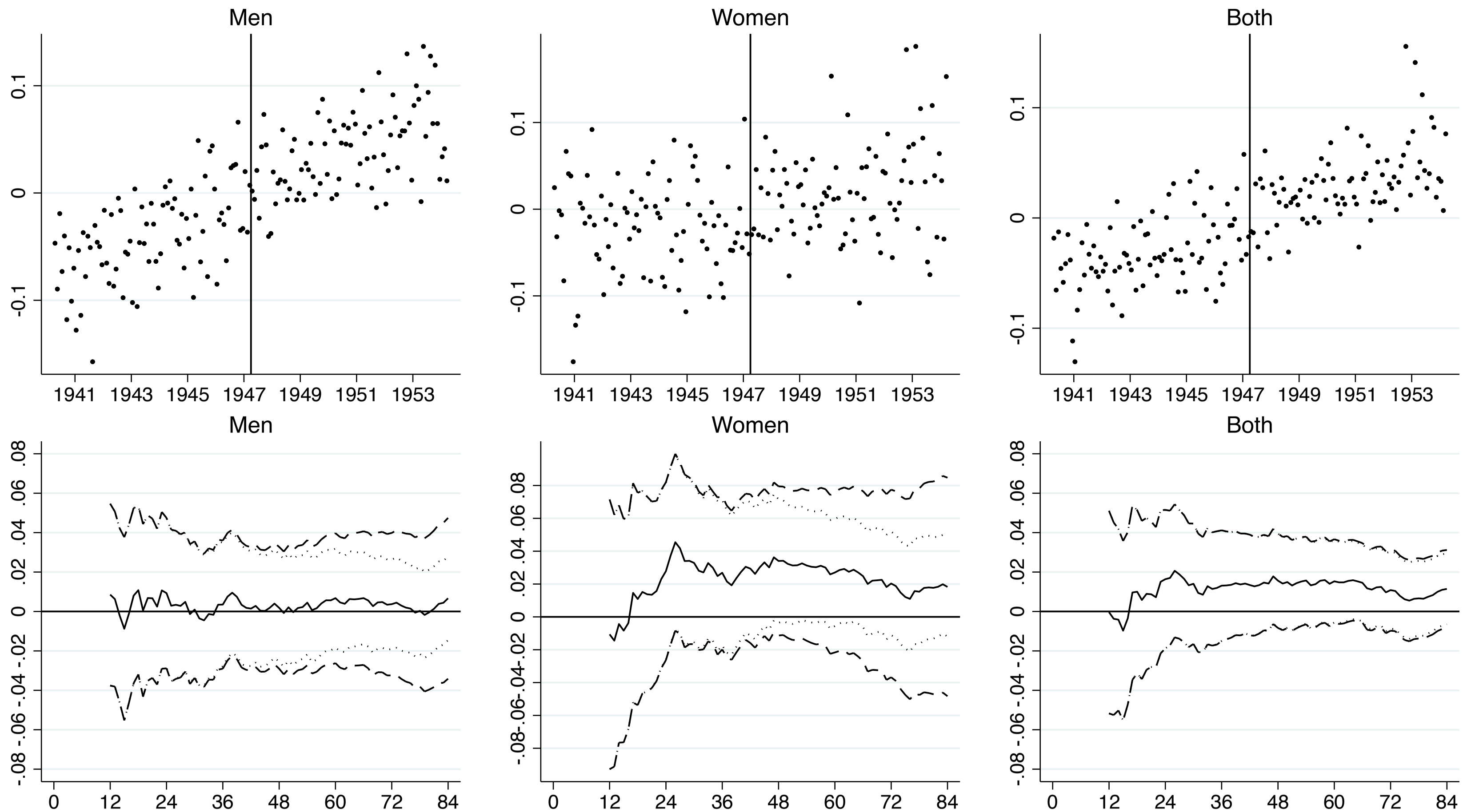
Notes: The graphs display regression discontinuity estimates of the impacts of the compulsory schooling reform on three dimensions of labor market activity (i.e., the estimates corresponding to the outcomes graphs in Figure 2). These estimates are based on least-squares estimation of equation (1). Estimates are displayed for every bandwidth from 3 months to 24 months. In addition to these estimates (solid lines), the graphs display 95% confidence intervals based on robust standard errors (dotted lines) and Kolesár and Rothe (2018) "honest" confidence intervals (dashed lines).

