

The Impact of State-Provided Education: Evidence from the 1870 Education Act*

Benjamin Milner[†]

This version: August 2021

Abstract

How does access to public education affect occupational outcomes and intergenerational mobility? The UK's 1870 Education Act, which introduced a public education system in England and Wales, provides a unique historical context in which to explore these questions. Using newly digitized historical records and a regression kink design, I find that public school access improved a child's chance of obtaining an occupation requiring literacy in adulthood by as much as 17 pp. I use a triple difference specification to show that the effect extended to children further removed from the kink, and that the quality of occupational outcomes increased with each additional year of schooling. To study the reform's effect on intergenerational mobility, I link father-son pairs across time, matching nearly 4 million individuals using full-count historical censuses. I find that by targeting the lower classes, public school introduction significantly improved intergenerational mobility, with the adult outcome gap between high- and low- class children decreasing by over 10%.

*I am indebted to my supervisor, Mauricio Drelichman, as well as to the other members of my committee - Felipe Valencia Caicedo, Marit Rehavi, and Nicole Fortin - for their guidance and encouragement. I am also thankful for the valuable feedback I received from many others, but especially Kevin Milligan, Hans-Joachim Voth, David Mitch, Angela Redish, Thomas Lemieux, Siwan Anderson, Matt Lowe, Rogerio Santarrosa, Nathan Canen, and Bert Kramer. This paper also benefited from comments at the EHA and CNEH conferences and at Vancouver School of Economics seminars, from financial assistance from the Centre for Innovative Data in Economics (CIDE), and from access to historical census records provided by the UK Data Archive and Findmypast.

[†]Affiliation: Assistant Professor, Department of Economics, University of Alberta. Contact: bmilner@ualberta.ca

1 Introduction

During the 20th century large improvements in education provision coincided with rapid and inclusive economic growth. Goldin & Katz (2009) called this era the “Human Capital Century”. Today education accounts for over 10% of global public expenditures, totalling nearly \$4 trillion annually (UNESCO 2019), and many point to it as the key to upward social mobility. The role of the state, however, continues to be a source of debate in both political and academic spheres. The rarity of long-term longitudinal data, as well as a dearth of counterfactuals, has made it difficult to causally identify the effects of public school, particularly on social mobility.

Exploiting the quasi-random nature by which public schools were introduced in England and Wales following the 1870 Education Act, I am able to address these issues.¹ The 1870 Act offers several advantages relative to modern settings: the selection of treated areas followed a well-defined rule, yielding clean causal identification; the previous system consisted exclusively of private schools; and the reform effectively targeted the working classes. The historical setting also enables me to track the treated from childhood well into adulthood. Armed with freshly digitized historical data on the rollout of public schools, as well as full-count individual-level census records, I use a regression kink design to show that access to public schooling increased a child’s probability of obtaining an occupation requiring literacy in adulthood by as much as 17 pp. I then use a triple difference specification to show that the effect extended to children further removed from the kink. Finally, to examine the effect of public schools on intergenerational mobility, I link nearly 4 million individuals across the full-count 1861, 1881, 1901 censuses. I find that by targeting children of the lower classes, public schools increased intergenerational mobility by narrowing the gap in occupational attainment between lower- and higher-class children in treated areas by over 10%.

Milton Friedman (1962) famously argued that public schools were unnecessary and that the state should instead support choice between private providers of education. Debate concerning the state’s role has continued since, producing a massive body of academic work. Yet, throughout this literature, the context studied is typically one of public school dominance (Hanushek 2002). Friedman’s counterfactual of little to no direct state involvement remains unexamined. Prior to 1870, all primary schooling in England and Wales was privately provided, funded by charity, fees, and government subsidies. The Education Act introduced a public system for the first time, with schools funded directly by local taxpayers and run

¹Throughout this paper I refer to state-provided schools as “public” and all others as “private”. Other works use these terms differently. In the literature, “public” occasionally refers to any school that received government subsidies, which included many church and charity schools as well as state schools. In other contexts, “public school” refers to the exclusive private schools attended primarily by the upper class.

by elected school boards accountable to the central government. It thus provides the ideal context to cleanly identify the effect of publicly provided education.

The Education Act also provides a unique opportunity to study the intergenerational mobility effects of targeted education reform. Recent works by Chetty et al. (2014) and Chetty & Hendren (2018) show that intergenerational income mobility varies greatly across the US, and that highly mobile areas also tend to have better schools. However, both works stress that the relationship they document is purely descriptive. Card et al. (2018) document a negative relationship between school quality and the intergenerational persistence of years of schooling using 1940 US census data, but do not examine adult outcomes. The few existing works that extend the analysis to adult outcomes suggest that improvements to school provision and quality actually tend to decrease intergenerational mobility. Parman (2011) shows that public school expansion in Iowa at the turn of the 20th century worsened intergenerational income mobility. Similarly, Grawe (2010) uses US census data from 1940-2000 to show that state-cohorts with low pupil-to-teacher ratios experienced less intergenerational mobility. Rauscher (2016) finds that 19th century compulsory school attendance laws in US states initially reduced intergenerational mobility, with it eventually returning to pre-reform levels. The reason typically given for this surprising result is that affluent children are better able to take advantage of improvements in public education, suggesting that improvements targeted at the disadvantaged might have better results. I show that the 1870 reform, through precisely such targeting, benefited almost exclusively working class children, improving intergenerational mobility as a result. These results justify the hope often put in public education, and have important implications for policy today.

The remainder of the paper is structured as follows. Section 2 frames this paper’s contribution by reviewing the related literature. Section 3 describes the 1870 Education Act and its historical context in detail. Section 4 describes the data and its collection. In Section 5 I describe the regression kink framework and results. In Section 6 I turn to the triple difference specification and results. Finally, in Section 7 I outline the census linking procedure, and use the linkage to demonstrate the impact public school introduction had on intergenerational mobility. I summarize the findings in Section 8 and discuss future avenues of research.

2 Literature Review

In addition to contributing to discussions comparing public and private schooling and the impact of public schools on social mobility, my work relates to several other strands of literature. In showing how education affected the job prospects of those who received it, it touches on the well-trodden topic of individual returns to education. While it is safe to say few casual relationships have been more thoroughly established than the modern link

between education and income, if and how the relationship has changed across time is less clear. Clark (2005) suggests that the return to skill in England varied widely from 1200-2000. Long (2006) finds a positive correlation between education and occupation outcomes in 19th century England. I prove that this correlation was in fact causal: better education improved adult occupation outcomes in the late 19th century.

I also add to the literature examining the role of education during what is known as the Second Industrial Revolution (1870-1914), a period characterized by rapid innovation propelled by the spread of people and ideas across newly formed railroad and communication networks (Mokyr 1998). While it is generally accepted that education played at most a small role in the First Industrial Revolution (Mitch, 1999), recent work has suggested that this changed during the Second.² Maloney & Valencia (2017) show the important role engineers played in this time, suggesting the importance of the upper tail of the education distribution. Squicciarini (2017) and Becker et al. (2011), in the contexts of France and Prussia respectively, find strong relationships between education and industrization at the regional level during this period. I build upon this literature by examining the effects of education during this time at the individual level, and measuring its impact on intergenerational mobility.

Beyond what it can teach us about the effect of public schooling in general, the impact of the 1870 Act is of great historical interest in and of itself. While some historians view it as a crucial turning point that resulted in a massive expansion of English school supply (Hamilton 1883; Middleton 1970; Armytage 1970), others, most famously West (1970), have suggested that school supply would have grown at a similar rate due to private investment in the absence of public intervention. This debate has until now remained unresolved in large part due to the lack of disaggregated school supply data. Addressing this, my work is the first to gather and analyse school data at the administrative level at which the Act was implemented - the parish. The results show that not only did the reform have a positive effect on school supply and attendance, but that this effect spilled over into adult outcomes.

More broadly, my work studies the effects of major education reform. Close counterparts in this regard are Duflo (2001, 2004), which also examine the impact of a large positive shock to school supply and observe positive labour market effects. Goldin & Katz (2008) is another example, showing that compulsory schooling and child labour laws played a small but significant role in the arrival of mass secondary schooling in early 20th century US. Montalbo (2019) and Blanc & Wacziarg (2020) both examine the effects of another historical education reform: France's 1833 Guizot Law. However, without the benefit of individual level data, Montalbo is necessarily silent on the effect on intergenerational mobility, while Blanc &

²An important caveat to this is that the very upper tail of the education distribution played a significant role in the First Industrial Revolution (Squicciarini & Voigtländer 2015).

Wacziarg’s detailed examination of a single village enables fascinating insights, but limits statistical power and external validity. By linking fathers and sons across time using the full count national census, I am able to push further than previous works and analyse how social class interacted with access to schooling.

I also add to the growing list of studies utilizing regression kink analysis. While its use has been growing in public finance, labour, and health, my paper is among the first to apply the technique in the economic history literature. It is also one of the first papers, besides Dong (2018) and Sohn & Lee (2019), to utilize a kink in the probability of receiving treatment, as opposed to in its magnitude.

Long (2006) is the contribution closest to my paper. It also uses late 19th century English census data to measure the impact of education on occupation outcomes. Using a nationally representative sample of 5,337 men linked across the 1851 and 1881 censuses, Long finds that 1851 school attendance is related to higher 1881 occupation status, even after controlling for father’s occupation and other family characteristics. My results support this, showing that those treated with improved primary education are more likely to have higher 1901 occupation status. However, several important differences exist between our works. I am able to address the endogeneity of the education choice, using an external policy shock (the 1870 Act) that varies by location. I also show how policy can affect intergenerational mobility. Finally, the linkage of nearly 4 million men across census years enables me to identify the policy’s mobility effects with precision, while still accounting for age and parish level differences.

3 The Education Act of 1870

William Forster – the 1870 Education Act’s architect and leading advocate in parliament – stated that the Act’s main purpose was to extend and improve “the elementary education chiefly of the working classes,” by increasing supply, attendance, and quality (McCann 1970, pg. 134). According to Middleton (1970, pg. 166), “it is not too extravagant to claim that [the 1870 Education Act] introduced a new type of society which radically altered the child’s place in the community.” This is not an overstatement. In the ten years following the its passage, the Act caused over 2,000 parishes to form school boards.^{3,4} These in turn created thousands of public schools, known as “board schools.” This represented the first public sector foray into school administration in English history, and was massive in scale. By 1900 nearly 5,700 board schools existed, attended by nearly 1,900,000 children (Lawson & Silver

³The roll out of boards over this period can be observed in Appendix Figure B.1.

⁴The parish, also known as the civil parish, represented the smallest unit of local government in the UK at the time. Their borders were originally based on those of the Church of England’s ecclesiastic parishes, but had diverged prior to 1870.

2013). While the creation of the modern English education system certainly did not end with the 1870 Act,⁵ it was, in the words of Morris (1972, pg.23), “in many ways the 1066 of English education ... both a climax and a new beginning.”

There were separate, contemporary legislative acts that addressed middle and upper class education: the 1869 Endowed Schools Act for the middle class, and the 1868 Public Schools Act for the upper class. Both were far more limited in scope, however. The Public Schools Act only addressed the nine leading primary schools in the nation, while the Endowed Schools Act attempted to protect endowments at schools catering to the 1/7 of the population labelled as middle class (Middleton 1970). In contrast, it was estimated in the legislation of the 1870 Act that the lower classes it served represented over 85% of the population.

Prior to 1870, pressure for large-scale education reform in England and Wales had been building for decades. Scotland, Germany, and large parts of the US had all provided universal primary education since the early 19th century. By 1870, a large and diverse coalition of liberals, non-conforming religious believers, industrialists, and trade unions were pushing for greater state involvement in education. Their motivations varied. In 1867, the franchise had been extended to nearly all working class men, and some thought it important to educate these new voters (Lawson & Silver 2013). Non-conforming religious believers disliked the Church of England’s central role in school provision (Richards 1970). Finally, both industrialists and trade unionists viewed education as key to improving the productivity of workers (McCann 1970). English industry was beginning to fall behind its American and German rivals, and many blamed the education, or lack thereof, of the English workforce. Forster himself stated that “upon the speedy provision of elementary education depends our industrial prosperity” (Middleton 1970, pg. 167).

Prior to the Act’s passage, schooling (where it existed at all) was provided through a mixture of private and church-run schools, which in turn were funded through a combination of charity, fees, and government subsidies.⁶ The Act sought to “fill in the gaps” where these institutions fell short of providing sufficient school space. To that end, school space was judged by government inspectors, and those parishes deemed to have insufficient space for children ages 5-12 were forced to elect school boards. These boards, under supervision from London, were to construct and run board schools, funded through a mixture of local land taxes and government subsidies and loans.

Gaps in school supply were very common prior to the reform. Appendix Figure 1a shows the geographic distribution of parish-level school supply across England and Wales at the

⁵Mitch (2019) suggests that the 1902 Education Act, which further centralized education authority and funded secondary schools, was perhaps even more consequential.

⁶For more detailed descriptions of education providers prior to 1870 see Gordon (1974), Mitch (1992).

time. The government judged parish primary school supply to be sufficient if spots were provided for at least 1/6 of the total population.⁷ This 1/6 rule, widely accepted throughout the second half of the 19th century, was based on the calculation that approximately 6/7 of the population belonged to the working class for whom parish schools were provided, and a little over 1/6 of the population was of school age (5-12).⁸ Examination of the 1871 survey of school space suggests only 30% of parishes, representing less than a quarter of the total population, met this requirement at the time.

The 1/6 rule was not hard and fast. Indeed, a comparison of Figures 1a and 1b – which chart pre-treatment school supply as well as where school boards were eventually formed – reveals that some parishes above the 1/6 threshold received boards, while many below did not. The most important reason for this was that parishes were given a grace period, during which they could attempt to fill any school insufficiency privately. Before a parish was compelled to form a school board, several steps had to be taken. First, the central Board of Education needed to determine that a parish lacked sufficient school space and provide notice of the deficiency to the parish.⁹ This involved sending an inspector to the parish in person, and appears to have been a time consuming process given that some parishes did not receive notice until 1874.¹⁰ After notice was given, local school proprietors and land owners were allowed one month to dispute the deficiency and demand a public inquiry.¹¹ If no inquiry was needed, or after it was held, a final notice of the deficiency was to be provided to the parish. If a deficiency still existed six months after the posting of this final notice and was not in the process of being alleviated through the construction of a school, the parish was to be compelled to form a school board.¹²

Both in funding and governance, school boards represented a significant break from the past. While private schools were largely funded by charity, boards were given the ability to tax local residents. In terms of governance, management of private schools was typically the domain of the Church and wealthy benefactors;¹³ in contrast, all local taxpayers were able to vote in school board elections.¹⁴ And while the central government had little authority

⁷“In ordinary cases . . . it may be assumed that accommodation in elementary school will be required for one sixth of the entire population. This has been the rule hitherto followed . . . and it is probably sufficient for all practical purposes.” The Board of Education’s instructions to inspectors. Parliamentary Papers, 1883, Vol. 53, paper C.3602.

⁸Parliamentary Papers, 1871, Vol. 22, paper C.406

⁹“The Elementary Education Act, 1870,” section 8, UK parliamentary paper, 1871, Vol. 22, paper C.406

¹⁰UK parliamentary paper, 1874, Vol. 18, paper C.1019

¹¹“The Elementary Education Act, 1870,” section 9

¹²Ibid, Section 10.

¹³Private schools eligible for government subsidy were expected to maintain certain standards and were known as “voluntary” schools. They still enjoyed significant autonomy, however, with the Board of Education unable to intervene in school governance even in clear cases of mismanagement (Gordon, 1974, pg. 8).

¹⁴“The Elementary Education Act, 1870,” schedule 2, part 1

over private school trustees, it could recall and replace any school board or board member deemed in violation of its mandate.¹⁵

Landowners and the established church were united in opposing “the dread intrusion of a School Board” (Thompson 1963, pg. 208). The opposition of the landowners is unsurprising, since school boards were partially funded by local land taxes. Further, landowners typically did not belong to the class that sent their children to board schools, and, unlike industrialists, they stood to gain little from a more educated workforce (Galor & Moav 2006; Galor et al. 2009). Goni (2018) empirically verifies the opposition of landowners to the Act, documenting a negative relationship between school board funding and landowner power across parishes. At the same time, the Church of England viewed the Act as a “source of great danger,” both to its own influence and to the moral fabric of the nation (Platten 1975).

This antipathy meant that, to avoid school boards, small school deficiencies were often filled privately by local elites during the grace period. However, the greater a parish’s pre-reform school supply deficit was, the more likely it was to form a board, presumably due to the increasing cost the local elite faced to fill it privately. Parishes with sufficient pre-reform school supply were allowed to voluntarily form school boards, but this option was rarely exercised. The result is that a kink in treatment probability existed at the 1/6 cutoff, as can be seen in Figure 2, which compares the empirical frequency of eventually receiving a school board with pre-reform school supply. Section 5 describes how I exploit this kink to identify the treatment effect of the reform.

Table 1 compares the summary statistics of parishes that received school boards with those that did not. Unsurprisingly, school supply in 1871 was lower among parishes that eventually received school boards. These parishes were on average more urban, with larger populations and a higher proportion of non-agricultural workers. While school boards clearly were not randomly distributed, for the identification strategies used here they do not have to be. Both the regression kink design and the triple difference-in-differences solve the selection problem based on minimal and flexible assumptions, as described in Sections 5 and 6.

Prior to the passing of the Act, Forster noted that

“there are vast numbers of children utterly untaught, or very badly taught, because there are too few schools, because many schools are bad schools, and because many parents either cannot, or will not, send their children to school.”¹⁶

As already established, school boards addressed the issue of “too few schools”. They also sought to address the problems of “bad schools” and parents who “cannot, or will not, send

¹⁵Ibid., section 23

¹⁶Middleton (1970), pg. 170

their children to school.” While board schools did charge fees to those able to pay, these were typically lower than those charged by private schools (Platten 1975), and were remitted for families deemed unable to pay.¹⁷ Additionally, board officers were given the right to fine parents of absentee children aged 10 and under, and the discretion to pass bylaws requiring attendance up to age 13. Outside the boards, schools were unable to mandate attendance until after 1880, and even then only for children up to age 10. Further, there is anecdotal evidence of higher teacher quality in board schools relative to private schools. An 1861 parliamentary report on education noted that “The teachers of them [for-profit schools] have often no special fitness, ... but have taken up the occupation in default or after the failure of other trades.”¹⁸ Others (Gardner 1984; Hurt 1971) have suggested that the operators of these schools were often themselves illiterate. In contrast, teachers at board schools were generally drawn from teacher training colleges (West 1970). Further, classes were examined annually by government inspectors to ensure standards of efficiency were met.¹⁹

Through multiple channels, including increased supply, lowered fees, better enforced attendance, and better quality of instruction, school boards brought improvements in education to the working classes. This motivates their use throughout this work as a source of a positive shock to education provision. Nonetheless, in Appendix C I verify this ‘stage zero’ effect, proving that school boards dramatically increased school supply and attendance.

4 Data

This project draws upon several primary sources never before used in empirical work. Pre-treatment school supply and population figures are gathered from a national survey administered immediately following the passage of the reform. Local parish officials were instructed to fill out returns requesting total local school capacity, defined as a tenth of school square footage. The results were published in an 1871 report to parliament.²⁰ Using optical character recognition software I digitized the records of all 14,094 parishes surveyed.²¹

This represents the universe of parishes laying outside of municipal boroughs. Parishes within municipal boroughs were not administered the survey and thus are not included in the data.²² Thus, parishes in the cores of many major cities, including London, are excluded from the analysis. Nonetheless, it would be wrong to view the remaining parishes as exclusively rural. Many large centres, such as West Bromwich and Torquay, did not have

¹⁷“The Elementary Education Act, 1870,” section 17

¹⁸UK parliamentary paper, 1861, Vol. 21, paper 2794, pg. 38.

¹⁹“The Elementary Education Act”, section 71.

²⁰UK parliamentary paper, 1871, Vol. 55, 201. A sample of the report, showing 54 parish records spread across two pages, is shown in Appendix Figure E.1.

²¹Characters classified as uncertain in the OCR output were verified by hand.

²²There were 221 municipal boroughs at the time, and each often contained many parishes.

borough status, while many densely populated suburbs, such as Croydon and West Ham, lay outside of boroughs. Over 24% of those included resided in parishes with a population density greater than 4 people per acre (or 2560 per square mile).²³ Further, the linked sample used in Section 7 includes adults residing in municipal boroughs, so long as data from their childhood parish of residence is available.

Pre-treatment school supply is mapped in Figure 1a. Few parishes at the time met the 1/6th rule for school supply, and many had no school space whatsoever.²⁴

To determine which parishes ended up receiving school boards, I use the Board of Education's 1878-79 Annual Report to Parliament, which records all 2,398 parishes that formed school boards, along with formation date for each.²⁵ I digitized this information by hand, and matched it to the 1871 survey data. As can be seen in Figure 1b, while boards were more common in some areas - in particular the Southwest and Wales - they were formed across the country, with more than one in every English and Welsh county.

To assess the 1870 Act's effect on total school supply, I hand-collected school data from the Board of Education's 1880-81 and 1888-1889 annual reports to Parliament.²⁶ Details on how these sources are linked to the 1871 survey are provided in Appendix C.