Figure F3 Impacts of compulsory schooling reform on having positive weekly earnings



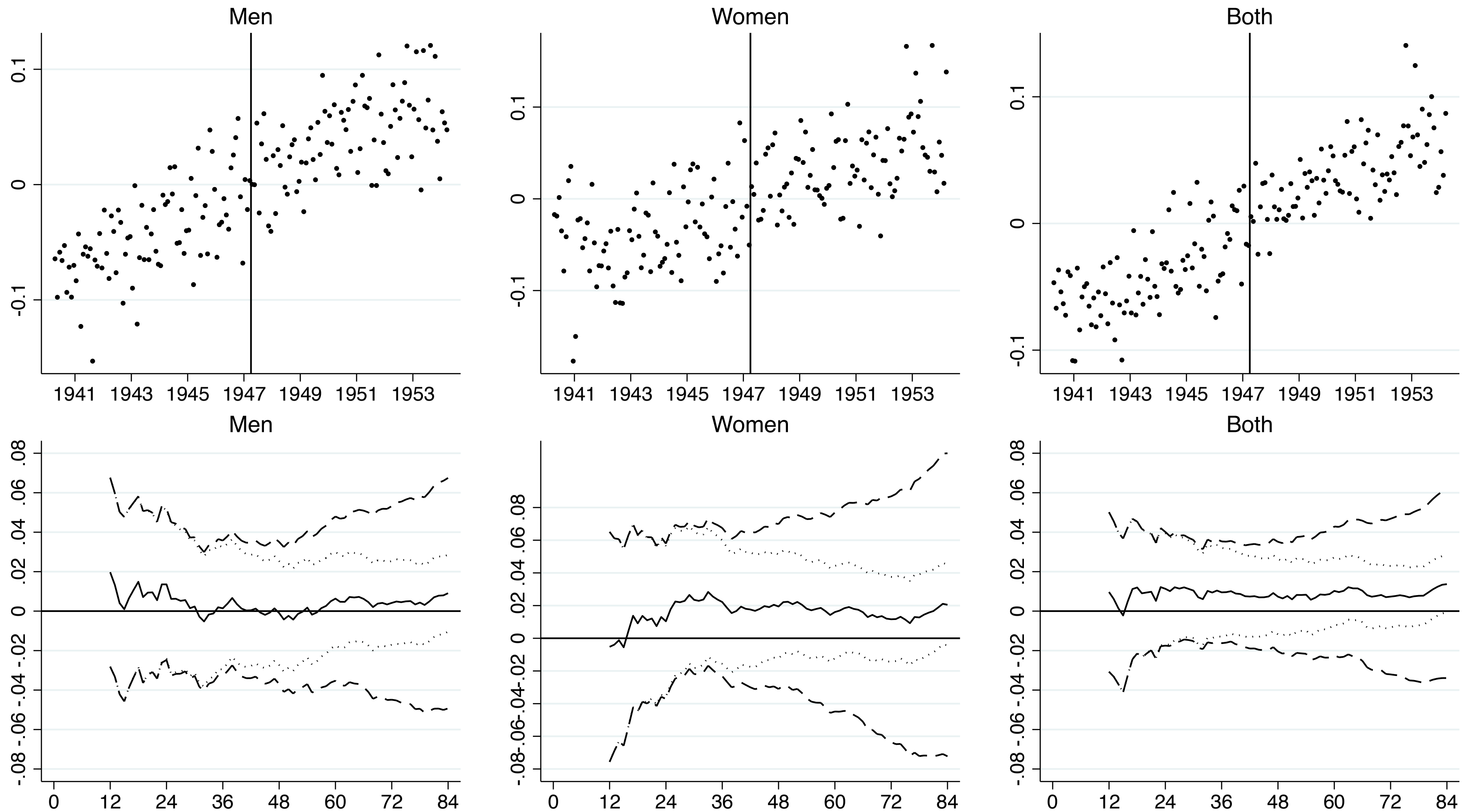
Notes: The graphs are analogous to those in Figure 3, except that the outcome is the cohort-specific fraction of individuals with positive weekly earnings.

Figure F4 Impacts of compulsory schooling reform on log annual earnings



Notes: The graphs are analogous to those in Figure 3, except that the outcome is log annual earnings.

Figure F5 Impacts of compulsory schooling reform on log weekly earnings (among full-time workers)



Notes: The graphs are analogous to those in Figure 3, except that the sample is restricted to full-time workers (defined as those working more than 40 weeks in a year).

Figure F6 Impacts of compulsory schooling reform on home ownership

Notes: The graphs in the top panel display the cohort-specific fraction of individuals that own their own home as reported on the 1991 Census. The sample includes only those born in England and Wales. Cohorts are indexed by the month in which they turned 14. The vertical line corresponds to 1 April 1947, when the compulsory schooling reform took effect. The graphs in the bottom panel display regression discontinuity estimates of the impacts of this reform on the outcome displayed in the top panel. These estimates are based on least-squares estimation of equation (1). Estimates are displayed for every bandwidth from 12 months to 84 months. In addition to these estimates (solid lines), the graphs in the bottom panel display 95% confidence intervals based on robust standard errors (dotted lines) and Kolesár and Rothe (2018) "honest" confidence intervals (dashed lines).