Several control variables come from other sources. Local religiosity is controlled for using denominational church attendance from the 1851 Religious Census.²⁷ Parish distance from London is obtained using GIS software, matching parishes to shape files of their historical boundaries.²⁸

Parish data are merged with individual data from the 1861, 1881, and 1901 censuses. This includes information on age, occupation, gender, family, and place of birth. The censuses were made available by the Integrated Census Microdata (ICeM) project, part of the UK Data Archive. Most of this paper's analysis is conducted using the universe of males ages 16-50 in each census. However, for reasons described in Section 7, I narrow my focus to those ages 25-45 for the census linked analysis. Females are not included in the analysis because the practice of changing surname at marriage severely limits the ability to link women across censuses. Additionally, the occupation reported for women at the time was often inaccurate, with many providing their husband's occupation as their own (Higgs 1996, pg. 98).

²³Appendix Table F.1 provides a detailed comparison between the general population and those residing in parishes included in the school data.

²⁴Occasionally a single school served multiple parishes. Appendix A details how parish school supply is determined in these cases, and demonstrates that the results of the paper are not dependent on this.

²⁵UK parliamentary paper, 1878-79, Vol. 23, paper C.2342. Appendix Figure E.2 shows a sample page.

²⁶Appendix Figure E.3 shows a sample page.

²⁷Southall & Ell (2004), "Great Britain Historical Database : Census Data : Religion Statistics, 1851." [data collection]. UK Data Service. SN: 4562, <http://doi.org/10.5255/UKDA-SN-4562-1>.

²⁸"Great Britain Historical GIS Project" (2011), University of Portsmouth.

I wish to identify which parish an individual spent their childhood in, so as to observe whether or not they were treated. However, for Sections 5 and 6, when I am using the complete, unlinked censuses, I only observe with certainty current parish of residence and county of birth. Those no longer residing in their birth county are obviously less likely to have grown up in their current parish of residence, hence they are dropped. In Section 7, however, I keep these movers, as using the linked sample I directly observe childhood parish of residence. While the linked sample is mainly used to address questions of social mobility, in Appendix D I use it to demonstrate that excluding movers in the previous sections did not bias results, and that it was indeed childhood place of residence that mattered for treatment.

Income was not recorded in the censuses of the period. Occupation, however, was, and I use it as a proxy for social status and return to human capital. The censuses report over 7 million unique occupation strings. Following each census, officials in the General Register Offices categorized each string into one of nearly 800 occupation groups that were broadly consistent across the censuses used. Using job adverts published in 19th century English periodicals, as well as other contemporaneous descriptions of occupations, Mitch (1992) estimates each occupation group’s use of literacy, specifying four categories of jobs: “literacy required”; “literacy likely to be useful”; “possible (or ambiguous) use of literacy”; and “unlikely to use literacy”. It is this categorization of occupations that I use as the dependent variable for most of the analysis, although to ease interpretation I group the first two and the last two together to create a binary variable.²⁹

To be clear, I do not rely on this categorization to determine how public education affected the spread of literacy. By many measures literacy was already widespread at the time of the Act; by 1870 around 80% of men were able to write their own name (West 1978). Instead, occupations that required literacy should be viewed as a proxy for those in which the returns to human capital were high. While basic levels of literacy were common, there was still large variation in the degree of literacy: at the time, many could read but were unable to write anything beyond their name. Jobs that required literacy were typically those where productivity increased with literacy ability. While literacy alone may have sufficed, those with greater mastery of it were more sought after (Mitch 1992). Indeed, high levels of literacy at the time was seen as a signal of general intelligence (Mitch 1992).

Occupations requiring literacy were also typically better compensated, and held in higher esteem (Mitch 1992). However, to more precisely proxy these outcomes, I also verify the results using HISCLASS, a break down of 19th occupations by social class which I describe in detail in Section 6.2.3.

²⁹In Appendix Table G.1 I show the 12 most common jobs by literacy requirement.

5 Regression Kink Identification

One of the goals of the 1870 Education Act was to “fill in the gaps”: to target parishes with insufficient school supply. As previously described, each parish was expected to provide enough school space to educate $1/6$ of its entire population, with each pupil expected to require ten square feet of school space. Thus, parishes with less than 0.166 school spots – equivalently 1.66 square feet of school space – per capita, were deemed to have insufficient supply, and were to be compelled to form school boards to address the shortfall. However, they were given a grace period to address the shortfall privately if they so wished. If this was accomplished, board formation was not compelled.

The outcome of this assignment rule can be seen in Figure 2. The larger the pre-reform supply shortfall, the more likely a parish was to receive a school board. This was likely because larger shortfalls were more costly to fill privately, making expansion of private supply less attractive relative to board formation. Crucially, though, past the cutoff of 0.166 school spots per capita, pre-reform school supply had little bearing on whether or not a parish received a school board. These sufficiently-supplied parishes were still allowed to form school boards voluntarily, but did so rarely. The result is a kink in the relationship between treatment probability and pre-reform school supply at the cutoff. Supporting the notion that this kink is caused by the Act’s arbitrary rule, and not instead by some kink in unobservables, Appendix Figure K.4 finds no kinks in the distributions of those variables shown in Table 1 to be correlated with board formation: population, religious non-conformism, farm employment, and secondary sector employment.

As discussed in Section 3, school boards were associated with improvements in school supply, attendance, and quality. If boards did indeed improve human capital among the treated, then the kink in treatment probability at the cutoff should induce a second stage kink in human capital at the cutoff as well. I use the 1901 proportion of jobs that make use of literacy within a parish among 19-30 year olds as a proxy for human capital accumulation. The 19-30 age range is the largest range that ensures all individuals included from treated parishes received at least four out of a possible eight years of treatment, while in the comparative age range in 1881 no individual received four years or more.³⁰ Figure 3 demonstrates that there does indeed appear to be a kink in this variable present at the cutoff. I exploit these first and second stage kinks in a Fuzzy Regression Kink (FRK) design to identify the effect of public school introduction.

³⁰A detailed breakdown of years of treatment by age, year, and date of board formation is given in Table H.1.

5.1 Regression Kink Assumptions

Identification using a FRK design rests on four assumptions, described formally in Appendix I. The first is simply that a kink in treatment probability exists. The existence of such a kink at the cutoff of 0.166 pre-reform school spots per capita is strongly supported both by the institutional setting and by the data displayed in Figure 2. A second assumption requires no other kinks in treatment probability near the cutoff, and no jump at the cutoff. Again, Figure 2 supports this.

The final assumptions require no kink to exist in the dependent variable near the cutoff given treatment status. Let Y represent the proportion of jobs within a parish that make use of literacy, and Y_1 and Y_0 its potential values given treatment or not, respectively. The third assumption implies that the relationship between Y_0 or Y_1 and pre-treatment school supply itself is not kinked at the cutoff, while the fourth assumption is that there is no kink in the distribution of unobservables at the cutoff. These final assumptions may be viewed as analogous to an exclusion restriction, as they imply that the instrument (the kink) is only correlated with Y through its association with the endogenous variable (treatment status).

The threat of a kink in the relationship between Y_0 or Y_1 and pre-treatment school supply itself at the cutoff seems unlikely: school space and human capital accumulation are likely related, but there is no reason to expect a kink, let alone one specifically at 1.66 square feet per capita. The same applies for unobservables - with the exception of sorting (which I address below), there is no economic or historical reason why there should be a kink in Y at the cutoff except because of the 1870 Act; as already discussed, no kink appears to exist in other pre-treatment covariates. Nevertheless, I test this assumption by examining a similar, but untreated, sample, as suggested by Landais (2015). In Figure 4 I examine the relationship between Y and pre-treatment school supply among the same age group (19-30 year olds) as Figure 3 but from twenty years prior (1881 as opposed to 1901). These men were too old to be treated, thus a kink at the cutoff in this sample would suggest an assumption violation. However, no kink is observed. While this result does not rule out a violation of assumptions, it does imply that such a violation likely did not exist in 1881.

One potential concern is that sorting at the cutoff could cause a kink in the distribution of unobservables. In Figure 5, I bin the parishes according to pre-treatment school supply R and plot the number of parishes in each bin. The density of R , $f_R(r)$, appears smooth around the cutoff $r_0 = 0.166$, suggesting sorting is not an issue. Nonetheless, in Appendix J I formally assess this smoothness using two tests. The first, the McCrary test, fails to reject its null that $f_R(r)$ is continuous at r_0 . In a normal RD, this would be sufficient, as continuity of the distribution of unobservables is all that is assumed. However, in RK, differentiability is also required. Thus, I conduct a second test, suggested by Card et al. (2012), the null

of which is that $f_R(r)$ is differentiable at r_0 . This test also fails to reject its null, providing further evidence against sorting.

Thus, all evidence suggests that the kink in Y at the cutoff is indeed caused by the 1870 Act. One final concern, however, is that the kink may not be caused solely by the Act's creation of public schools, but also by private school expansion resulting from the Act. To test this alternative mechanism, data on total post-treatment parish school supply were collected from the Board of Education's 1880 Report to Parliament. Appendix Figure K.5 bins parishes by their pre-reform school supply per capita, then plots each bin's average 1880 school supply per capita. No kink in post-treatment school supply appears near the 1870 cutoff,³¹ suggesting that it is indeed the creation of public schools, and the improved quality and attendance associated with them, that drive the kink observed in Figure 3. Nonetheless, as a robustness check, FRK regressions directly controlling for 1880 supply are included in the analysis below.

5.2 Regression Kink Identification and Results

The available evidence strongly supports the FRK assumptions. This enables identification of a local average treatment effect of public schools. Let $G(r) = E[Y|R = r]$, and treatment $T = 1_{\{P(R)-U \geq 0\}}$, where U is normalized such that $U \sim Unif(0, 1)$, so $P(R)$ is the probability of treatment given pre-treatment school supply. The treatment effect is equal to the size of the second stage kink, divided by the size of the first stage kink:³²

$$\frac{G'_+ - G'_-}{P'_+ - P'_-} = E[Y_1 - Y_0 | U = P(r_0), R = r_0] = \tau \quad (1)$$

See Appendix I for proof.

It is important to make clear what this measures: the average treatment effect at r_0 among those of type $U = P(r_0)$. These are the compliers: the parishes that would not be treated to the right of the kink but would be treated to the left. Thus, τ identifies the effect of public schools on the marginal parish, the one just on the edge of being treated or not.

To identify G'_+ , G'_- , P'_+ , and P'_- , local linear regression is used. On either side of the cutoff, observations within a distance of h are used to estimate parameters for the following equations:

$$T = \alpha_0 + \alpha_1 R + \alpha_2 [1_{[R \geq r_0]} * (R - r_0)] + \epsilon \quad (2)$$

$$Y = \gamma_0 + \gamma_1 R + \gamma_2 [1_{[R \geq r_0]} * (R - r_0)] + v \quad (3)$$

³¹This is perhaps not surprising. School supply growth is evident even among those parishes already above the 0.166 floor in 1870, supporting West's (1970) notion that private forces were actively increasing school supply irrespective of public intervention. Thus, while the floor may have been binding in the short run for many of parishes near the cutoff, by 1880 it likely rarely was. Of course, the proportion that found the floor binding in 1880 presumably increased as 1870 supply decreased, as indeed Appendix Figure K.5 suggests.

³² $F'_+ = \lim_{r \downarrow r_0} \frac{\partial F(r)}{\partial r}$ and $F'_- = \lim_{r \uparrow r_0} \frac{\partial F(r)}{\partial r}$.

Then the treatment effect described above can be estimated as

$$\hat{\tau} = \frac{\hat{G}'_+ - \hat{G}'_-}{\hat{P}'_+ - \hat{P}'_-} = \frac{\hat{\gamma}_2}{\hat{\alpha}_2} \quad (4)$$

Note that Equation 3 is simply the reduced form of the following:

$$Y = \beta_0 + \beta_1 R + \tilde{\tau} T + \mu \quad (5)$$

The treatment effect $\tilde{\tau}$ is estimated using 2SLS on an optimal bandwidth $h = 0.064$.^{33,34} Table 2 shows the results. The first stage is very strong, as was suggested by Figure 2. The second stage estimate suggests that parishes near the kink treated with a school board experience an over 17 pp increase in the proportion of jobs requiring literacy in the treated age range. To put this effect in perspective, in 1901 around 45% of the total population were employed in jobs requiring literacy. Thus, public schools appear to have had large and long lasting effects, dramatically improving the job prospects of treated children.

These estimates are not sensitive to bandwidth selection. Figure 6 plots the second stage estimates and confidence intervals for alternative bandwidths.³⁵ While the confidence interval grows as the bandwidth (and thus sample size) decreases, the size of the coefficient is quite consistent. This also suggests that asymptotic bias, which should decrease as the bandwidth shrinks, is small.

Note that the change in treatment probability at the cutoff is independent of all covariates given the assumptions, and thus their inclusion is not necessary to obtain consistent estimates. However, such inclusion can improve estimates by reducing both sampling variability (Lee & Lemieux, 2010) and asymptotic bias (Frolich & Huber, 2019). The 2SLS regressions are thus rerun, but including those variables shown in Table 1 to be correlated with treatment: population, religious non-conformism, farm employment, and secondary sector employment. The proportion of jobs within a parish in 1881 that make use of literacy among those aged 19-30 – in other words, the dependent variable lagged by 20 years – is also included as a control: as previously noted, this earlier cohort was too old to be treated by the 1870 Act. The top panel of Figure 7 plots for various bandwidths the estimated treatment effect $\tilde{\tau}$ obtained when including these controls.³⁶ Encouragingly, the results are very similar to those obtained without covariates.

To rule out the alternative mechanism that the 1870 Act’s imposition of a floor on total

³³Given the system of Equations 2 and 5 is exactly identified, using 2SLS to estimate $\tilde{\tau}$ yields an estimate identical to $\hat{\tau}$. The instrument is α_2 , measuring the first stage kink. Lemma A.2 of Calonico et al. (2014) (henceforth CCT) demonstrates that, under certain regularity conditions and when $h \rightarrow 0$ as $N \rightarrow \infty$, $(\hat{\tau} - \tau)$ is asymptotically normal, justifying the use of standard 2SLS errors.

³⁴The optimal bandwidth was selected by applying the method of CCT to the baseline model.

³⁵I tested all bandwidths $h = 0.064 + c * (0.005)$, where $c \in \{-4, -3, \dots, 3, 4\}$

³⁶Full regression results for bandwidth $h = 0.064$ when including controls are given in Table K.1.

local school supply is driving results, the regressions are also rerun including 1880 school supply per capita as a control in addition to those covariates discussed above. The results of this exercise should be viewed with caution: given that one of the key mandates of school boards was to increase local school supply, directly controlling for 1880 school supply may underestimate the effect of school boards. Nonetheless, the estimated treatment effect remains large regardless of bandwidth, as shown in the bottom panel of Figure 7. Though the confidence intervals are often larger,³⁷ even in this most stringent (perhaps overcontrolled) specification the treatment effect when $h = 0.064$ suggests a 12.5 pp increase in literate jobs due to school boards.

As a final robustness test, I examine a “donut” regression kink. First proposed in Barreca et al. (2011), this test addresses concerns that observations at or very near the cutoff may violate one of the identifying assumptions. To ensure results are not driven by some strange subsample very near the kink, only observations laying within an outer bandwidth but outside of an inner bandwidth are used in the estimation, with the inner bandwidth creating a “donut hole” of unused observations around the kink. In Figure L.6 I show the results of a series of such estimations, keeping the outer bandwidth constant at 0.064 but varying the size of the inner bandwidth.³⁸ As expected, the confidence interval grows with the size of the inner bandwidth (as sample size shrinks), but the size of the estimated effect is strongly positive, ranging from 0.122 to 0.229.

In sum, all specifications suggest that public schools increased the proportion of people holding jobs requiring literacy by well over 10 pp among treated ages in parishes near the cutoff. This result is robust to a wide range of bandwidths, the inclusion of controls, and to a “donut” design, and demonstrates that attending public schools increased human capital and improved adult outcomes.

6 Triple Difference

While the FRK design allows for considerable endogeneity, it is also restrictive, only identifying the treatment effect on compliers near the kink. This is not an issue when treatment effects are relatively stable across the population, but this is difficult to assess when viewing a FRK in isolation.

The nature of the reform allows for another identification technique: a difference-in-difference-in-differences, or Triple Difference (DDD). Using this dimensions of time, place, and age, I estimate an average treatment effect on the treated (ATT) across the entire

³⁷This is in part due to sample size. Data limitations (see Appendix C for details) mean 1880 school supply is unavailable for a significant subset of parishes. Thus, of the 5979 parishes within bandwidth $h = 0.064$ of the cutoff, the full set of covariates is available for only 4197.

³⁸I tested all inner bandwidths $h = 0 + c * (0.0025)$, where $c \in \{0, 1, 2, 3, 4, 5, 6, 7, 8\}$

sample, which I then compare with the FRK LATE. This method does not rely on a kink at the cutoff, which can thus be directly controlled for. As well, by making use of all the data, it provides much tighter estimates, which are necessary to identify intergenerational mobility effects in Section 7.

6.1 Model and Theory

The three dimensions used in the DDD framework are time, place, and age. The time dimension compares 1881, when the reform was in its infancy and few had been affected by it, with 1901. Assume that no one was treated in 1881 (this assumption will be relaxed shortly). The place dimension compares parishes that received school boards with those that did not. The age dimension compares those of the correct age to have been schooled within parishes with school boards with those not. For now, assume that the treated age range does not vary by parish (this will also be relaxed shortly). At its most basic, the model can be written as:

$$Y_{itpa} = \delta + \gamma_1 t + \gamma_2 S_p + \gamma_3 A_a + \gamma_4 (S_p * t) + \gamma_5 (S_p * A_a) + \gamma_6 (t * A_a) + \gamma_7 (S_p * A_a * t) + u_{itpa} \quad (6)$$

where t is a year dummy (1 if 1901, 0 if 1881), S_p is a school board dummy (1 if parish p received a school board, 0 otherwise), and A_a is an age dummy (1 if a is a treated age, 0 otherwise). For the moment I will remain agnostic concerning what the dependent variable Y represents. Some simple arithmetic reveals that γ_7 is equal to the following:

$$\gamma_7 = [(\bar{Y}_{1BS} - \bar{Y}_{2BS}) - (\bar{Y}_{1NS} - \bar{Y}_{2NS})] - [(\bar{Y}_{1BU} - \bar{Y}_{2BU}) - (\bar{Y}_{1NU} - \bar{Y}_{2NU})] \quad (7)$$

where B represents parishes with boards and N represents parishes without, 1 represents 1881 and 2 represents 1901, S represents those of the correct age to be schooled within boards and U those not. Thus the treatment effect, γ_7 , represents the difference between two DID, one on those of the correct age to be treated and another on those not.

The ATT over the entire population is identified by γ_7 if two assumptions hold. Both are modified versions of the traditional DID assumptions: parallel trends, and no spillovers. The parallel trends assumption in a DDD is much more relaxed than the one found in a DID, as it allows for a difference in trends between treated and untreated parishes in the absence of the reform, as long as the difference in trends was common across the two age groups. To be clear, this does not imply the two age groups needed common trends in the absence of the reform. Put another way, to violate the assumption, not only must treated and untreated parishes have experienced different growth rates in the absence of the reform, but this difference in growth rates between treated and untreated parishes would itself also have to differ by age group. Even without this added flexibility, differences in parish trends can partially be accounted for by adding a vector Z_p of pre-treatment parish-level covariates,

interacted with time. Note in particular that I directly control for any kink in trends caused by pre-treatment school supply at the previously described cutoff, ensuring the DDD results are identified off variation fully independent from that used to identify the FRK results.³⁹

The second assumption is that there are no spillover effects into untreated parishes or ages. Spillovers across parishes would be a concern in large cities, where multiple parishes exist within the same labour market. However, as described in Section 4, parishes inside municipal boroughs - which most large cities were - are excluded. For most remaining parishes it is reasonable to view each as a distinct labour market. As for spillovers across ages, this is unlikely if one assumes the old and young are separate labour markets. For the sake of robustness, in Appendix N I use a simple DID to test for spillover effects among the old and find no evidence of any.

Two issues stand in the way of implementing the model described in Equation 6. First, due to the reform's staggered roll out, treated ages vary by parish. In 1901 a 35-year-old residing in a parish that formed a board in 1870 would have been 4 at that time, enabling them to spend a full 8 years in a board school from age 5-12. However, someone from the same birth cohort but residing in a parish that formed a board in 1879 would have been too old to receive any treatment. And while the correct age range can be inferred for parishes that ended up receiving the reform, it is unclear what age range they should be compared against in parishes that never did.

I address this difficulty by adding more flexibility to the model, replacing the single school board dummy with a set of fixed effects, one for each possible reform arrival year (from 1870 to 1879, inclusive) as well as one for untreated parishes. Additionally, instead of a single age dummy, fixed effects for each year of age are included. As before, these place and age dummies are interacted with each other and with the time dummy. The result is as follows:

$$Y_{itpa} = \mu_{ra} + \lambda_{rt} + \alpha_{at} + \gamma_7(A_{art}) + (Z_p)\Psi + (Z_p * t)\Omega + u_{itpa} \quad (8)$$

where r represents the arrival year group parish p belongs to, μ_{ra} are the age-arrival year fixed effects, λ_{rt} are the arrival year trend effects, α_{at} are the age trend effects, and A_{art} is a dummy equal to 1 if the individual is from the 1901 census of the age to be treated in parish of type r . The old model is nested in this new model, as the new model simply divides the age and board dummies into finer categories.⁴⁰

A second issue is treatment intensity: some receive the full 8 years of instruction in a

³⁹As a robustness check, I also run regressions directly controlling for any kink in trends at the parish-age level (See Appendix M).

⁴⁰Goodman-Bacon (2019) describes how to interpret a similar model - where multiple groups receive treatment at different times, necessitating the inclusion of fixed effects - in a DD setting. The estimated treatment effect can be viewed as a weighted average of all possible estimators that compare one treatment year group with another.

board school, while children already of school age when the board arrives are only partially treated. To further complicate matters, in the base year of 1881 some individuals had already been partially treated. For instance, in 1881 a 16-year-old residing in a parish that formed a school board in 1870 would have been 5 at the time. If we allow a year for the board to build a school, they had the opportunity to attend a board school for 7 years, from ages 6-12. In 1901, a 16-year-old residing in the same parish had the opportunity to attend a board school for all 8 primary years, 5-12. Thus, the latter cohort received only one additional year of treatment. To address this, I use the difference in years of possible board school attendance TY_{tra} as the treatment variable in place of the treatment dummy A_{apt} . I calculate the number of years each year-parish-age cohort could have attended board schools, and subtract from this the number of years the equivalent parish-age cohort in 1881 could have attended. In this way, the treatment variable is always 0 if the year is 1881, and is the difference between 1881 treatment and 1901 treatment by parish-age group if the year is 1901.⁴¹

The final model is:

$$Y_{itpa} = \mu_{ra} + \lambda_{rt} + \alpha_{at} + \gamma_7(TY_{tra}) + (Z_p)\Psi + (Z_p * t)\Omega + u_{itpa} \quad (9)$$

Additional flexibility can be added by including parish fixed effects η_p in the place of the vector of constant parish controls:

$$Y_{itpa} = \mu_{ra} + \lambda_{rt} + \alpha_{at} + \gamma_7(TY_{tra}) + \eta_p + (Z_p * t)\Omega + u_{itpa} \quad (10)$$

6.2 Triple Difference Results

Table 3 presents the main results of the full sample triple difference. I again use the probability of obtaining an occupation that requires literacy as my main dependent variable. Unlike the FRK, however, the unit of observation is the individual instead of the parish, enabling me to control for age and partial treatment. The results suggest that each additional year of access to public school increased the probability of obtaining a literate occupation by around 0.13 pp. Given that primary schooling lasted 8 years (from ages 5 to 12 inclusive), this implies that full treatment improved chances by around 1 pp. The exclusion of parish fixed effects has very little effect on the size of the estimated treatment effect, although it does result in a large increase in standard errors. I also include a number of parish level controls interacted with time to ensure parish level differences in growth rates are not driving results; the estimated treatment effect coefficient is relatively unaffected by their inclusion.⁴²

⁴¹Table H.1 provides a detailed breakdown of how this difference varies by age and year of school board formation.

⁴²Controls included (all interacted with a 1901 dummy): 1871 parish population; 1871 parish population squared; distance from London; distance from London squared; 1871 school supply per capita; 1871 school supply per capita interacted with cutoff (0.166) dummy; 1851 Church of England attendees per capita; 1851 Catholic attendees per capita; 1851 other religious service attendees per capita.

6.2.1 Robustness

A placebo test of the identifying assumption - that the difference in trends between those in treated vs untreated parishes would not have varied systematically across ages in the absence of the reform - can be conducted using 1861 census data. Suppose we mistakenly thought the Education Act was passed 20 years earlier, in 1850, and that all parishes that received boards formed them 20 years prior to their actual formation date. If the identifying assumption is true, we should not see any significant results when running the same regressions as in Table 3 but using 1881 data in the place of 1901 data and 1861 data in the place of 1881 data. Table 4 shows the results; as expected, all estimates are insignificant. Indeed, the estimated treatment effect is very close to zero when parish fixed effects are included.

In Appendix M I conduct additional robustness checks. I rerun the regressions excluding the partially treated, for whom $TY_{tra} \notin \{0, 8\}$, so as to ensure that results are not being distorted by some non-linear relationship between treatment intensity and outcomes. To ensure outliers are not driving results, I rerun the regressions excluding individuals residing in the top and bottom ventile in terms of pre-treatment school supply per capita, then do the same for 1871 population. Finally, I allow the effect of pre-treatment school supply on literate occupations to vary not only by year and side of the $r_0 = 0.166$ cutoff, but by also by age as well. In each of these robustness tests, the estimated treatment effect remains statistically significant and essentially unchanged in size.

6.2.2 Comparison with FRK results

While the baseline DDD treatment effects may seem small, bear in mind that only 45% of the population held literate occupations. Further, this effect is a population average, but the reform was directed at the lower classes. As I will show using linked census data below, the effect was much larger among the poor. Nonetheless, the ATT measured by the triple difference remains smaller than the LATE given by the fuzzy regression kink.

Of course, there is no reason why these two estimates should be equal. They estimate different things: the FRK LATE, the effect of public schools in parishes at the kink just on the edge of receiving treatment; the DDD ATT, the average effect of public schools across all treated individuals. Yet the sources of their difference can help us better understand how the effect of public schools varied across the population.

In Appendix N, I rule several possible sources of the difference: that treatment effects are on average larger near the kink; that it is the result of using different levels of observation (parish vs individual); that the DDD result is biased due to spillovers across ages.

Perhaps the most likely explanation for the difference is that the effects of public schools were on average larger in the parishes on the margin of receiving them. Remember that the FRK estimates the effect on compliers near the kink - the parishes that would not have been

treated were they to the right of the kink, but would have been treated were they on the left. Were the treatment effect on the compliers larger than that on the always takers, the FRK estimate would be larger than the DDD estimate, even over the same sample.

This scenario implies that the parishes with the most to gain from public schools were also the most reluctant to receive them. This is more plausible than it may seem, given that the local elites who influenced the treatment decision were unlikely to benefit from it given that the reform was aimed at the children of the lower classes. Unfortunately, as compliers and always takers cannot be differentiated in the data, this theory is impossible to test.

It is also possible that the DDD is underestimating the effect due to a violation in its parallel trends assumption. If, in the absence of the reform, the young would have shown more gains in literate occupations relative to the old in untreated parishes than in treated ones, the DDD would be biased downward. To test this, I rerun the regressions of Table 3 but using only those parishes that received boards.⁴³ The results are shown in Table 5. The treatment effect (0.225-0.278 pp per additional year of public schooling, or 1.8-2.224 pp from the full 8 years) is larger than the baseline estimates (0.126-0.136 pp per additional year), but the differences between the two estimation strategies are never significant. The results nonetheless suggest that, if anything, the baseline DDD results are an underestimation.

6.2.3 Class Results

I have thus far examined the effect public schooling had on the probability of obtaining a literate job, a good proxy for occupations with a high return to human capital. Literate jobs were also typically better compensated, and viewed in higher esteem (Mitch 1992).

To make the link between public schooling and positive adult outcomes even clearer, I test the impact public schooling had on the probability of obtaining a job viewed as high class. I use HISCLASS, a ranking of occupations by social class developed by Maas & Van Leeuwen (2016) using 19th century occupation data from Canada and Western Europe.⁴⁴ It divides occupations into seven ordered social classes.⁴⁵ For comparability with my literate occupation variable, I create a binary class variable, grouping classes 1-3 and classes 4-7. The resulting class dummy is equal to 1 for high class occupations and 0 otherwise. Just over 45% of the male population in 1901 was employed in a high class occupation, very similar to the proportion employed in literate occupations. The Pearson correlation coefficient between the occupation class and literacy dummies is 0.5772, suggesting that while closely related,

⁴³This is possible due to the staggered roll out of the reform over the course of a decade, which created variation in the treated age range across treated parishes, allowing me to separate age-trend effects from the treatment effect.

⁴⁴Note that HISCLASS and Mitch's literacy ranking are derived from two separate occupation classifications (HISCO (Van Leeuwen et al., 2002) and Armstrong (1972), respectively), and thus any correlation is not the result of shared sources.

⁴⁵Ordered classes are listed in Appendix Table G.2

the two variables are indeed measuring different things.

I run the baseline DDD regressions using the class dummy as the dependent variable. The results, shown in Table 6, suggest that an additional year of access to public schooling increased the odds of obtaining a high class occupation by 0.15-0.3 pp, and thus access for the full 8 years of primary school increased the odds by 1.2-2.4 pp. When parish fixed effects are included, the treatment effects are very similar in size to literate occupation results. Thus, it is clear that the introduction of public school education improved access to jobs that were not only more human capital intensive, but also higher class.

7 Census Linkage and Mobility Estimation

In modern minds, Victorian England and Wales may conjure images of extreme class rigidity. Reading Hardy or Dickens, one is led to believe that ambition among the lower classes was constantly thwarted (*Jude the Obscure*, 1895), and that social mobility was only possible via large inheritances (*Great Expectations*, 1861). But as Long (2013) and Long & Ferrie (2013) have shown, there was a surprising amount of mobility during the period. Further, Long (2013) shows that social mobility increased significantly between his two comparison periods (1851-1881 and 1881-1901). Was this increase in mobility driven by the introduction of public schooling in 1870?

The analysis thus far has shown that public school introduction dramatically improved job prospects on average within treated age-parish-year groups. But there is good reason to believe heterogeneous treatment effects existed within such groups across social classes. The lower, middle, and upper classes all largely attended different schools. The 1870 Education Act was aimed at improving education among the working classes, estimated to represent about 85% of the population.⁴⁶ Even within that 85% there was considerable variation, as this included everyone from farm day labourers to skilled tradesmen. If, in the absence of public schools, those near the top of the distribution were better able to endow their children with human capital than those at the bottom, then the arrival of public schools may have served to level the playing field, improving social mobility. To test this I need to observe both adult outcomes and childhood class, something that can only be done by linking individuals across censuses.

7.1 Linkage Description

I link young people (ages 5-25) in the full 1881 census to their adult selves (ages 25-45) in the full 1901 census, and find 2,357,948 unique matches. For a pre-treatment comparison group, I also link those aged 5-25 in the full 1861 census to their 1881 selves and find 1,522,047 unique matches. To construct these links, I use a modified version of the Abramitzky, Boustan, and

⁴⁶UK parliamentary paper, 1871, Vol. 22, paper C.406.

Eriksson (2012, 2014, 2019a) method, described in detail in Appendix O, using first and last name, birth year, and birth place (county and parish). Summary statistics of the linkages are provided in Appendix P.

I only link ages 25-45 – as opposed to recreating the 16-50 age range used in Sections 5 and 6 – for several reasons. Individuals aged 16-19 in the later census were not born 20 years prior, and thus do not appear in the earlier census. Since I am interested in where an individual attended primary school, which began at age 5, beginning the early census age range at 5 was appropriate. Finally, father’s class is only observed when he lived in the same household in the earlier census, and thus would only be observable for 50 year olds if they resided with their father when they were 30. For this reason I limit the top age to 45.

Linkages using historical full-count censuses are relatively rare. Appendix Table R.1 compares notable linkages in the historical literature with those constructed here. While match rates of 15-30% are common in the literature, I am able to link 37.1% of men age 0-25 in 1861 to 1881, and 42.2% of men age 0-25 in 1881 to 1901. The reason for the high match rate is that, unlike historical US censuses where birthplace was only listed at the state level, the censuses of England and Wales included birth parish. Matching on this much finer level of geography greatly increases the probability that a match will be unique.⁴⁷

To test the frequency of false positives, I make use of middle initials. Given that middle initial was not used in the matching process (as they are only available for a subset of individuals), a middle initial shared by both sides of a match is a strong indicator that the match is correct. In Appendix R I describe the test in detail. The results suggest that the PPVs (Positive Prediction Value, as the ratio of true matches to total matches is known in the machine learning literature) for the 1861-1881 and 1881-1901 linkages are approximately 87.2% and 93.5%, respectively. The literature suggests that these are very good: Bailey et al. (2019) state that common match procedures lead to PPVs between 66% and 90%, while Abramitzky et al. (2019b) are slightly more optimistic, arguing that modern procedures lead to PPVs between 70% and 95%.⁴⁸

Both the 1861-1881 and 1881-1901 baseline linked samples are representative of the total population (See Appendix Tables P.1 and P.2). Column 4 of Appendix Tables P.1 and P.2

⁴⁷Long & Ferrie (2013) also match English and Welsh census data using birth parish, while not achieving as high a match rate. This is likely because (a) they link across a longer period (30 vs 20 years), over which more individuals would have died or emigrated, and (b) they did not have access to the standardized birth parish variable recently constructed by I-CeM researchers, which addresses the issue of parishes with multiple and changing names.

⁴⁸I also construct two other, more conservative linkages of the 1861-1881 and 1881-1901 censuses, described in detail in Appendix S. These reduce the linkage rate, but lead to even fewer false positives. If false positives significantly affect analysis, one would expect results obtained using the more conservative linkages to be closer to the truth than those obtained using the baseline method. Table S.1 demonstrates that, if anything, the mobility effects described below actually get larger when using the more conservative methods.

shows summary statistics of the subset of linked individuals who resided with their father in the earlier census, enabling measurement of childhood class (using father’s occupation).⁴⁹ These individuals are about a year and a half younger than the general population, as one would expect given younger children are more likely to reside with their fathers. Concerning class, however, these individuals seem representative of the overall population.

7.2 Mobility Regressions and Results

I test the effect of the introduction of public schools on intergenerational mobility using the triple difference framework described in Section 6. As before, I estimate the effect an additional year of access to public schools had on the probability of obtaining a literate occupation in adulthood.⁵⁰ However, using the linked sample, I now observe the occupation an individual’s father held 20 years prior, which I use as a proxy for the individual’s childhood social class (using the class dummy described in Section 6.2.3). By interacting childhood social class with the triple difference, I can compare the effects of public schools on high and low class children. The resulting triple difference equation is as follows:

$$Y_{itpac} = \mu_{rac} + \lambda_{rtc} + \alpha_{atc} + \gamma(TY_{tra}) + \zeta(TY_{tra} * FC_i) + \eta_{pc} + (Z_p * t)\Omega + (Z_p * t * FC_i)\Pi + u_{itpac} \quad (11)$$

Observe that, in addition to the treatment years variable TY_{tra} , all lower level terms have been interacted with father’s class (FC_i) as well, so as to avoid biasing the interaction of interest. In the place of Equation 10’s age-arrival year and parish fixed effects (μ_{ra} and η_p), and arrival year and age trend effects (λ_{rt} and α_{at}), Equation 11 has age-arrival year-class and parish-class fixed effects (μ_{rac} and η_{pc}), and arrival year-class and age-class trend effects (λ_{rtc} and α_{atc}). Parish trend controls have also been interacted with FC_i . The effect on lower class children of an additional treatment year is estimated by γ , while ζ is the difference between this effect and that on higher class children.

7.2.1 Baseline Mobility Results

The first two columns of Table 7 report the results of regression on Equation 11 with and without parish trend controls. Children of fathers with lower class occupations saw their probability of obtaining a literate job in adulthood increase by approximately 0.2 pp with each additional year of access to public school. Thus, those treated for a full eight years saw their chance increase by around 1.6 pp. In contrast, children of fathers with higher class

⁴⁹Individual’s are only linked to their fathers in the census if they reside in the same household. This provides an extra impetus for linking to childhood, as co-residence with one’s father was rare in adulthood.

⁵⁰In Appendix T I estimate effects using other measures of occupation quality, including OCCSCORE - a commonly used measure based on the median income of occupation practitioners in the US in 1950 - and a ranking of occupations based on the likelihood of practitioners’ children obtaining prestigious occupations. The results described below hold in each case.

occupations saw zero or perhaps even slightly negative effects from treatment. Again, this fits with the 1870 Education Act’s stated purpose of improving the education of working class children, and with the historical reality that upper and middle class children largely attended private schools unaffected by the reform.

The gains made by lower class children allowed them to partially close the gap between them and high class children. Consider that among untreated children in 1901, those with high class fathers were 20.0 pp more likely than their low class counterparts to obtain literate occupations in adulthood.⁵¹ Each additional year of access to public schools decreased this gap by 0.29-0.31 pp. Thus, having access to public schools from age 5 to 12 reduced the outcome gap by over 10%, to 17.5-17.7 pp.

7.2.2 Comparison of Brothers

The linked sample enables another method of identification: comparing the outcomes of brothers. Many brothers residing in the same household experienced differing levels of treatment based on age. By including brother fixed effects, one can therefore control for genetic and household effects while still identifying the treatment effect. Of course, age must still be controlled for along with its interactions, but all parish, time, and class fixed effects and controls are subsumed by the brother fixed effects. The resulting equation is shown below, with κ_b representing the brother fixed effect:

$$Y_{itpacb} = \mu_{rac} + \alpha_{atc} + \kappa_b + \gamma(TY_{tra}) + \zeta(TY_{tra} * FC_b) + u_{itpacb} \quad (12)$$

Brothers are only identified in the census if they resided in the same household, which typically only occurred in childhood. Thus, this method is only possible using the linked sample, in which childhood household is observable. The results are shown in the third column of Table 7. While the significance of the interaction term disappears - likely due to the much smaller sample size - the effect sizes are even larger than those previously estimated, implying that full treatment would have reduced the occupational outcome gap between high and low class children by over 15%.

7.2.3 Using Finer Categories of Class

I have so far defined class as a binary variable. Of course, in reality Victorian class structure was more complex than this. To better understand how access to public school differentially effected various parts of the social hierarchy, I make use of the original seven ranked classes of the HISCLASS variable. I divide my sample into seven groups, each containing all the children of fathers belonging to one of the ranked classes. For each group I then regress on the probability of obtaining a literate occupation in adulthood using Equation 10. The treatment years coefficients of each regression are reported in Table 8. The smaller sample

⁵¹Controlling of age and parish of birth. For derivation, see Appendix U

sizes of each regression - the result of dividing into seven different classes - lead to larger standard errors. The trend across regressions nonetheless confirms that, in general, the impact of public schools increased the further down the social ladder one began.

8 Conclusion

The creation of England and Wales' public school system in 1870 provides a unique opportunity to observe how instituting a public school system on top of a previously existing private system affected outcomes, and how an education reform specifically aimed at the poor can improve social mobility across generations. I find that the introduction of public schools resulted in a large positive shock to education provision, increasing school supply and attendance. More importantly, children lucky enough to have access to public schooling were more likely to obtain good occupations in adulthood. In the parish just on the margin of receiving treatment, public schools increased the probability of obtaining a human capital-intensive occupation by 17 pp. Each additional year of public school access significantly improved a child's chance of obtaining a high social status occupation in adulthood. Further, by targeting children of the lower classes, public schools decreased the gap in occupational attainment between lower- and higher-class children by over 10%.

Further work remains to be done to disentangle how public school provision improved educational outcomes. Whether public schools improved private schools by providing competition in the education marketplace, for example, is an open question. As well, while this paper focused entirely on males due to data restrictions, the effects of public school provision on females, who were also provided for by the 1870 Education Act, is obviously of interest.

This paper moves the literature forward on multiple fronts. In particular, it contributes to comparisons of public and private schools by demonstrating that a private only system will not always be optimal. And perhaps more significantly from a policy standpoint, the result that public school provision increased social mobility goes against previous findings of the opposite, justifying the hope that education policy properly targeted at disadvantaged children will increase equality of opportunity.

References

- Abramitzky, R., L. Boustan, and K. Eriksson (2019a). To the new world and back again: Return migrants in the age of mass migration. *ILR Review* 72(2), 300–322.
- Abramitzky, R., L. P. Boustan, and K. Eriksson (2012). Europe’s tired, poor, huddled masses: Self-selection and economic outcomes in the age of mass migration. *American Economic Review* 102(5), 1832–56.
- Abramitzky, R., L. P. Boustan, and K. Eriksson (2014). A nation of immigrants: Assimilation and economic outcomes in the age of mass migration. *Journal of Political Economy* 122(3), 467–506.
- Abramitzky, R., L. P. Boustan, K. Eriksson, J. J. Feigenbaum, and S. Pérez (2019b). Automated linking of historical data. Technical report, National Bureau of Economic Research.
- Armstrong, W. A. (1972). The use of information about occupation. *Nineteenth-century society: Essays in the use of quantitative methods for the study of social data*, 191–310.
- Armytage, W. (1970). The 1870 Education Act. *British Journal of Educational Studies* 18(2), 121–133.
- Bailey, M., C. Cole, M. Henderson, and C. Massey (2019). How well do automated methods perform in historical samples? Evidence from new ground truth. Technical report, National Bureau of Economic Research.
- Baker, R. B., J. Blanchette, and K. Eriksson (2018). Long-run impacts of agricultural shocks on educational attainment: Evidence from the boll weevil. Technical report, National Bureau of Economic Research.
- Barreca, A. I., M. Guldi, J. M. Lindo, and G. R. Waddell (2011). Saving babies? Revisiting the effect of very low birth weight classification. *The Quarterly Journal of Economics* 126(4), 2117–2123.
- Becker, S. O., E. Hornung, and L. Woessmann (2011). Education and catch-up in the Industrial Revolution. *American Economic Journal: Macroeconomics* 3(3), 92–126.
- Blanc, G. and R. Wacziarg (2020). Change and persistence in the age of modernization: Saint-Germain-d’Anxure, 1730-1895. *Explorations in Economic History*, 101352.
- Calónico, S., M. D. Cattaneo, and R. Titiunik (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica* 82(6), 2295–2326.
- Card, D., C. Domnisoru, and L. Taylor (2018). The intergenerational transmission of human capital: Evidence from the golden age of upward mobility. Technical report, National Bureau of Economic Research.
- Card, D., D. Lee, Z. Pei, and A. Weber (2012). Nonlinear policy rules and the identification and estimation of causal effects in a generalized regression kink design. Technical report, National Bureau of Economic Research.
- Card, D., D. S. Lee, Z. Pei, and A. Weber (2015). Inference on causal effects in a generalized regression kink design. *Econometrica* 83(6), 2453–2483.
- Chetty, R. and N. Hendren (2018). The effects of neighborhoods on intergenerational mobility II: County-level estimates. *The Quarterly Journal of Economics* 133(3), 1163–1228.
- Chetty, R., N. Hendren, P. Kline, and E. Saez (2014). Where is the land of opportunity? The geography of intergenerational mobility in the United States. *The Quarterly Journal of Economics* 129(4), 1553–1623.