Table F1 Impacts of compulsory schooling reform on log annual earnings

	Bandwidth: 2 years			Bandwidth: 4 years				Bandwidth: 6 years			
	(1)	(2)	(3)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Panel A: Male											
Reform	0.0093	0.0079	0.0012	0.0019	0.0011	-0.0017	-0.0016	0.0037	0.0046	0.0037	0.0042
	(0.019)	(0.019)	(0.020)	(0.014)	(0.014)	(0.013)	(0.017)	(0.012)	(0.011)	(0.012)	(0.013)
Observations	48	48	48	96	96	96	84	144	144	144	132
Panel B: Female											
Reform	0.028	0.026	0.028	0.034	0.033	0.036	0.039	0.020	0.021	0.020	0.011
	(0.027)	(0.028)	(0.028)	(0.019)	(0.019)	(0.019)	(0.027)	(0.016)	(0.016)	(0.016)	(0.020)
Observations	48	48	48	96	96	96	84	144	144	144	132
Panel C: Both											
Reform	0.016	0.015	0.011	0.016	0.015	0.014	0.016	0.011	0.011	0.011	0.0073
	(0.018)	(0.018)	(0.019)	(0.012)	(0.012)	(0.012)	(0.015)	(0.010)	(0.0100)	(0.010)	(0.012)
Observations	48	48	48	96	96	96	84	144	144	144	132

Notes: The estimates reported in the table are analogous to those reported in Table 2, except that the outcome here is log annual earnings.

Table F2 Impacts of compulsory schooling reform on log weekly earnings (among full-time workers)

	Bandwidth: 2 years			Bandwidth: 4 years				Bandwidth: 6 years			
	(1)	(2)	(3)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Panel A: Male											
Reform	0.014	0.013	0.0012	-0.00094	-0.00099	-0.0055	-0.0072	0.0042	0.0049	0.0042	0.0058
	(0.019)	(0.019)	(0.019)	(0.013)	(0.013)	(0.013)	(0.016)	(0.011)	(0.011)	(0.011)	(0.012)
Observations	48	48	48	96	96	96	84	144	144	144	132
Panel B: Female											
Reform	0.010	0.0099	0.016	0.019	0.020	0.022	0.015	0.012	0.013	0.012	0.0022
	(0.024)	(0.023)	(0.023)	(0.016)	(0.016)	(0.016)	(0.021)	(0.013)	(0.013)	(0.013)	(0.015)
Observations	48	48	48	96	96	96	84	144	144	144	132
Panel C: Both											
Reform	0.011	0.011	0.0059	0.0073	0.0076	0.0059	0.0028	0.0074	0.0085	0.0074	0.0049
	(0.014)	(0.014)	(0.014)	(0.0097)	(0.0097)	(0.0098)	(0.012)	(0.0079)	(0.0079)	(0.0079)	(0.0088)
Observations	48	48	48	96	96	96	84	144	144	144	132

Notes: The estimates reported in the table are analogous to those reported in Table 2, except that the sample is restricted to full-time workers (defined as those working more than 40 weeks in a year).

References used in Online Appendices

Autor, D. H., Levy, F., and Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly journal of economics*, 118(4):1279–1333.

Board of Education (1937). Homework: Board of Education Educational Pamphlet no. 110. His Majesty's Stationery Office, London, UK.

Buscha, F. and Dickson, M. (2015). The wage returns to education over the life-cycle: Heterogeneity and the role of experience.

Card, D. E. (1999). The causal effect of education on earnings. In Ashenfelter, O. and Card, D. E., editors, *The Handbook of Labor Economics Vol 3a*. Elsevier/North Holland, Amsterdam.

Card, D. E. (2009). Immigration and inequality. *American Economic Review*, 99(2):1–21.

Clark, D. and Royer, H. (2013). The effect of education on adult mortality and health: Evidence from Britain. *American Economic Review*, 103(6):2087 – 2120.

Devereux, P. and Hart, R. A. (2011). Forced to be rich? Returns to compulsory schooling in Britain. *Economic Journal*, 120:1345–1364.

Dolton, P. and Sandi, M. (2017). Returning to returns: Revisiting the British education evidence. *Labour Economics*, 48:87–104.

Gelman, A. and Imbens, G. (2019). Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business & Economic Statistics*, 37(3):447–456.

Harmon, C. and Walker, I. (1995). Estimates of the economic return to schooling for the UK. *American Economic Review*, 85(5):1278–1296.

Heckman, J. J., Lochner, L., and Todd, P. E. (2008). Earnings functions and rates of return. *Journal of Human Capital*, 2(1):1–31.

HMSO (1947). The new secondary education: Ministry of Education Pamphlet no. 9. London, UK.

HMSO (1948). Education in 1947: Report of the Ministry of Education to Parliament. London, UK.

Kolesár, M. and Rothe, C. (2018). Inference in regression discontinuity designs with a discrete running variable. *American Economic Review*, 108(8):2277–2304.

Lambert, P. S. and Prandy, K. (2018). Camsis project webpages: Cambridge social interaction and stratification scales. Retreived 11 January 2023 from http://www.camsis.stir.ac.uk/occunits/distribution.html.

Oreopoulos, P. (2006). Estimating average and local average treatment effects of education when compulsory school laws really matter. *American Economic Review*, 96(1):152–175.

Oreopoulos, P. (2008). Estimating average and local average treatment effects of education when compulsory schooling laws really matter: Corrigendum. *American Economic Review*. https://assets.aeaweb.org/asset-server/articles-attachments/aer/contents/corrigenda/corr_aer.96.1.152.pdf.