- Clark, G. (2005). The condition of the working class in England, 1209-2004. *Journal of Political Economy* 113(6), 1307–1340.
- Dong, Y. (2018). Jump or kink? Regression probability jump and kink design for treatment effect evaluation. *Working Paper*.
- Dufo, E. (2001). Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment. *American Economic Review* 91(4), 795–813.
- Dufo, E. (2004). The medium run effects of educational expansion: Evidence from a large school construction program in Indonesia. *Journal of Development Economics* 74(1), 163–197.
- Feigenbaum, J. J. (2015). A new old measure of intergenerational mobility: Iowa 1915 to 1940. *Unpublished Manuscript*.
- Friedman, M. (1962). *Capitalism and Freedom*. University of Chicago press.
- Frölich, M. and M. Huber (2019). Including covariates in the regression discontinuity design. *Journal of Business & Economic Statistics* 37(4), 736–748.
- Galor, O. and O. Moav (2006). Das human-kapital: A theory of the demise of the class structure. *The Review of Economic Studies* 73(1), 85–117.
- Galor, O., O. Moav, and D. Vollrath (2009). Inequality in landownership, the emergence of human-capital promoting institutions, and the Great Divergence. *The Review of Economic Studies* 76(1), 143–179.
- Gardner, P. (1984). *The Lost Elementary Schools of Victorian Britain*. Croom Helm.
- Goldin, C. D. and L. F. Katz (2008). Mass secondary schooling and the state: The role of state compulsion in the high school movement. In *Understanding Long-Run Economic Growth: Geography, Institutions, and the Knowledge Economy*, pp. 275–310. University of Chicago Press.
- Goldin, C. D. and L. F. Katz (2009). *The race between education and technology*. Harvard University Press.
- Goni, M. (2018). Landed elites and education provision in England and Wales. Evidence from school boards, 1870–99. *Manuscript, University of Vienna*.
- Goodman-Bacon, A. (2018). Difference-in-differences with variation in treatment timing. Technical report, National Bureau of Economic Research.
- Gordon, P. (1974). *The Victorian School Manager: A Study In the Management of Education 1800-1902*. Woburn Press.
- Grawe, N. D. (2010). Primary and secondary school quality and intergenerational earnings mobility. *Journal of Human Capital* 4(4), 331–364.
- Hamilton, R. (1883). Popular education in England and Wales before and after the Elementary Education Act of 1870. *Journal of the Statistical Society of London* 46(2), 283–349.
- Hanushek, E. A. (2002). Publicly provided education. *Handbook of Public Economics* 4, 2045–2141.
- Higgs, E. (1996). *A Clearer Sense of the Census*. HMSO.
- Hurt, J. S. (1971). Professor West on early nineteenth-century education. *The Economic History Review* 24(4), 624–632.
- Landais, C. (2015). Assessing the welfare effects of unemployment benefits using the regression kink design. *American Economic Journal: Economic Policy* 7(4), 243–78.
- Lawson, J. and H. Silver (2013). *A social history of education in England*. Routledge.

- Lee, D. S. and T. Lemieux (2010). Regression discontinuity designs in economics. *Journal of Economic Literature* 48(2), 281–355.
- Long, J. (2006). The socioeconomic return to primary schooling in Victorian England. *The Journal of Economic History* 66(4), 1026–1053.
- Long, J. (2013). The surprising social mobility of Victorian Britain. *European Review of Economic History* 17(1), 1–23.
- Long, J. and J. Ferrie (2013). Intergenerational occupational mobility in Great Britain and the United States since 1850. *American Economic Review* 103(4), 1109–37.
- Maas, I. and M. V. Leeuwen (2016). *HISCLASS*. IISH Dataverse.
- Maloney, W. and F. Valencia Caicedo (2017). Engineering growth: Innovative capacity and development in the Americas. *CESifo Working Paper Series No. 6339*.
- McCann, W. P. (1970). Trade unionists, artisans and the 1870 education act. *British Journal of Educational Studies* 18(2), 134–150.
- McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics* 142(2), 698–714.
- Middleton, N. (1970). The Education Act of 1870 as the start of the modern concept of the child. *British Journal of Educational Studies* 18(2), 166–179.
- Mitch, D. (1992). *The rise of popular literacy in Victorian England: The influence of private choice and public policy*. University of Pennsylvania Press.
- Mitch, D. (1999). The role of education and skill in the British Industrial Revolution. In J. Mokyr (Ed.), *The British Industrial Revolution: An Economic Perspective, 2nd edition*, pp. 241–279.
- Mitch, D. (2019). The elementary education act of 1870: Landmark or transition? In *School Acts and the Rise of Mass Schooling*, pp. 301–324. Springer.
- Mokyr, J. (1998). The Second Industrial Revolution, 1870-1914. In V. Castronovo (Ed.), *Storia dell'economia Mondiale*, pp. 219–245.
- Montalbo, A. (2019). Education and economic development. The influence of primary schooling on municipalities in nineteenth-century France. Unpublished manuscript.
- Morris, N. (1972). 1870: The rating option. *History of Education* 1(1), 23–42.
- Parliament, U. K. (1861). Report of the commissioners appointed to inquire into the state of popular education in England. Vol. 21, paper 2794.
- Parliament, U. K. (1871a). Report of the Committee of Council on Education (England and Wales); 1870-71. Vol. 22, paper C.406.
- Parliament, U. K. (1871b). Return of civil parishes in England and Wales under Education Act, of population, rateable value, number of schools and scholars in attendance. Vol. 55.
- Parliament, U. K. (1874). Report of the Committee of Council on Education (England and Wales); with appendix 1873-74. Vol. 18, paper C.1019.
- Parliament, U. K. (1879). Report of the Committee of Council on Education (England and Wales); with appendix. 1878-79. Vol. 23, paper C.2342.
- Parliament, U. K. (1883). Committee of Council on Education. Instructions to H.M. Inspectors of schools, 1871, relative to inquiries into school supply of their districts. Vol. 53, paper C.3602.
- Parliament, U. K. (1889). Report of the Committee of Council on Education (England and Wales); with appendix 1888-89. Vol. 29, paper C.5804.

- Parman, J. (2011). American mobility and the expansion of public education. *The Journal of Economic History* 71(1), 105–132.
- Platten, S. G. (1975). The conflict over the control of elementary education 1870–1902 and its effect upon the life and influence of the church. *British Journal of Educational Studies* 23(3), 276–302.
- Rauscher, E. (2016). Does educational equality increase mobility? Exploiting nineteenth-century US compulsory schooling laws. *American Journal of Sociology* 121(6), 1697–1761.
- Richards, N. J. (1970). Religious controversy and the school boards 1870–1902. *British Journal of Educational Studies* 18(2), 180–196.
- Schurer, K. (2019). *Integrated Census Microdata (I-CeM) Names and Addresses, 1851-1911: Special Licence Access*. UK Data Service. [data collection].
- Schurer, K. and E. Higgs (2014). *Integrated Census Microdata (I-CeM), 1851-1911*. UK Data Service. [data collection].
- Sohn, H. and S.-W. Lee (2019). Causal impact of having a college degree on women’s fertility: Evidence from regression kink designs. *Demography* 56(3), 969–990.
- Southall, H. (2011). Great Britain historical GIS project. Data Collection.
- Southall, H. R. and P. Ell (2004). *Great Britain Historical Database : Census Data : Religion Statistics, 1851*. UK Data Service. Data Collection.
- Squicciarini, M. P. (2017). Devotion and development: Religiosity, education, and economic progress in 19th-century France. Unpublished working paper, Northwestern University.
- Squicciarini, M. P. and N. Voigtlander (2015). Human capital and industrialization: Evidence from the age of enlightenment. *The Quarterly Journal of Economics* 130(4), 1825–1883.
- Thompson, F. (1963). *English landed society in the nineteenth century*. London: Routledge and Kegan Paul.
- United Nations Educational, S. and C. Organization (2019). Global education monitoring report: Migration, displacement and education: Building bridges, not walls.
- West, E. G. (1970). *Education and the state: A study in political economy, 2nd edition*. The Institute of Economic Affairs.
- West, E. G. (1978). Literacy and the industrial revolution. *The Economic History Review* 31(3), 369–383.

Tables & Figures

Table 1: **Summary Statistics: Treated vs Untreated Parishes**

VARIABLES	1871		1881		1901	
	Board	No Board	Board	No Board	Board	No Board
Population of parish	1,604	718.5	1,838	837.2	2,008	930.4
School Supply per capita	0.0865	0.134				
Distance to London	208.81	194.92				
1851 Religious Attendance Per Capita						
Church of England	0.332	0.373				
Roman Catholic	0.0080	0.0117				
Other Religious Services	0.378	0.307				
Percent Occupied Within Sector						
Agriculture			0.178	0.289	0.136	0.215
Secondary			0.520	0.395	0.536	0.442
Tertiary			0.229	0.246	0.290	0.303
Occupation Requires Literacy			0.311	0.310	0.367	0.363
Parishes	2399	10711	2399	10711	2399	10711

Table 2: RK 2SLS Results at CCT Bandwidth

First Stage: School Board	
1871 School Supply/Pop	-2.8753*** (0.2410)
(Cutoff Dummy)*(1871 School Supply/Pop - 0.1666)	2.7606*** (0.5007)
Constant	0.5526*** (0.0336)
Second Stage: Proportion of jobs requiring literacy, ages 19-30	
School Board	0.1744** (0.0872)
1871 School Supply/Pop	0.2308 (0.1649)
Constant	0.1586*** (0.0366)
Parishes	5979
Bandwidth	0.064
Instrument F-Stat	30.39

Cutoff Dummy = $1_{[1871 \text{ School Supply/Pop} > 0.1666]}$. Robust standard errors in parentheses. Unit of observation is the parish. The second stage estimates the increase in the proportion of men aged 19-30 in 1901 employed in occupations requiring literacy due to the formation of a school board in the parish.*** p<0.01, ** p<0.05, * p<0.1

Table 3: **Full Sample Triple Difference**

	(1)	(2)	(3)
	Literacy Required	Literacy Required	Literacy Required
Years of Treatment	0.00136 (0.00108)	0.00130*** (0.000439)	0.00126*** (0.000446)
Observations	3,896,995	3,896,989	3,864,788
Controls	NO	NO	YES
Parish Fixed Effects	NO	YES	YES
Arrival-Age Fixed Effects	YES	YES	YES
Age Trend Effects	YES	YES	YES
Arrival Trend Effects	YES	YES	YES

Robust standard errors in parentheses, clustered at the parish level. Unit of observation is the individual. Controls (all interacted with a 1901 dummy): 1871 parish population; 1871 parish population squared; distance from London; distance from London squared; 1871 school supply per capita; 1851 Church of England attendees per capita 1851 Catholic attendees per capita; 1851 other religious service attendees per capita. *** p<0.01, ** p<0.05, * p<0.1

Table 4: **1861 Placebo**

	(1)	(2)	(3)
	Literacy Required	Literacy Required	Literacy Required
Years of Treatment	0.000788 (0.00142)	-0.000045 (0.000466)	-0.000037 (0.000474)
Observations	3,229,655	3,229,650	3,202,470
Controls	NO	NO	YES
Parish Fixed Effects	NO	YES	YES
Arrival-Age Fixed Effects	YES	YES	YES
Age Trend Effects	YES	YES	YES
Arrival Trend Effects	YES	YES	YES

Robust standard errors in parentheses, clustered at the parish level. Unit of observation is the individual. Controls (all interacted with a 1881 dummy): 1871 parish population; 1871 parish population squared; distance from London; distance from London squared; 1871 school supply per capita; 1851 Church of England attendees per capita 1851 Catholic attendees per capita; 1851 other religious service attendees per capita. *** p<0.01, ** p<0.05, * p<0.1

Table 5: **Only Treated**

	(1)	(2)	(3)
	Literacy Required	Literacy Required	Literacy Required
Years of Treatment	0.00278* (0.00163)	0.00225* (0.00118)	0.00240** (0.00119)
Observations	928,995	928,979	923,479
Controls	NO	NO	YES
Parish Fixed Effects	NO	YES	YES
Arrival-Age Fixed Effects	YES	YES	YES
Age Trend Effects	YES	YES	YES
Arrival Trend Effects	YES	YES	YES

Robust standard errors in parentheses, clustered at the parish level. Unit of observation is the individual. Controls (all interacted with a 1901 dummy): 1871 parish population; 1871 parish population squared; distance from London; distance from London squared; 1871 school supply per capita; 1851 Church of England attendees per capita 1851 Catholic attendees per capita; 1851 other religious service attendees per capita. *** p<0.01, ** p<0.05, * p<0.1

Table 6: **Full Sample Triple Difference. Dependent Variable: Class Dummy**

	(1)	(2)	(3)
VARIABLES	Class Dummy	Class Dummy	Class Dummy
Years of Treatment	0.00296*** (0.000741)	0.00158*** (0.000423)	0.00156*** (0.000426)
Observations	4,030,974	4,030,968	3,998,042
Controls	NO	NO	YES
Parish Fixed Effects	NO	YES	YES
Arrival-Age Fixed Effects	YES	YES	YES
Age Trend Effects	YES	YES	YES
Arrival Trend Effects	YES	YES	YES

Robust standard errors in parentheses, clustered at the parish level. Unit of observation is the individual. Controls (all interacted with a 1901 dummy): 1871 parish population; 1871 parish population squared; distance from London; distance from London squared; 1871 school supply per capita; 1851 Church of England attendees per capita 1851 Catholic attendees per capita; 1851 other religious service attendees per capita.
*** p<0.01, ** p<0.05, * p<0.1

Table 7: Social Mobility Triple Difference

VARIABLES	(1) Literacy Required	(2) Literacy Required	(3) Literacy Required
Years of Treatment	0.00210*** (0.000719)	0.00198*** (0.000719)	0.00237* (0.00127)
(Years of Treatment)*(Father Class Dummy)	-0.00309** (0.00136)	-0.00290** (0.00137)	-0.00324 (0.00245)
Observations	1,379,666	1,365,910	716,101
Controls Included	NO	YES	N.A.
Birth Parish-Class Fixed Effects Included	YES	YES	N.A.
Arrival-Age-Class Fixed Effects Included	YES	YES	YES
Age-Class Trend Effects Included	YES	YES	YES
Arrival-Class Trend Effects Included	YES	YES	N.A.
Brother Fixed Effects Included	NO	NO	YES

Robust standard errors in parentheses, clustered at parish level. Controls (all interacted with a 1901 dummy, and including interactions with father's class dummy): 1871 parish population; 1871 parish population squared; distance from London; distance from London squared; 1871 school supply per capita; 1851 Church of England attendees per capita 1851 Catholic attendees per capita; 1851 other religious service attendees per capita. *** p<0.01, ** p<0.05, * p<0.1

Table 8: **Triple Differences By Father's Class**

Father's Class	Years of Treatment	Observations
Higher Professionals	-0.00469 (0.00387)	35,560
Lower Professionals	-0.00106 (0.00239)	116,467
Skilled Workers	0.00027 (0.00145)	249,216
Farmers	0.00286 (0.00191)	145,305
Semi-Skilled Workers	0.00089 (0.00122)	348,709
Unskilled Workers	0.00387* (0.00217)	114,540
Unskilled Farm Labourers	0.00226* (0.00117)	347,875

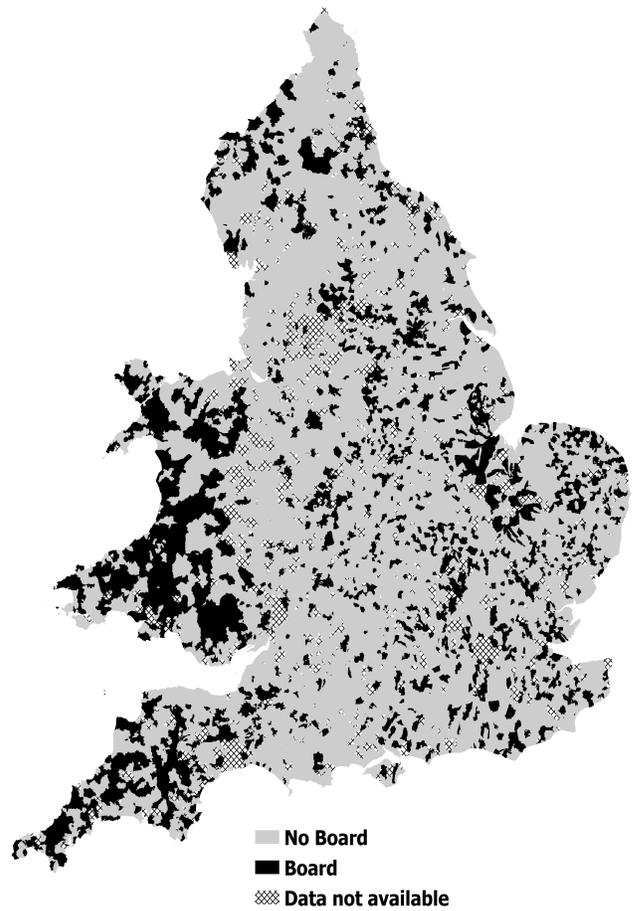
Robust standard errors in parentheses, clustered at the parish level. Unit of observation is the individual. Controls (all interacted with a 1901 dummy): 1871 parish population; 1871 parish population squared; distance from London; distance from London squared; 1871 school supply per capita; 1851 Church of England attendees per capita 1851 Catholic attendees per capita; 1851 other religious service attendees per capita. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Figure 1: School Supply and Board Formation, England and Wales

(a) Pre-treatment School Supply

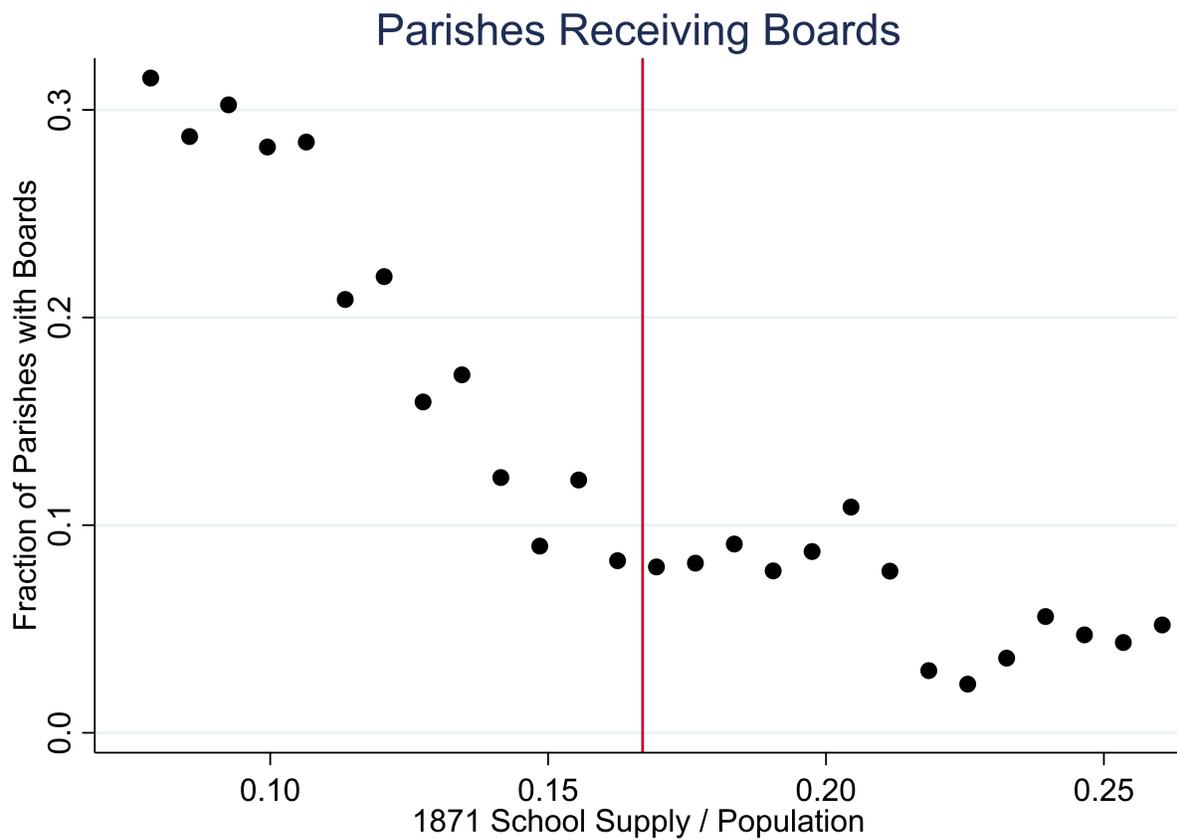


(b) School Boards



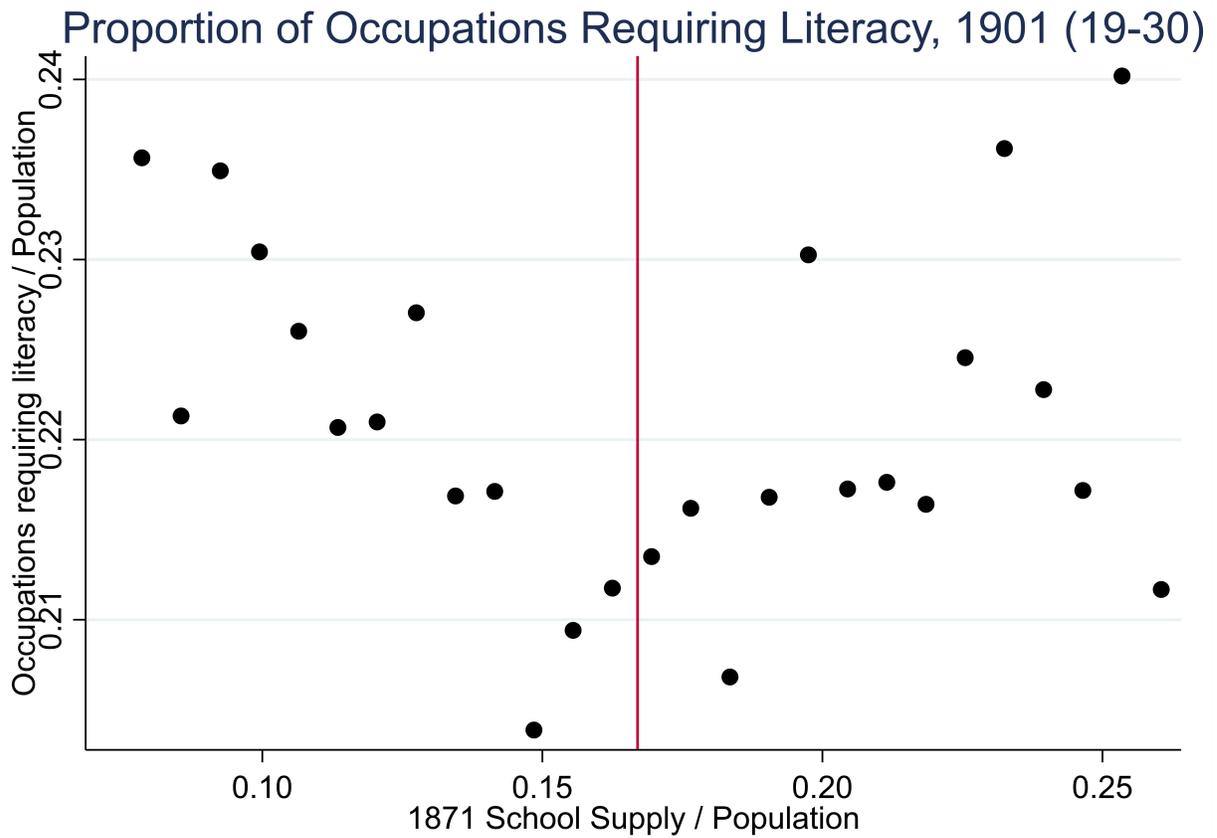
Note: Figure (a) maps the 1871 school supply to population ratio at the parish level. Parishes were considered sufficiently supplied if their ratio was at or above 1/6, or 0.167. Figure (b) maps which parishes received school boards between 1870 and 1879. Data are unavailable for parishes within municipal boroughs.

Figure 2: First Stage Kink



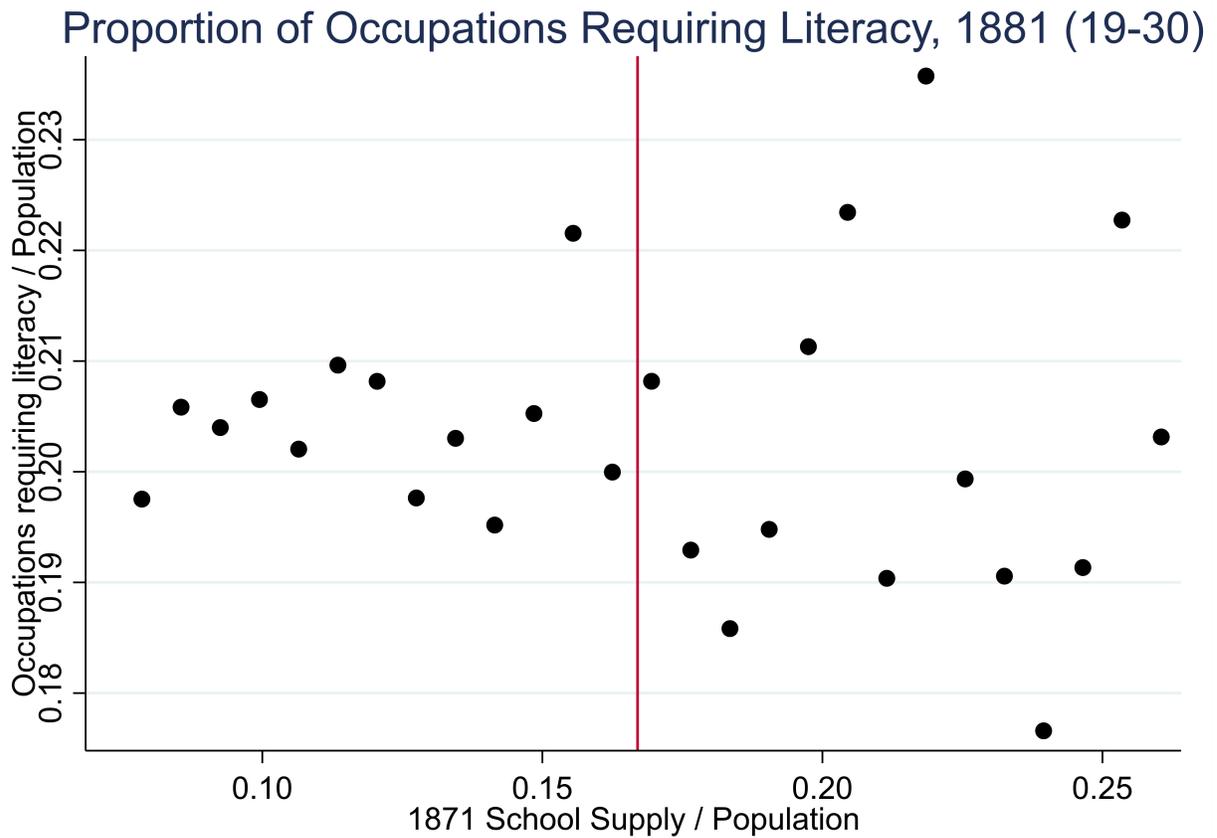
Note: Parishes are binned by 1871 school supply to population ratio using a bin width of 0.07. Figure 2 plots the 1871 parish school supply to population ratio against the fraction of parishes within each bin that received school boards between 1870 and 1879. The red line denotes the 0.166 cutoff.

Figure 3: Second Stage Kink



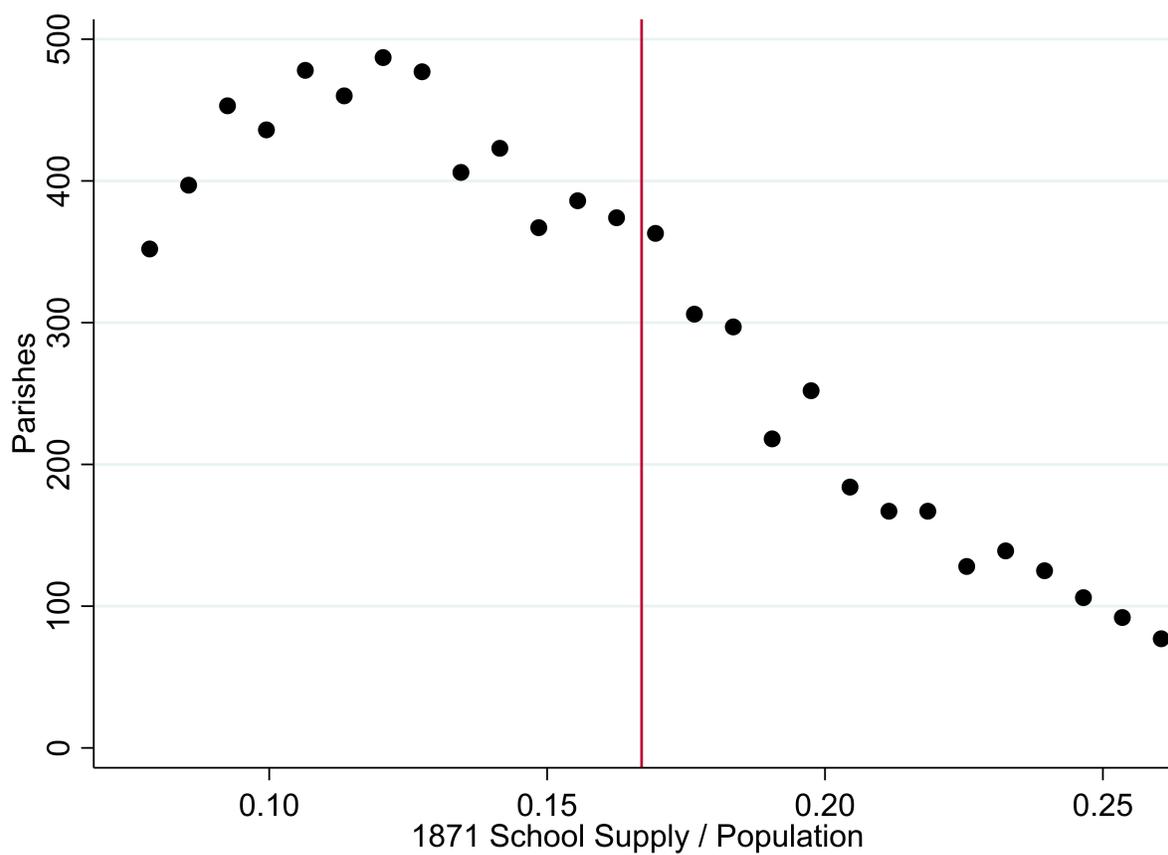
Note: Parishes are binned by 1871 school supply to population ratio using a bin width of 0.07. Figure 3 plots the 1871 parish school supply to population ratio against the average proportion of men aged 19-30 employed in occupations requiring literacy in 1901 within each bin. These men were of the correct age to have been treated by public schools. The red line denotes the 0.166 cutoff.

Figure 4: Second Stage Kink Placebo Test: Same Age, Untreated Year



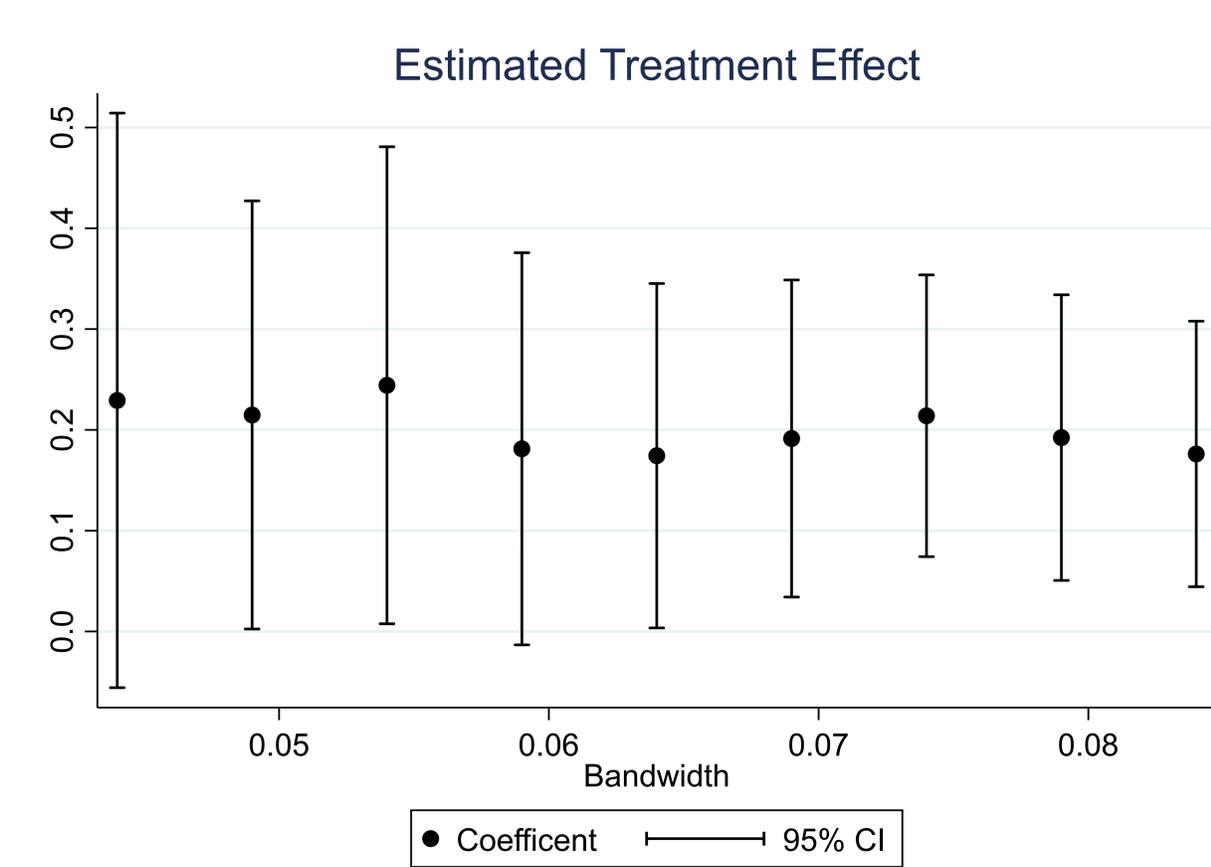
Note: Parishes are binned by 1871 school supply to population ratio using a bin width of 0.07. Figure 4 plots the 1871 parish school supply to population ratio against the average proportion of men aged 19-30 employed in occupations requiring literacy in 1881 within each bin. These men were too old to have been treated by public schools. The red line denotes the 0.166 cutoff.

Figure 5: Running Variable Density



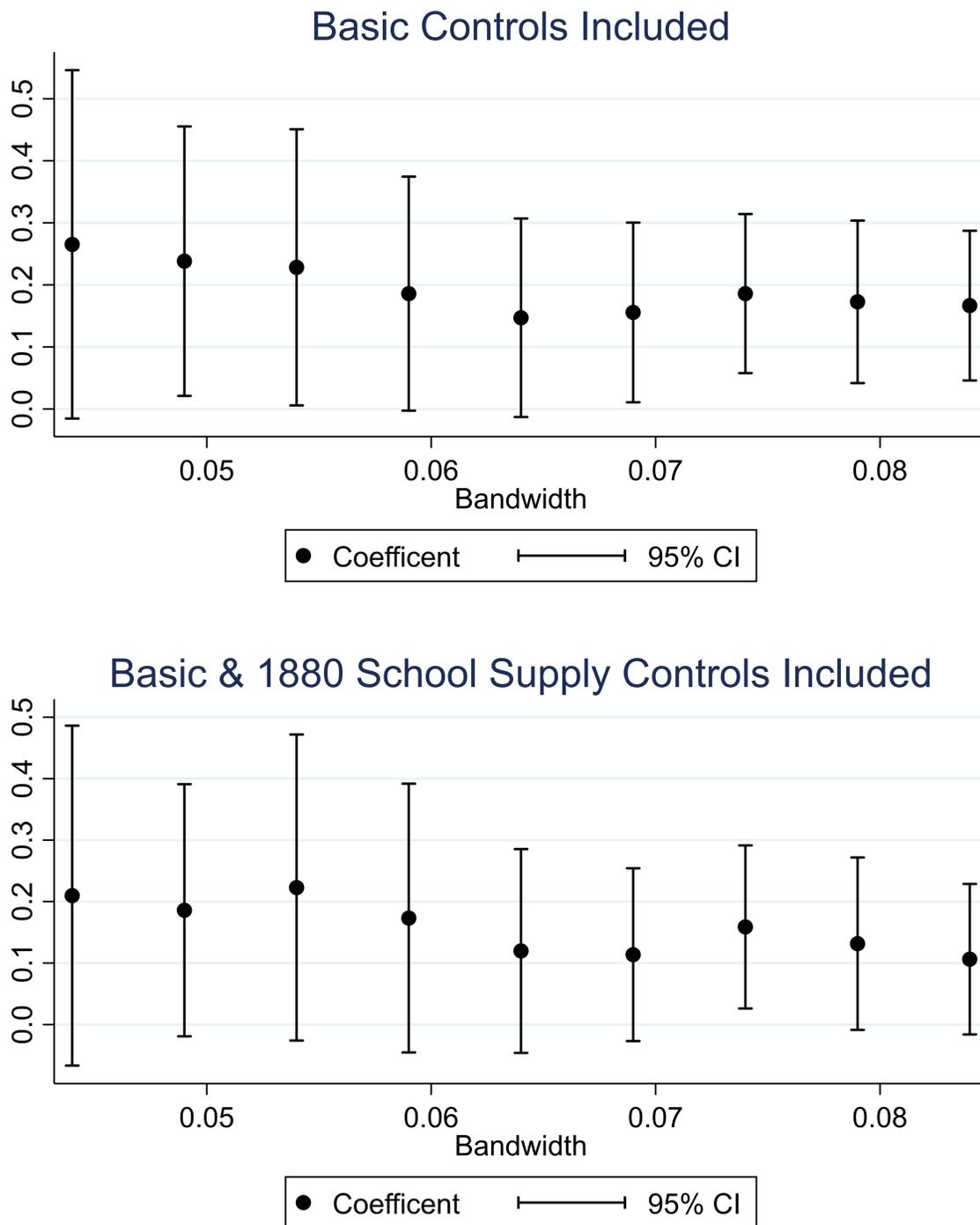
Note: Parishes are binned by 1871 school supply to population ratio using a bin width of 0.07. Figure 5 plots the 1871 parish school supply to population ratio against the number of parishes within each bin. The red line denotes the 0.166 cutoff.

Figure 6: Varying Bandwidth



Note: Figure 6 plots for varying bandwidths the second stage estimates from the regression kink analysis, representing the treatment effect of school boards on the proportion of men aged 19-30 employed in occupations requiring literacy. Estimates using the following bandwidths are plotted: $h = 0.064 + c * (0.005)$, where $c \in \{-4, -3, \dots, 3, 4\}$. The optimal bandwidth $h = 0.064$ was selected using the method of Calonico et al. (2014).

Figure 7: Varying Bandwidth, Covariates Included



Note: Figure 7 plots for varying bandwidths the estimated treatment effect of school boards on the proportion of men aged 19-30 employed in occupations requiring literacy in 1901. The top panel results are obtained from the regression kink analysis with the following covariates included: 1871 population; 1851 attendance per capita at churches outside the Church of England or Roman Catholic traditions; 1881 agricultural employment per capita; 1881 secondary sector employment per capita; proportion of men aged 19-30 employed in occupations requiring literacy in 1881. The bottom panel results include all the above mentioned covariates, as well as 1880 school supply per capita. Estimates using the following bandwidths are plotted: $h = 0.064 + c \cdot (0.005)$, where $c \in \{-4, -3, \dots, 3, 4\}$.

A Appendix: Multi-Parish Schools

Schools of the period were typically meant to serve only those residing in the local parish. This feature arose from to the Church of England’s historical influence, with local clergy traditionally responsible not only for the spiritual but also the temporal needs of their flock. However, the 1871 survey of school space indicates that in a minority of parishes (11.5%), local children made use of schools in nearby parishes. Unsurprisingly, these parishes were typically very small, with an average 1871 population of just under 170. The patterns of school board formation also indicate that the duty of school provision was occasionally shared: by 1879, 34% of boards spanned more than one single parish. These “unified” boards, as they were called, typically consisted of multiple very small parishes, or a very small parish (or parishes) grouped with a single larger parish.

This paper uses parish school supply and population figures recorded in the 1871 survey, and assumes the authorities of the period used the same figures when determining whether supply was sufficient given population. Given that the 1871 survey was originally administered for precisely this purpose, this assumption seems reasonable. However, care must be taken in considering how authorities applied data from the 1871 survey. It seems likely that for those parishes described as sharing school space in 1871, as well as those joined together in unified boards, authorities making treatment decisions considered total school supply across all parishes within the group, and compared this with total group population. Thus, it is this group school supply to population ratio that is used in these cases for the analysis in the paper.

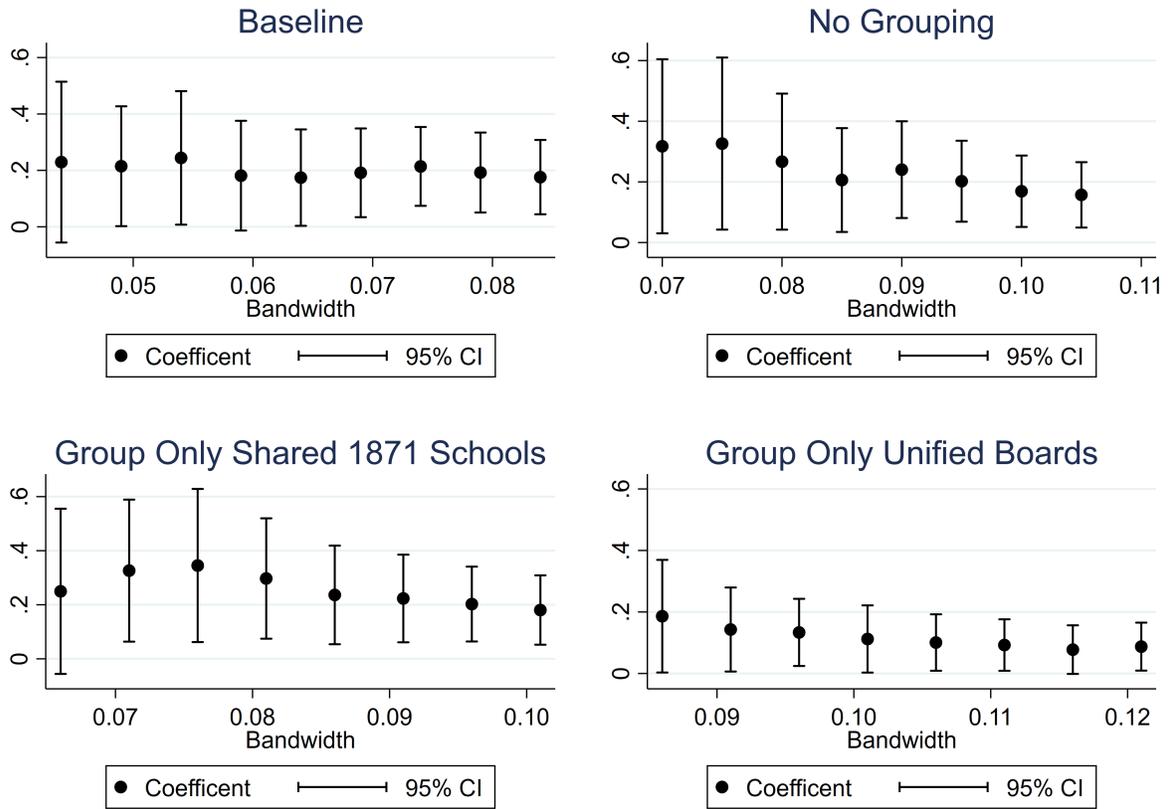
It is important to note, however, that the results of the paper are in no way dependant on this. Table A.1 shows the results of the FRK using other methods of determining school supply per capita from the 1871 figures. Column 1 reproduces the results shown in Table 2, obtained using the preferred method. Column 2 provides the results obtained by simply using each individual parish’s own school supply and population, making no adjustments for grouped parishes. The method used in column 3 only groups those parishes listed as sharing schools in the 1871 survey are grouped together, while that for column 4 only groups together those joined within the same unified board. While the estimated treatment effect varies across cases, it is strongly positive and significant regardless of method, never dropping below 10 pp. It should be noted that while each specification is estimated using its own optimal CCT bandwidth, in each case the results are robust to varying the bandwidth, as can be seen in Figure A.1.

Table A.1: RK 2SLS, Various Grouping Methods

First Stage: School Board	(1)	(2)	(3)	(4)
1871 School Supply/Pop	-2.87528*** (0.24098)	-2.1538*** (0.1770)	-1.9508*** (0.1705)	-3.1424*** (0.1537)
D*(1871 School Supply/Pop - 0.1666)	2.76059*** (0.50073)	1.9769*** (0.3606)	1.801*** (0.3496)	2.6137*** (0.3174)
Second Stage: Proportion of jobs requiring literacy, ages 19-30, 1901				
School Board	0.17437** (0.08718)	0.2403*** (0.0814)	0.2363** (0.0931)	0.1007** (0.0468)
1871 School Supply/Pop	0.23084 (0.16488)	0.1843 (0.1168)	0.1629 (0.1212)	0.1225 (0.1032)
Parishes	5979	6725	7463	7642
Bandwidth	0.064	0.090	0.086	0.106
Instrument F-Stat	30.39	30.05	26.54	67.82

$D = 1_{[1871 \text{ School Supply/Pop} > 0.1666]}$. Column 1 reproduces the results shown in Table 2, obtained using the preferred method for determining pre-treatment school supply per capita. Column 2 provides the results obtained by using each individual parish's own school supply and population, making no adjustments for grouped parishes. The method used in column 3 only adjusts for those parishes listed as sharing schools in the 1871 survey are grouped together, while that used in column 4 only adjusts for those joined within the same unified board.*** p<0.01, ** p<0.05, * p<0.1

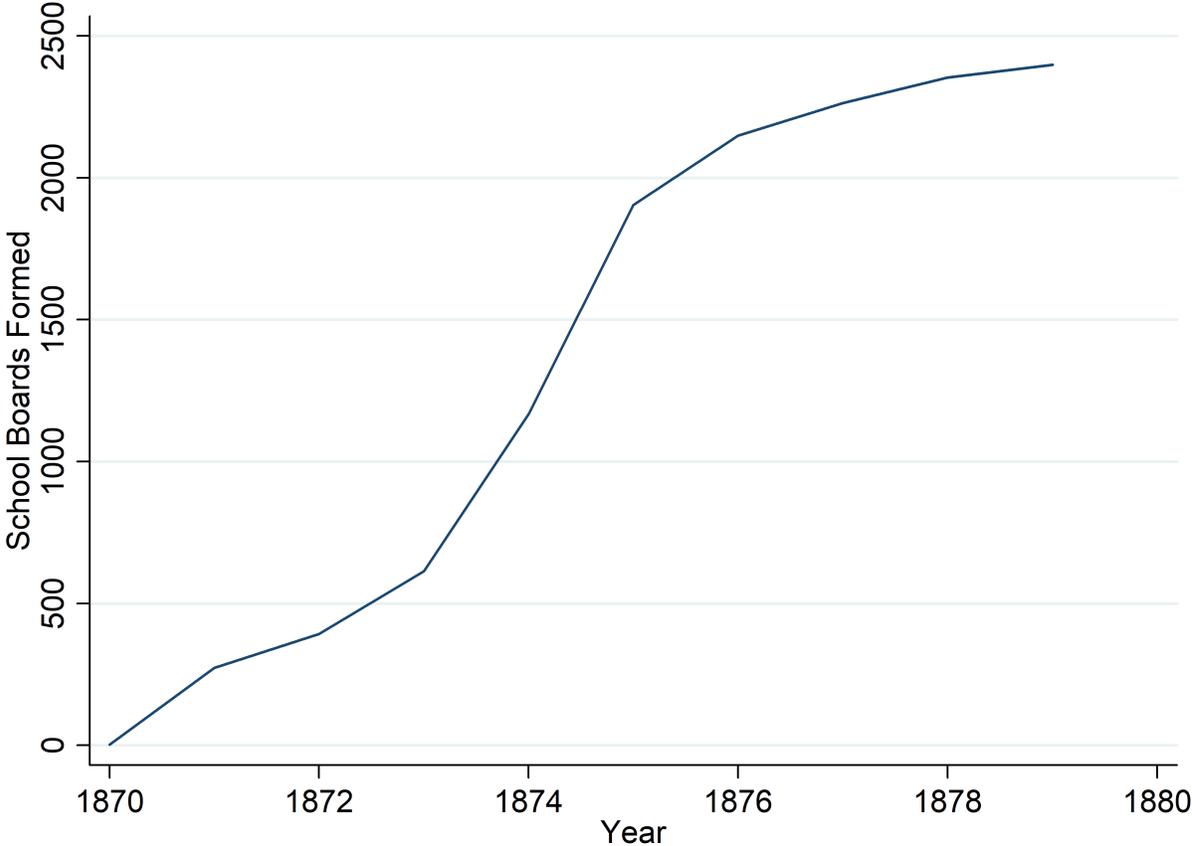
Figure A.1: Estimated Treatment Effect Using Various Bandwidths, Grouping Methods



Note: Each figure plots for varying bandwidths the estimated treatment effect of school boards on the proportion of men aged 19-30 employed in occupations requiring literacy in 1901. In the top left, school supply per capita is determined using the preferred method, summing 1871 school supply and population across parishes when said parishes either shared a school according to the 1871 survey, or were eventually were joined together in a unified school board. In the top right, only a parishes own school supply and population are used to determine supply per capita. In the bottom left, school supply and population are only summed across parishes when said parishes share a school according to the 1871 survey. In the bottom right, school supply and population are only summed across parishes that were eventually were joined together in a unified school board. In each case, estimates using the following bandwidths are plotted: $h = B + c * (0.005)$, where $c \in \{-4, -3, \dots, 3, 4\}$, where B is the optimal bandwidth calculated given the data.

B Appendix: Board Formation by Year

Figure B.1: School Board Formation by Year



Note: Figure B.1 plots the take up of public school boards following the passage of the 1870 Education Act.

C Effect of Boards on School Supply and Attendance

In this section I verify the positive effects school boards had on school supply and attendance. These ‘stage zero’ results are included as motivation for using the school board reform as a proxy for improved education.

A simple difference-in-differences (DID) specification, shown in Equation 13, is used to estimate the impact of school boards on school supply.⁵² Pre-reform school space and population are drawn from the 1871 education survey, while post-reform figures are from either the 1880 or 1888 Board of Education Report to Parliament, depending on the sample.

$$\frac{\text{School Supply}}{\text{Population}}_{pt} = \omega_p + \gamma_1 t + \gamma_2 (S_p * t) + (Z_p * t)\beta + \epsilon_{pt} \quad (13)$$

where t is a year dummy equal to 1 if from 1880 or 1888 (depending on the sample) and 0 if from 1871, ω_p is a parish fixed effect, S_p is a school board dummy (1 if parish received school board, 0 otherwise), and Z_p is a vector of pre-treatment parish controls (interacted with time to allow trend to vary on them).

The 1880 and 1888 school supply figures, drawn from the Board of Education’s reports to Parliament in those years, were originally given at the school level, organized by location. To allow comparison with the pre-treatment 1871 school supply figures, which are only available at the parish level, where possible I aggregate the 1880 and 1888 data to the parish level and match to the 1871 survey. Note that a parish lacking any school in 1880 (1888) would be missing from the the 1880 (1888) report, and thus go unmatched. However, this is not the only possible explanation for a lack of a match: while the majority of locations given in the 1880 and 1888 reports are at the parish level, a significant minority of locations are conglomerations of more than one parishes, or are ambiguous; in these cases, I am unable to match back to the 1871 survey. Just under 24% of parishes included in the 1871 survey and described as having at least one school at that time are unmatched to the 1880 report; while presumably some of these went unmatched due to their schools closing between 1871 and 1880, it seems likely that most did not match due to the location issue described above.

Given this ambiguity, I conduct the DID analysis using two separate approaches. First, I use only the sample of matched parishes. This analysis can be viewed as assessing the effect of school boards on parish school supply given supply was non-zero at $t = 1$. If, as expected, school boards decreased the likelihood of a parish having no school in the late period, this method would underestimate boards’ total effect on school supply. This estimates should thus be considered a lower bound.

In the second approach I rerun the DID analysis but include all parishes, assuming zero

⁵²A parish’s school supply was at the time defined as (total school square footage)/10.

school supply in the late period for those unmatched to the 1880 (or 1888) data. As previously discussed, this will falsely attribute zero school supply to parishes which went unmatched due to data ambiguity. However, assuming this measurement error is uncorrelated with school board formation, the estimated treatment effect of school boards will be unbiased.

Results from the first approach are shown in Table C.1: columns (1), (2), and (3) give the results comparing 1871 with 1880, with and without parish fixed effects and pre-treatment controls, while columns (4), (5), and (6) compare 1871 with 1888. Even given its likely underestimation, the estimated treatment effect of school boards is positive throughout, and appears to grow over time. The coefficients suggest a 1.55-1.97 pp increase in the school supply ratio due to school boards by 1880, although this effect is not statistically significant. In contrast, by 1888, the effect of school boards is strongly significant across both specifications, increasing the school supply ratio by 2.6 pp. In 1871 the average treated parish had a school supply ratio of 8.65%, implying that by 1888 school boards had increased school supply ratio 30% in these parishes.

Table C.2 reports the results using the full sample, assuming that those parishes unmatched to the 1880 (or 1888) report had zero school supply in the late period. As expected, the estimated treatment effects are larger than those obtained using only the matched sample. Interestingly, in this case the effect of school boards on school supply seems to diminish slightly over time – school boards are associated with a 5.93-6.03 pp increase in the school supply ratio in 1880, but only a 4.25-4.96 pp increase in 1888. Regardless, the results strongly suggest that boards led to a large increase in school supply that persisted long after they were initially formed.

While the DID approach controls for constant parish-level unobservables, it relies on the parallel trends assumption. That is to say, we must assume that, in the absence of the reform, both parishes that received the reform and those that did not would have experienced on average the same growth in school supply per capita, given the pre-treatment controls. Unfortunately, parish-level school supply prior to 1871 is not available, and thus testing for parallel trends prior to this period is impossible. Admittedly, one could think of reasons why the assumption might not hold. As described in Section 3, parishes had some control over whether or not they received a board. Thus, if only those parishes that were most eager to expand their school supply received boards, the parallel trends assumption would likely be violated and these results could not be interpreted causally. If it were the case that boards primarily went to eager parishes, one would expect most to have been formed within a year or two of the passage of the reform, for as previously stated any parish could voluntarily receive a school board whenever they desired. However, this was not the case; parishes that ended up with boards typically did not form them until several years after the reform was

passed (this can be observed in Figure B.1). Thus, it seems likely that the opposite was true: parishes that took it upon themselves to supply school space through private means in the years immediately following the reform were not obliged to institute boards, and thus only those parishes least eager to expand school supply ended up receiving boards. Given this, the results shown here strongly support the notion that school boards caused school supply to increase, refuting the claims made by West (1970) that the boards actually held back school growth.

Attendance figures also suggest boards had a positive impact on education. As attendance figures are only available for 1888, using DID is not possible. However, simple regressions with controls, shown in Table C.3, suggest boards were correlated with higher school attendance per capita. Not surprisingly, school supply per capita is highly correlated with attendance rates, but even controlling for this, boards appear to have a strong positive impact. Given the inability to control for parish unobservables, these results should obviously be interpreted with caution. However, they line up with the fact that boards had more authority to enforce school attendance and also subsidized fees. As well, the likely increase in school quality brought by boards may have made attendance more attractive to parents and students.

Whether or not these results can be interpreted causally, they demonstrate that there was an improvement in the provision of education in parishes that received boards relative to those that did not. Indeed, these results likely tell only part of the story, as the evidence of superior instruction given at board schools, discussed in Section 3, is not captured here. However, they motivate the use of the school board reform as a proxy for improved education.

Table C.1: School Supply Per Capita Diff-in-Diff, Matched Parishes Only

	1871-1880 Comparison		1871-1888 Comparison	
	(1)	(2)	(3)	(4)
	School Supply to Pop Ratio			
(Board)*(Time Dummy)	0.0155 (0.0116)	0.0197 (0.0122)	0.0260** (0.0122)	0.0259*** (0.0119)
Time Dummy	0.0977*** (0.0050)	0.0977*** (0.0326)	0.1131*** (0.0044)	0.1095*** (0.0259)
Board	-0.0631*** (0.0022)		-0.0639*** (0.0024)	
Constant	0.1496*** (0.0017)		0.1507*** (0.0017)	
Observations	17,466	17,352	16,730	16,621
Pre-trend Controls	NO	YES	NO	YES
Parish Fixed Effects	NO	YES	NO	YES

Robust standard errors in parentheses. Pre-trend controls (all interacted with a time dummy): 1871 parish population; distance from London; distance from London squared; 1851 Church of England attendees per capita 1851 Catholic attendees per capita; 1851 other religious service attendees per capita. *** p<0.01, ** p<0.05, * p<0.1

Table C.2: School Supply Per Capita Diff-in-Diff, All Parishes

	1871-1880 Comparison		1871-1888 Comparison	
	(1)	(2)	(3)	(4)
	School Supply to Pop Ratio			
(Board)*(Time Dummy)	0.0603*** (0.0096)	0.0593*** (0.0096)	0.0425*** (0.0089)	0.0496*** (0.0089)
Time Dummy	0.0217*** (0.0035)	0.0335 (0.0233)	0.0310*** (0.0032)	0.0775*** (0.0197)
Board	-0.0478*** (0.0020)		-0.0478*** (0.0020)	
Constant	0.1338*** (0.0014)		0.1338*** (0.0014)	
Observations	26,152	25,982	26,152	25,982
Pre-trend Controls	NO	YES	NO	YES
Parish Fixed Effects	NO	YES	NO	YES

Robust standard errors in parentheses. Pre-trend controls (all interacted with a time dummy): 1871 parish population; distance from London; distance from London squared; 1851 Church of England attendees per capita 1851 Catholic attendees per capita; 1851 other religious service attendees per capita. *** p<0.01, ** p<0.05, * p<0.1

Table C.3: **1888 School Attendance Rate**

VARIABLES	School Attendance to Pop Ratio	School Attendance to Pop Ratio
Board	0.0222*** (0.00290)	0.0210*** (0.00257)
Observations	7,156	7,156
Controls	NO	YES

Robust standard errors in parentheses. Controls: 1871 parish population; distance from London; distance from London squared; 1871 school supply per capita; 1851 Church of England attendees per capita 1851 Catholic attendees per capita; 1851 other religious service attendees per capita. *** p<0.01, ** p<0.05, * p<0.1

D Appendix: Childhood vs Adult Parish of Residence

If access to public schooling during childhood is truly driving results, then it should be the case that childhood parish of residence is what matters, not adult residence. When using the complete, unlinked censuses in Sections 5 and 6, I am unable to view childhood parish of residence, and address this issue by dropping individuals residing in a county other than the one of their birth, as these individuals had likely moved since childhood. Using the census linkages, however, I directly observe childhood place of residence, and so can keep these movers. To show that dropping them in the previous sections had little effect on my results, I rerun the regression displayed in Column (3) of Table 3, but using the linked samples. The result is shown in Column (1) of Table D.1.

As a placebo test, I determine how many years an individual would have been treated had they resided in their adult parish of residence during childhood. In Column (2) of Table D.1, I show the results using this adult residence treatment variable. As expected, the results are insignificant and close to zero. Finally, in Column (3) of Table 3, I run a horse race regression, including the treatment variables associated with both childhood and adult parish of residence. The results again strongly suggest that it is childhood access to public schools that is driving results.

Table D.1: **Childhood vs. Current Parish of Residence**

	(1)	(2)	(3)
	Literacy Required	Literacy Required	Literacy Required
Years of Treatment (Childhood)	0.000948** (0.000481)		0.00131** (0.000638)
Years of Treatment (Current)		-0.000251 (0.000561)	-0.00102* (0.000611)
Observations	2,042,790	1,491,469	1,433,523
Childhood Arrival-Age Fixed Effects	X		X
Childhood Arrival Trend Effects	X		X
Childhood Parish Fixed Effects	X		X
Current Arrival Trend Effects		X	X
Current Arrival-Age Fixed Effects		X	X
Current Parish Fixed Effects		X	X
Age Trend Effects	X	X	X

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Figure E.2: 1879 School Boards

School Boards.	Union.	A.D.	Popula- tion. (1871.)	No. of Members.	Date of Election.
CAMBRIDGE—cont.					
<i>Dullingham</i> - -	Newmarket - -	4	818	6	21 May 1875
(c.) <i>Foxton</i> - -	Royston - -	4	413	5	16 Oct. 1876
<i>Gamlingay</i> - -	Caxton and Ar- rington.	4	2,063	6	31 Dec. 1874
(c.) <i>Great and Little Ab- ington (U.D.).</i>	Linton - -	4	639	5	26 Nov. 1873
<i>Great Abington</i> -					
<i>Little Abington</i> -					
<i>Haddenham</i> - -	Ely - -	4	2,055	5	3 Jan. 1874
(c.) <i>Harston</i> - -	Chesterton - -	4	917	5	23 Mar. 1876
(c.) <i>Haurton</i> - -	Chesterton - -	4	289	5	23 Mar. 1876
(c.) <i>Impington</i> - -	Chesterton - -	4	387	5	28 Sept. 1875
(c.) <i>Leverington</i> - -	Wisbeach - -	31	1,315	5	20 Dec. 1875
<i>Littleport</i> - -	Ely - -	4	3,869	5	14 April 1874
(c.) <i>Maneu</i> - -	North Witchford	4	1,311	5	4 April 1876
<i>March</i> - -	North Witchford	4	5,854	7	1 June 1871
(c.) <i>Pampisford</i> - -	Linton - -	4	355	5	9 Aug. 1875
<i>Parson Drove</i> - -	Wisbeach - -	31	808	5	13 Jan. 1875
(c.) <i>Rampton</i> - -	Chesterton - -	4	256	5	9 Jan. 1874
<i>Sawston</i> - -	Linton - -	4	1,729	5	8 April 1872
<i>Soham</i> - -	Newmarket - -	4	4,283	5	10 Mar. 1871
<i>Staplesford</i> - -	Chesterton - -	4	594	5	27 July 1875
(c.) <i>Swaffham Priory</i> - -	Newmarket - -	4	1,369	5	(no election)
<i>Thorney</i> - -	Peterborough - -	24	2,099	5	11 Feb. 1875
<i>Upwell. See Upwell (Norfolk).</i>					
<i>Waterbeach</i> - -	Chesterton - -	4	1,619	5	12 Feb. 1875
(c.) <i>Whittlesey St. Mary and St. Andrew.</i>	Whittlesey - -	4	7,002	7	26 Nov. 1875
(c.) <i>Wicken</i> - -	Newmarket - -	4	1,133	5	8 Mar. 1876
<i>Willingham</i> - -	Chesterton - -	4	1,619	5	26 Feb. 1873
			67,452		

Note: Sample page from the list of school boards provided by the Board of Education to Parliament in 1879. Source: UK parliamentary paper, 1878-79, Vol. 23, paper C.2342.

Figure E.3: 1888 Schools

548

Schools aided by Parliamentary Grants.

No. of Dist. or Union in Census Tables.	Name and Denomination of School, and Month when Inspection is due.	No. of Scholars for whom Accommodation is provided.	Annual Grants.			No. of Dist. or Union in Census Tables.	Name and Denomination of School, and Month when Inspection is due.	No. of Scholars for whom Accommodation is provided.	Annual Grants.		
			Average Attendance.	Amount.					Average Attendance.	Amount.	
			£	s.	d.				£	s.	d.
DEVON—cont.											
271	Rewe and Netherex - Ch. 5	85	37	27	3 10	274	Staverton - - - Ch. 6	86	47	37	17 5
286	Roborough - - - N. 12	97	51	42	13 9	280	Stockland - - - Ch. 3	66	60	58	2 0
271	Rockbeare - - - Ch. 11	66	68	49	13 0	282	Stockleigh Pomeroy - Ch. 6	66	38	36	7 10
	Rockbeare, Marsh Green - - - Ch. 11	60	45	47	18 9	271	Stoke Canon - - - N. 5	107	57	47	16 0
284	Romansleigh and Mariansleigh - - - 3	81	34	19	6 2	279	Stoke Damerel, Boys - B. 11	205	422	432	9 8
284	Rose Ash - - - N. 3	92	51	40	7 0	288	Stoke Damerel, Girls - B. 11	288	295	15	10 0
269	Rousdon - - - Ch. 1	163	83	85	14 4	275	Stoke Fleming - - - 5	110	100	87	10 0
276	St. Budeaux - - - N. 4	148	122	10	0 0	274	Stoke Gabriel - - - 6	145	95	68	18 6
	St. Budeaux, Bull Point N. 4	73	36	31	10 0	273	Stokeinteighcad - - - 7	80	64	53	14 4
291	St. Giles-in-the-Heath N. 2	61	29	34	19 9	275	Stokenham, Hucombe 5	110	98	74	13 0
286	St. Giles-in-the-Wood - N. 2	207	140	122	10 0		Stokenham - - - 5	213	125	102	16 6
275	St. Mary Church - - - N. 3	493	351	307	2 6		Stokenham, Evening - Ch. 11	36	19	8	16 0
	St. Mary Church, Hele N. 3	100	53	38	8 6	285	Stoke Rivers - - - 11	36	27	35	17 0
	St. Mary Church, Priory - - - R. 3	215	119	109	0 4	278	STONEHOUSE, EAST:				
290	St. Mary Tavy - - - 4	216	85	74	7 6		East Street - - - 11	598	312	273	0 0
273	St. Nicholas, Shaldon - 3	216	212	195	6 0		High Street - - - 11	554	556	502	1 0
280	St. Peter Tavy - - - N. 4	116	39	35	1 6	283	National - - - 1	594	450	533	15 0
271	St. Thomas - - - 5	883	786	688	12 6	280	St. Paul's Infant - - Ch. 11	223	158	152	7 0
	St. Thomas, Exwick - 5	157	123	106	2 9	275	Stoodleigh - - - Ch. 11	90	72	60	15 0
275	Salcombe, Infant - - 5	153	112	96	8 0	288	Stowford - - - Ch. 7	78	34	23	1 0
	Salcombe - - - N. 5	329	149	139	7 5	275	Street - - - N. 5	117	55	38	1 11
270	Salcombe Regis - - - N. 4	62	36	25	10 0	288	Sutcombe - - - 5	69	52	45	10 0
281	Sampford Courtenay 6	150	88	74	15 6	277	Sutton-on-Plym - - - N. 11	607	478	418	5 0
	Sampford Courtenay Sticklepath - - - 6	53	51	44	12 6	285	Swimbridge - - - P. 11	160	101	93	9 3
283	Saunpford Peverell - - N. 11	165	81	64	15 6	280	Swimbridge, Travellers' Rest - - - P. 11	41	26	26	12 4
282	Sausford - - - N. 7	200	120	93	10 8	286	Sydenham Damerel - N. 4	91	60	49	5 0
	Sausford, East Village Ch. 7	42	36	31	9 0	270	Taddiport, Infant - Ch. 2	68	37	23	2 6
	Sausford, New Buildings 7	44	31	21	3 8	270	Talaton - - - Ch. 4	82	39	27	11 0
269	Seaton - - - B. 1	209	138	104	0 5	276	Taleford - - - Ch. 4	71	47	41	2 6
276	Shaugh Prior - - - 4	78	44	27	9 0	276	Tamerton Pliott, Endd. 10	175	159	133	9 0
	Shaugh Prior, Lee Moor 4	179	94	58	17 1	280	TAVISTOCK:				
	Shaugh Prior, Lee Moor, Evening - - - N. -	—	10	3	14 0		Gulworthy - - - 4	120	69	67	10 0
286	Shebbear - - - 2	157	79	64	16 11		Morcelham - - - 4	57	30	27	15 0
	Shebbear, Newton, St. Petrock - - - 2	60	43	45	14 7	285	National - - - 4	510	263	230	6 10
286	Sheepwash - - - N. 2	106	61	43	19 9		Plymouth Road - - - 4	760	365	288	12 7
270	Sheldon - - - N. 3	57	26	33	10 6	271	Tawstock, Harracott - N. 10	83	45	36	16 6
275	Sherford - - - 5	72	55	43	11 11	273	Tawstock, Hollivell - Ch. 11	128	60	50	9 0
273	Shipway Collaton - - 2	40	34	31	19 6	271	Tedburn, St. Mary's - 6	120	89	83	7 0
285	Shirwell - - - 11	168	27	26	17 9	273	Teignrace - - - N. 7	65	31	35	10 3
282	Shobrooke - - - 6	116	91	72	6 3	273	TRIGSMOUTH:				
269	Shute - - - 2	148	74	64	11 0		Exeter Road - - - 3	463	294	215	2 4
270	Sidbury - - - N. 4	246	136	111	11 6		Upper Brook Street - 3	427	351	322	18 6
270	Sidbury Sidford, Infant N. 4	57	16	10	0 0		Roman Catholic - - 3	90	66	46	19 6
270	Sidmouth, All Saints' - N. 4	227	214	187	3 3	283	Templeton - - - P. 11	42	36	37	13 0
	Sidmouth, Boys - - P. 4	169	66	49	0 6	283	Thornbury - - - N. 5	53	37	28	18 3
	Sidmouth, Girls and Infants - - - P. 4	230	108	75	12 6	283	Thorverton - - - N. 11	253	154	134	1 0
183	Silverton - - - N. 11	117	91	69	2 5	281	Throuleigh and Gidleigh 6	81	32	22	11 4
	Silverton, Endowed - 11	103	85	70	2 6	275	Thurleston - - - Ch. 5	77	45	35	1 9
275	Slapton - - - 5	122	90	68	3 8	270	Tipton, St. John's - N. -	*			
270	Smeathorpe - - - 3	42	194	179	11 2	283	TIVERTON:				
274	South Brent - - - 12	265	15	6	19 11		Bampton Street - - 10	145	102	76	19 0
275	South Huish - - - Ch. 5	120	38	37	0 4		Bolham - - - 10	161	59	41	17 3
270	South Leigh - - - 3	60	61	45	15 4		Chevithorne - - - 10	190	94	69	6 8
275	South Milton - - - Ch. 5	102	107	85	7 6		Core - - - 19	84	47	44	13 6
284	South Molton - - - B. 3	414	251	235	14 2		Elmora - - - 19	210	152	129	13 6
	South Molton - - - N. 3	78	71	64	14 3		Heathcote - - - 10	784	538	584	4 0
275	South-Pool - - - N. 5	206	159	131	13 0		National - - - 10	551	256	218	2 6
181	South Tawton - - - 5	48	20	34	9 8		Wittheigh - - - 10	57	39	22	19 6
	South Tawton, Langdown 6	95	45	23	2 6	271	Topsham - - - 11	327	285	244	19 9
271	Sout. Woolford, Infant N. 11	84	29	19	12 5	273	Tor - - - N. 1	553	479	417	19 5
271	Sowton - - - Ch. 11		88	77	0 0	273	TORQUAY:				
276	Sparkwell (Plympton) Village - - - 6	147	51	43	8 9		Abbey Road - - - B. 1	595	232	201	12 4
281	Spreyton - - - 6	75	88	77	0 0		Abbey Road - - - R. 1	86	63	49	4 3
							National - - - 1	439	288	245	3 5
							St. John's - - - N. 1	202	131	110	2 10
							St. Luke's, Girls & Infants, Ch. 1	161	90	64	5 3
							Trust, Trinity - - - N. 1	345	166	149	7 7
							Upton Vale - - - B. 1	521	202	187	6 2
							Totnes - - - Ca. 12	346	98	81	9 9
							Totnes Grove - - - 12	424	260	203	4 10

Note: Sample page from the list provided by the Board of Education to Parliament of schools, both board and private, receiving public grants 1888. Board (public) schools are italicized. Source: UK parliamentary paper, 1889. Vol. 29, paper C.5804.

F Appendix: Full English & Welsh Population vs Population Matched to School Data

Only parishes situated outside of municipal boroughs were included in the 1871 survey of school space used to determine pre-treatment school supply. Thus, parishes in the cores of many major cities, including London, are excluded from the analysis. Nonetheless, it would be wrong to view the remaining parishes as exclusively rural. Many cities did not have borough status, while many densely populated suburbs lay outside of boroughs. Summary statistics from the 1881 census comparing the general population with the population residing in parishes included in the 1871 survey are provided in Table F.1.

Table F.1: Full 1881 Population vs Population Matched to School Data

VARIABLES	(1) Full	(2) Matched to School Data
Age	30.38	30.35
Population Density > 4/ acres	0.560	0.245
Percent Occupied Within Sector		
Agriculture	0.147	0.252
Secondary Sector	0.466	0.437
Tertiary Sector	0.315	0.240
Literacy Required	0.386	0.311
N	6,072,649	3,018,783

G Appendix: Literacy and Class Categorizations

Using job adverts published in 19th century English periodicals, as well as other contemporaneous descriptions of occupations, Mitch (1992) grouped English Census occupations into four categories: “literacy required”; “literacy likely to be useful”; “possible (or ambiguous) use of literacy”; and “unlikely to use literacy”. I group the first two and the last two together to create a binary variable, which is the primary dependent variable used in analysis. In Appendix Table G.1 I show the 12 most common jobs by literacy requirement in 1881.

I also make use of HISCLASS, a ranking of occupations by social class developed by Maas & Van Leeuwen (2016) using 19th century occupation data from Canada and Western Europe. It divides occupations into seven ordered social classes, listed in Table G.2.

Table G.1: **Common Jobs by Literacy Requirement**

Literacy Required or Likely to be Useful	Unlikely to Use Literacy
Commercial or Business Clerks	Coal Miner
Carpenter	General Labourer
Farm Owner/Steward	Agricultural Labourer
Grocer	Carman/Carrier/Carter
Blacksmith	Painter
Police	Cotton processor
Railway official	Bricklayer/Mason
Commercial traveler	Engine Stoker
Schoolmaster/teacher	Gardener
Innkeeper	Baker
Insurance agent	Coachman
Postman	Brick / tile / terra-cotta maker

Table G.2: **HISCLASS Categories**

Class Labels	Rank
Higher managers/professionals	1
Lower managers/professionals, clerical and sales positions	2
Foremen and skilled workers	3
Farmers and fishermen	4
Lower skilled workers	5
Unskilled workers	6
Lower and Unskilled farm labourers	7

H Appendix: Exposure to School Boards, 1901 vs 1881

Table H.1: Years of Treatment Difference by Board Arrival Year and Age

Age	1870	1871	1872	1873	1874	1875	1876	1877	1878	1879
16	1	2	3	4	5	6	7	8	8	8
17	2	3	4	5	6	7	8	8	8	8
18	3	4	5	6	7	8	8	8	8	8
19	4	5	6	7	8	8	8	8	8	8
20	5	6	7	8	8	8	8	8	8	8
21	6	7	8	8	8	8	8	8	8	8
22	7	8	8	8	8	8	8	8	8	8
23	8	8	8	8	8	8	8	8	8	8
24	8	8	8	8	8	8	8	8	8	8
25	8	8	8	8	8	8	8	8	8	8
26	8	8	8	8	8	8	8	8	8	8
27	8	8	8	8	8	8	8	8	8	7
28	8	8	8	8	8	8	8	8	7	6
29	8	8	8	8	8	8	8	7	6	5
30	8	8	8	8	8	8	7	6	5	4
31	8	8	8	8	8	7	6	5	4	3
32	8	8	8	8	7	6	5	4	3	2
33	8	8	8	7	6	5	4	3	2	1
34	8	8	7	6	5	4	3	2	1	0
35	8	7	6	5	4	3	2	1	0	0
36	7	6	5	4	3	2	1	0	0	0
37	6	5	4	3	2	1	0	0	0	0
38	5	4	3	2	1	0	0	0	0	0
39	4	3	2	1	0	0	0	0	0	0
40	3	2	1	0	0	0	0	0	0	0
41	2	1	0	0	0	0	0	0	0	0
42	1	0	0	0	0	0	0	0	0	0
43	0	0	0	0	0	0	0	0	0	0
44	0	0	0	0	0	0	0	0	0	0
45	0	0	0	0	0	0	0	0	0	0
46	0	0	0	0	0	0	0	0	0	0
47	0	0	0	0	0	0	0	0	0	0
48	0	0	0	0	0	0	0	0	0	0
49	0	0	0	0	0	0	0	0	0	0
50	0	0	0	0	0	0	0	0	0	0

This table documents the difference in years of exposure to board schools in 1901 vs 1881, by age and year of school board formation, assuming that exposure only occurs between ages 5-12 (inclusive), and that board schools were formed one year after the arrival of a school board. Example: an 18-year-old in 1901 residing in a parish that formed a school board in 1873 was exposed to board schools for 8 years (ages 5-12), while a 18-year-old in 1881 in the same parish was exposed to board schools for only 2 years (ages 11-12); thus the difference in years of exposure is 6.

I Appendix: Regression Kink Model

In the following description of the Regression Kink model, I rely heavily on the framework laid out in Dong (2018). Let $Y = y(T, R, W)$ represent the proportion of jobs within a parish that make use of literacy, where T is a binary treatment variable representing whether or not a parish received a school board, and R represents pre-reform school supply per capita. Importantly, treatment status T may be correlated with the error term W . Let $Y_t = y(t, R, W)$ for $t = 0, 1$, so Y_1 and Y_0 represent a parish's potential outcomes given treatment or not, respectively. Let treatment $T = 1_{\{P(R) - U \geq 0\}}$, where U is normalized such that $U \sim Unif(0, 1)$, so $P(R)$ is the probability of treatment given R . For notational simplicity, let $F_+ = \lim_{r \downarrow r_0} F(r)$ and $F_- = \lim_{r \uparrow r_0} F(r)$. Further, let $F'_+ = \lim_{r \downarrow r_0} \frac{\partial F(r)}{\partial r}$ and $F'_- = \lim_{r \uparrow r_0} \frac{\partial F(r)}{\partial r}$. Then the identification rests on the following assumptions:

Assumption 1 - *There exists a point $R = r_0$ such that $F'_+ \neq F'_-$.*

Assumption 2 - *$P(r)$ is continuously differentiable everywhere near r_0 except at r_0 , and is continuous at r_0 .*

Assumption 3 - *In the neighborhood of r_0 : (i) $y(0, R, W)$ and $y(1, R, W)$ are continuous, and differentiable w.r.t. R ; (ii) $\frac{\partial y(0, R, W)}{\partial R}$ and $\frac{\partial y(1, R, W)}{\partial R}$ are continuous.*

Assumption 4 - *In the neighborhood of r_0 , $f_{R, U|W=w}$ and $\frac{\partial f_{R, U|W=w}}{\partial R}$ are continuous.*

Let $G(r) = E[Y|R = r]$. If these assumptions hold, then:

$$\frac{G'_+ - G'_-}{F'_+ - F'_-} = E[Y_1 - Y_0 | U = P(r_0), R = r_0] = \int [y(1, r_0, w) - y(0, r_0, w)] \frac{f_{R, U|W=w}(r_0, P(r_0))}{f_{R, U}(r_0, P(r_0))} dF_W(w) = \tau \quad (14)$$

An intuitive proof is provided below. For a more detailed derivation, see Card et al. (2015).

Proof:

$$\begin{aligned} G(r) &= E[Y|R = r] = E[Y_1 * T | R = r] + E[Y_0 * (1 - T) | R = r] \\ &= E[(Y_1 - Y_0) * T | R = r] + E[Y_0 | R = r] \\ &= E[(Y_1 - Y_0) * 1_{\{U \leq P(r)\}} | R = r] + E[Y_0 | R = r] \end{aligned}$$

Thus, by Assumptions 2 and 3, for r in neighbourhood of r_0 but not equal to r_0 ,

$$\begin{aligned} \frac{\partial}{\partial r} G(r) &= P'(r) E[(Y_1 - Y_0) | U = P(r), R = r] \\ &+ \int_0^{P(r)} \frac{\partial}{\partial r} E[(Y_1 - Y_0) | U = u, R = r] du + \frac{\partial}{\partial r} E[Y_0 | R = r] \end{aligned}$$

Then,

$$G'_+ - G'_- = P'_+ E[(Y_1 - Y_0)|U = P_+, R = r_0] - P'_- E[(Y_1 - Y_0)|U = P_-, R = r_0] \\ + \int_0^{P_+} \frac{\partial}{\partial r} E[(Y_1 - Y_0)|U = u, R = r] du - \int_0^{P_-} \frac{\partial}{\partial r} E[(Y_1 - Y_0)|U = u, R = r] du$$

But by Assumption 2, $P_+ = P_- = P(r_0)$, so

$$\frac{G'_+ - G'_-}{P'_+ - P'_-} = E[Y_1 - Y_0|U = P(r_0), R = r_0] = \tau \quad (15)$$

This represents the average treatment effect at r_0 among those of type $U = P(r_0)$. These are the compliers: the parishes that would not be treated to the right of the kink ($T = 1_{\{P_+ - U \geq 0\}} = 0$) but would be treated to the left ($T = 1_{\{P_- - U \geq 0\}} = 1$). In the presence of heterogeneous treatment effects, this is somewhat restrictive. However, as pointed out in Lee & Lemieux (2010) in the RD context, it is less restrictive than one might assume, as the effect may be viewed as a weighted average of treatment effects at the kink across the entire population of unobservable types W . This is made clear in Card et al. (2015), which demonstrates that one may also write Equation 15 as

$$\tau = \int [y(1, r_0, w) - y(0, r_0, w)] \frac{f_{R,U|W=w}(r_0, P(r_0))}{f_{R,U}(r_0, P(r_0))} dF_W(w) \quad (16)$$

where the term $\int [y(1, r_0, w) - y(0, r_0, w)]$ represents the average treatment effect at the kink across all unobservable types W , which in turn is weighted by the probability a complier is of type $W = w$ (represented by the term $\frac{f_{R,U|W=w}(r_0, P(r_0))}{f_{R,U}(r_0, P(r_0))}$).

J Appendix: Running Variable Density Tests

I assess the smoothness of $f_R(r)$ using two tests: the McCrary test, and a test of differentiability at a point suggested by Card et al. (2012).

The McCrary test, common in the Regression Discontinuity literature, tests the continuity of the running variable at the cutoff by binning observations of the running variable, smoothing bin heights using local linear regression within some bandwidth, then comparing the heights of the bins to the left and right of the cutoff. For the test, I use the CCT optimal bandwidth calculated for the RK, $h = 0.064$. McCrary (2006) suggests that the test is robust to varying bin width b so long as $h/b > 10$. Nonetheless, I conduct the test twice using two different bin widths, $b = 0.003$ and $b = 0.006$. I report the results in the top window of Table J.1. For both bandwidths, the test fails to reject the null that $f_R(r)$ is continuous at the cutoff.

Card et al. (2012) suggests a test for the differentiability of the running variable at the cutoff. Similar to the McCrary test, this test bins observations of the running variable, then regresses the number of observations in each bin on polynomial terms of the running variable (centered at the kink) and the interaction between the centered running variable and the kink. The coefficient on the linear interaction term estimates the change in the first derivative at the cutoff. Its significance thus serves as a test of differentiability at the cutoff. I use a 2nd order polynomial, as the trend in $f_R(r)$ to the left of the cutoff, shown in Figure 5, suggests a quadratic. Again, I run the test twice using two different bin widths, $b = 0.003$ and $b = 0.006$. I again use the CCT optimal bandwidth calculated for the RK, $h = 0.064$. The coefficients on the linear interaction term are shown in the bottom window of Table J.1. In both specifications, the test fails to reject the null of no kink in $f_R(r)$ at the cutoff.

Table J.1: **Smoothness tests of 1871 School Supply density at the cutoff**

McCrary Test of Continuity	
Bin Width	Log Difference in Height
0.003	0.07031 (0.05427)
0.006	0.06398 (0.05428)
Card Test of Differentiability	
Bin Width	Estimated Kink
0.003	-510.86 (1438.96)
0.006	-618.88 (1323.86)

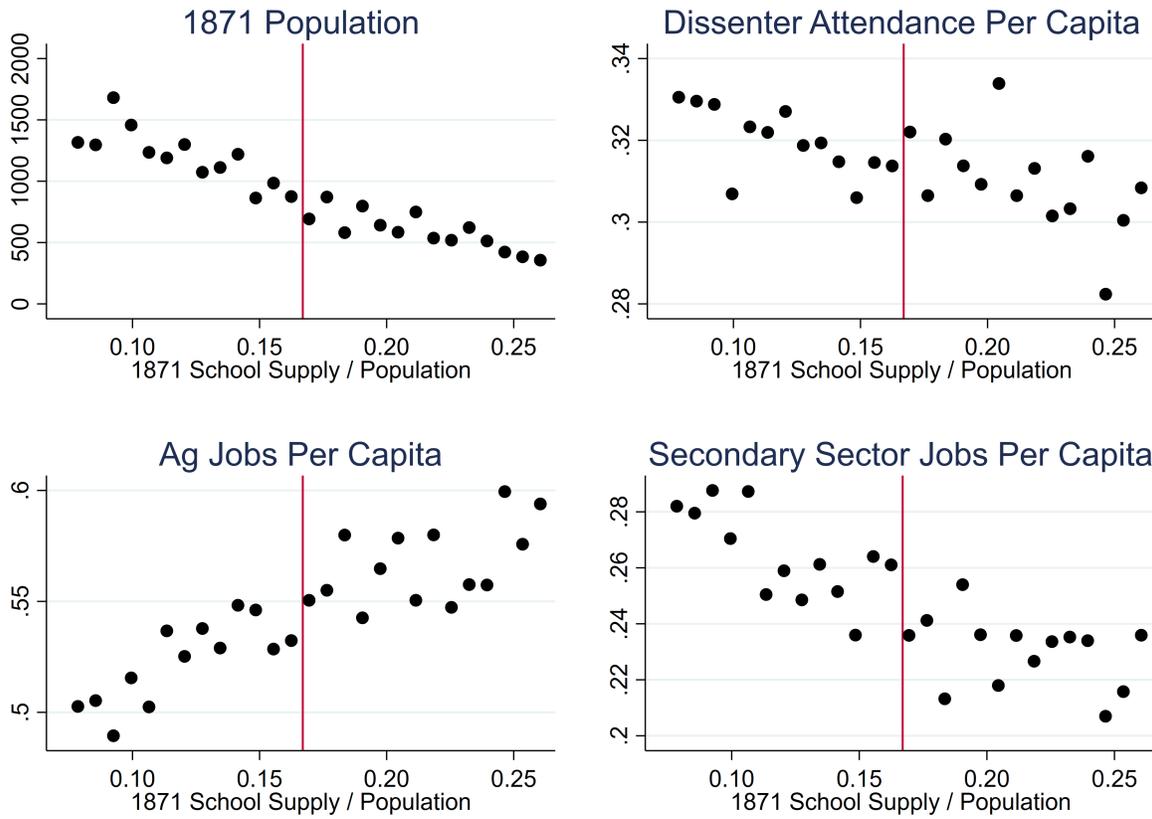
K Appendix: Regression Kink Covariates

One possible concern with the FRK design is that the kink in the relationship between pre-treatment school supply and school board formation, observed in Figure 2, may be driven by some kink in unobservables. If this were the case, one might expect to observe a kink

in other observable pre-treatment variables correlated with board formation. To test this, Figure K.4 plots the relationship between pre-treatment school supply and those variables shown in Table 1 to be correlated with treatment status: 1871 population, 1851 dissenting church attendance per capita, 1881 agricultural occupations per capita, and 1881 secondary sector occupations per capita. Note that 1881 occupation data, while not observed pre-treatment, may nonetheless be considered unaffected by the Acct, as the vast majority of adults at that time were untreated. As expected, no kink is observed in any of these variables at the cutoff, supporting the notion that the kink observed in board formation is indeed driven by the 1870 Act.

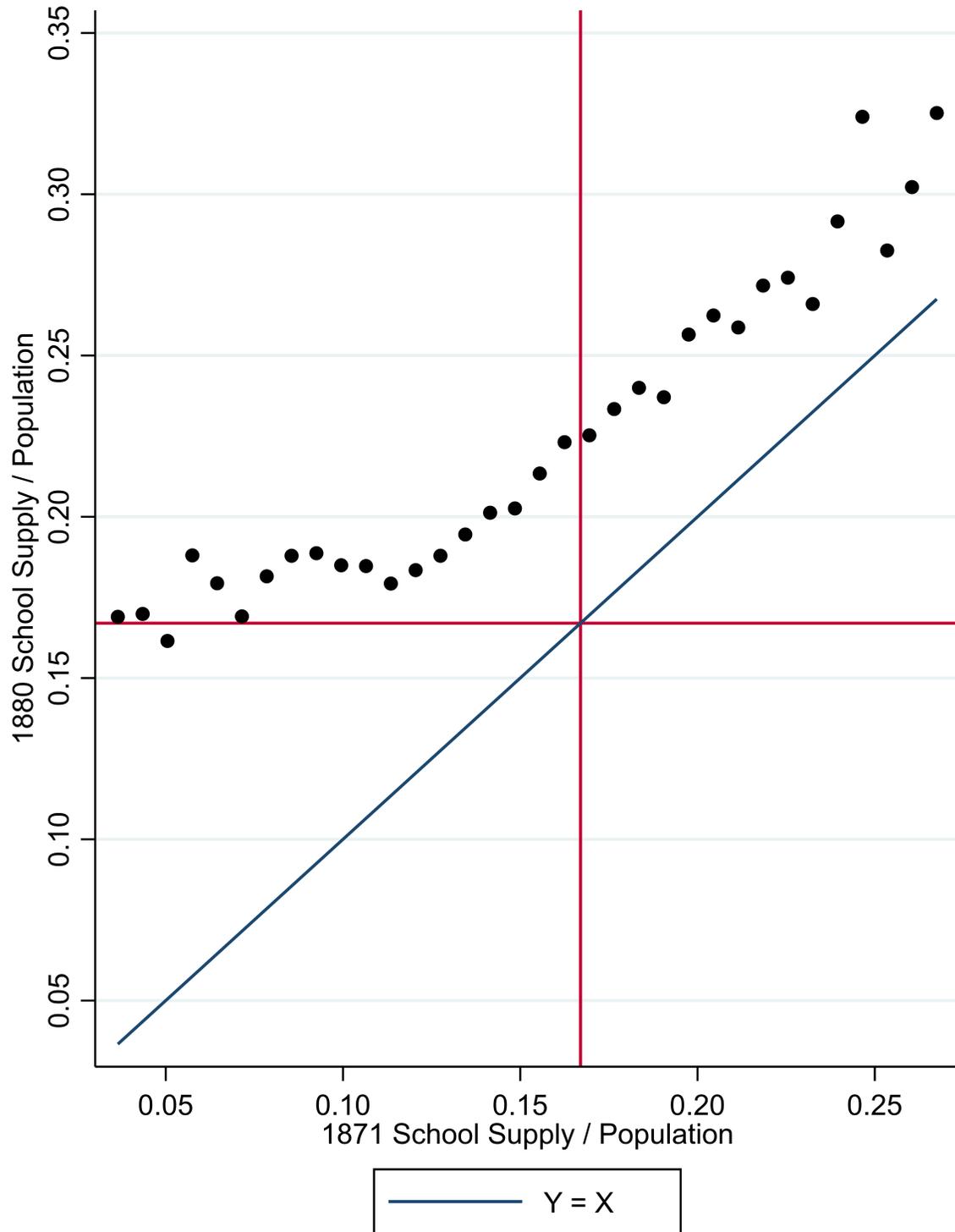
Nonetheless, as a precaution, the FRK analysis was conducted including controls as well. The results of this FRK analysis at the bandwidth $h = 0.064$ are provided in Table K.1. The results given in the first column, with no covariates included, are the same as those shown in Table 2. Column 2 includes as covariates those variables shown in Table 1 to be correlated with treatment status: 1871 population, 1851 dissenting church attendance per capita, 1881 agricultural occupations per capita, and 1881 secondary sector occupations per capita. Column 3 includes these same covariates, as well as 1880 school supply per capita. Note that the results in this final column should be viewed with caution: given that one of the key mandates of school boards was to increase local school supply, directly controlling for 1880 school supply may underestimate the effect of school boards.

Figure K.4: Covariates



Note: Parishes are binned by 1871 school supply to population ratio using a bin width of 0.07. In each case, the red line denotes the 0.166 cutoff. The top left figure plots the average parish 1871 population within each bin. The top right figure plots the average proportion of the local population within each bin that attended a church outside of the Church of England or Roman Catholic traditions according to the 1851 religious census. The bottom left figure plots the average proportion of men employed in agricultural occupations in 1881 within each bin. The bottom right figure plots the average proportion of men employed in secondary sector occupation in 1881 within each bin.

Figure K.5: School Supply Per Capita, 1871 vs 1880



Note: Parishes are binned by 1871 school supply to population ratio using a bin width of 0.07. Figure 5 plots the 1871 parish school supply to population ratio against the average 1880 parish school supply to population ratio (calculated using population figures from the 1881 Census) within each bin. The red lines denote the 0.166 cutoff on both the x and y axes, while the blue line plots the equivalence relationship between 1880 and 1871 school supply per capita.

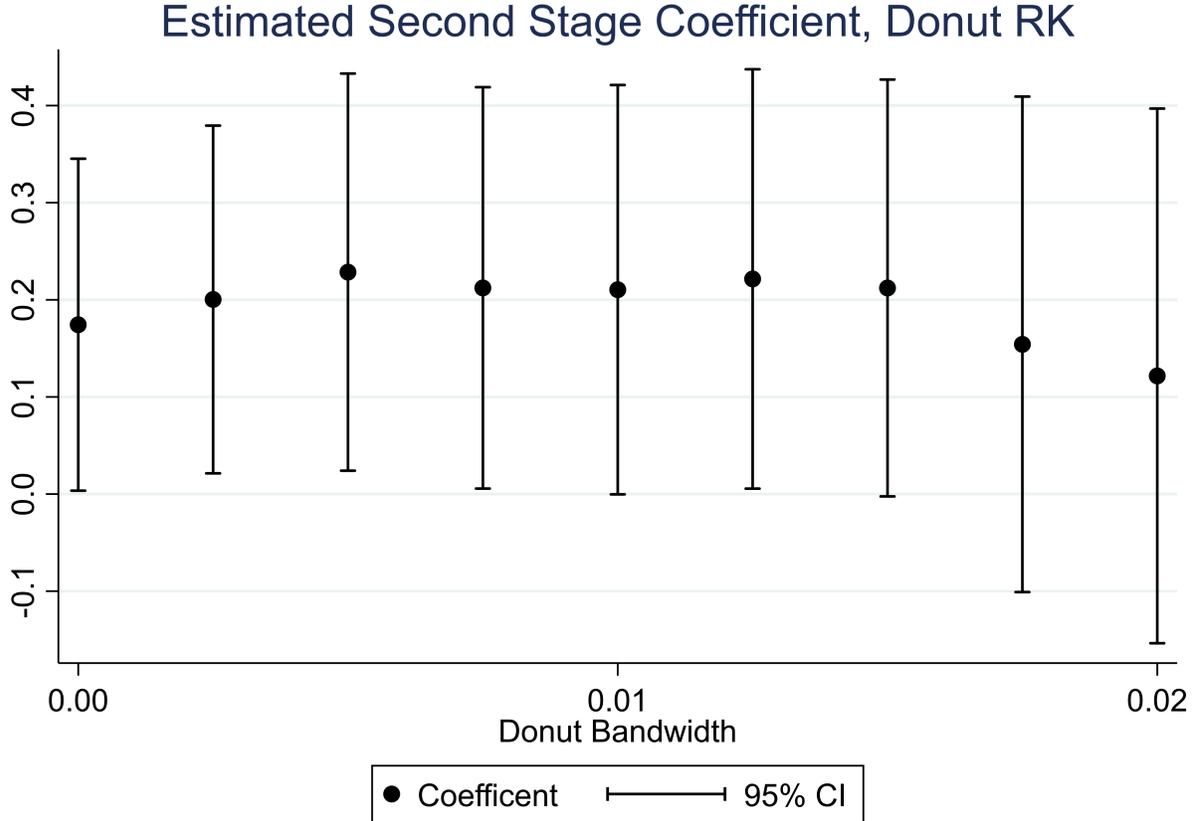
Table K.1: **RK 2SLS with Covariates**

First Stage: School Board			
1871 School Supply/Pop	-2.8753*** (0.2410)	-2.6444*** (0.2373)	-2.5977*** (0.2850)
(Cutoff Dummy)*(1871 School Supply/Pop - 0.1666)	2.7606*** (0.5007)	2.6141*** (0.4917)	2.8451*** (0.5883)
Prop. of Jobs Requiring Literacy, ages 19-30, 1881		0.0754** (0.0333)	0.1077*** (0.0420)
1880 School Supply/Pop			-0.2681*** (0.0643)
Second Stage: Proportion of jobs requiring literacy, ages 19-30, 1901			
School Board	0.1744** (0.0872)	0.1470* (0.0816)	0.1248 (0.0842)
1871 School Supply/Pop	0.2308 (0.1649)	0.2446* (0.1402)	0.1725 (0.1390)
Prop. of Jobs Requiring Literacy, ages 19-30, 1881		0.2578*** (0.0158)	0.2777*** (0.0195)
1880 School Supply/Pop			0.0479 (0.0343)
Other Controls	NO	YES	YES
Parishes	5979	5887	4197
Bandwidth	0.064	0.064	0.064
Instrument F-Stat	30.39	28.26	23.39

Cutoff Dummy = $1_{[1871 \text{ School Supply/Pop} > 0.1666]}$. Regressions are conducted at the parish level. The second stage estimates the increase in the proportion of men aged 19-30 in 1901 employed in occupations requiring literacy due to the formation of a school board in the parish. "Other Controls" are those variables shown in Table 1 to be correlated with treatment status: 1871 population, 1851 dissenting church attendance per capita, 1881 agricultural occupations per capita, and 1881 secondary sector occupations per capita. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

L Appendix: Donut Regression Kink Design

Figure L.6: RK Donut Analysis (Exclusion around kink)



Note: Figure L.6 plots for varying inner bandwidths the second stage estimates from the regression kink donut analysis. Only observations laying within an outer bandwidth but outside an inner bandwidth around the cutoff are used in the estimation. In all cases, the optimal bandwidth $h = 0.064$ is used as the outer bandwidth. Estimates using the following inner bandwidths are plotted: $h = 0 + c * (0.0025)$, where $c \in \{0, 1, 2, 3, 4, 5, 6, 7, 8\}$.

M Appendix: DDD Robustness Checks

Note that using TY_{tra} relies on the assumption that each additional year residing in a parish with a board school between the ages of 5-12 had on average the same effect. In truth, there is reason to doubt this. For one, at the older end of the 5-12 age range attendance was less frequent due to labour market opportunities, and thus having a board school could have made less of a difference at these ages. Further, even if attendance were steady across ages, it is plausible that education showed either diminishing or increasing returns. To test if this assumption biases my results, I replace years of treatment TY_{tra} with a full treatment dummy by dropping all those partially treated, for whom $TY_{tra} \notin \{0, 8\}$. This drops all individuals whose parish-age groups were only partially treated in either 1881 or 1901, meaning they had access to public schools for some of the years from 5 to 12 but not all of them. This

allows me to use the specification described in Equation 8, the results of which are reported in Table M.1. Note that the treatment coefficient now represents the effect of a full 8 years of public schooling. As before, full treatment improves the chance of obtaining a literate occupation by around 1 pp.

To ensure outliers are not driving results, I drop individuals residing in the top and bottom ventile in terms of pre-treatment school supply per capita and rerun the regressions of Table 4. I then do the same for 1871 population. The results are shown in Table M.2 and demonstrate that dropping the tails has little impact on the estimated treatment effect.

Finally, I allow the effect of pre-treatment school supply on literate occupations to vary not only by year and side of the $r_0 = 0.166$ cutoff, but by also by age as well:

$$Y_{itpa} = \mu_{ra} + \lambda_{rt} + \alpha_{at} + \gamma_7(TY_{tra}) + \eta_p + (Z_p * t)\Omega + \phi_{1at}(R_p) + \phi_{2at}[1_{[R \geq r_0]} * (R_p - r_0)] + u_{itpa} \quad (17)$$

where R_p represents pre-treatment parish school supply. The results are shown in Table M.3; the estimated treatment effects remain largely unchanged.

Table M.1: **Effect of Full Treatment, Dropping Partially Treated**

VARIABLES	(1) Literacy Required	(2) Literacy Required	(3) Literacy Required
Treatment	0.0124 (0.00782)	0.0108*** (0.00369)	0.0104*** (0.00374)
Observations	3,509,823	3,509,817	3,479,801
Controls	NO	NO	YES
Parish Fixed Effects	NO	YES	YES
Arrival-Age Fixed Effects	YES	YES	YES
Age Trend Effects	YES	YES	YES
Arrival Trend Effects	YES	YES	YES

Robust standard errors in parentheses, clustered at the parish level. Unit of observation is the individual. Controls (all interacted with a 1901 dummy): 1871 parish population; 1871 parish population squared; distance from London; distance from London squared; 1871 school supply per capita; 1851 Church of England attendees per capita 1851 Catholic attendees per capita; 1851 other religious service attendees per capita. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table M.2: **Full Sample Triple Difference, Tails Dropped**

VARIABLES	School Tails Dropped		Pop Tails Dropped	
	Years of Treatment	0.00116** (0.000458)	0.00115** (0.000464)	0.00128** (0.000538)
Observations	3,515,470	3,490,956	2,127,948	2,116,353
Controls	NO	YES	NO	YES
Parish Fixed Effects	YES	YES	YES	YES
Arrival-Age Fixed Effects	YES	YES	YES	YES
Age Trend Effects	YES	YES	YES	YES
Arrival Trend Effects	YES	YES	YES	YES

Robust standard errors in parentheses, clustered at the parish level. Unit of observation is the individual. Controls (all interacted with a 1901 dummy): 1871 parish population; 1871 parish population squared; distance from London; distance from London squared; 1871 school supply per capita; 1851 Church of England attendees per capita 1851 Catholic attendees per capita; 1851 other religious service attendees per capita. *** p<0.01, ** p<0.05, * p<0.1

Table M.3: **Full Sample Triple Difference, Controlling for Interactions Between Age, Year, School Supply, and FRK Cutoff**

	(1) Literacy Required	(2) Literacy Required	(3) Literacy Required
Years of Treatment	0.00139 (0.00103)	0.00119*** (0.000460)	0.00115** (0.000465)
Observations	3,896,995	3,896,989	3,864,788
Other Controls	NO	NO	YES
Parish Fixed Effects	NO	YES	YES
Arrival-Age Fixed Effects	YES	YES	YES
Age Trend Effects	YES	YES	YES
Arrival Trend Effects	YES	YES	YES

Robust standard errors in parentheses, clustered at the parish level. Unit of observation is the individual. All specifications include the full set of interactions between age, year, and parish 1871 school supply, as well as between age, year, parish 1871 school supply centered at the cutoff used in FRK analysis (0.1666), and the cutoff. Other controls (all interacted with a 1901 dummy): 1871 parish population; 1871 parish population squared; distance from London; distance from London squared; 1851 Church of England attendees per capita 1851 Catholic attendees per capita; 1851 other religious service attendees per capita. *** p<0.01, ** p<0.05, * p<0.1

N Appendix: Differences between FRK and DDD estimates: Disproven Theories

One possible explanation why the LATE measured by the FRK is much larger than the DDD ATT is that treatment effects were much larger in parishes with pre-treatment school supply near the kink. To test this, I run the triple difference regression on the sample of individual within the optimal bandwidth used in the RK. The results are shown in Table N.1. The estimated treatment effects are very similar to those shown in Table 3, suggesting that the use of parishes near the kink is not driving the difference.

Another possibility is that the difference is the result of using different levels of observation: in the FRK the parish, while in the DDD the individual. In the FRK, the use of the parish seems most appropriate, as that is the level at which the first stage treatment decision occurs. However, for comparability, I conduct the FRK using the individual as the unit of analysis, clustering errors at the parish level. I show the results in Figure N.7. The treatment effect is actually even larger when using the optimal bandwidth of 0.064 - around 24.7 pp as opposed to 17.6 pp. Thus, it seems that level of observation is also not the explanation for the differences.

A third possibility is that the DDD is underestimating the effect due to a violation in its “no spillovers” assumption. As discussed above, such a violation is unlikely for various reasons, but cannot be ruled out. Imagine, for instance, that the introduction of public schools drew human capital intensive industries to treated parishes, which in turn increased the probability of employment in literate occupations for the entire parish population, even those too old to attend the public schools. This would bias the DDD estimate downward. To directly test for spillovers across age groups, I run a simple difference-in-differences on only those over 35, all of whom were too old to be fully treated. Under the “no spillovers” assumption, one would expect the treatment effect estimated by the DID interaction term to be close to zero. The results are shown in Table N.2. As expected, the treatment effect is never significant, and is essentially zero when parish fixed effects are included, supporting the “no spillovers” assumption.

Figure N.7: Population Weighted RK

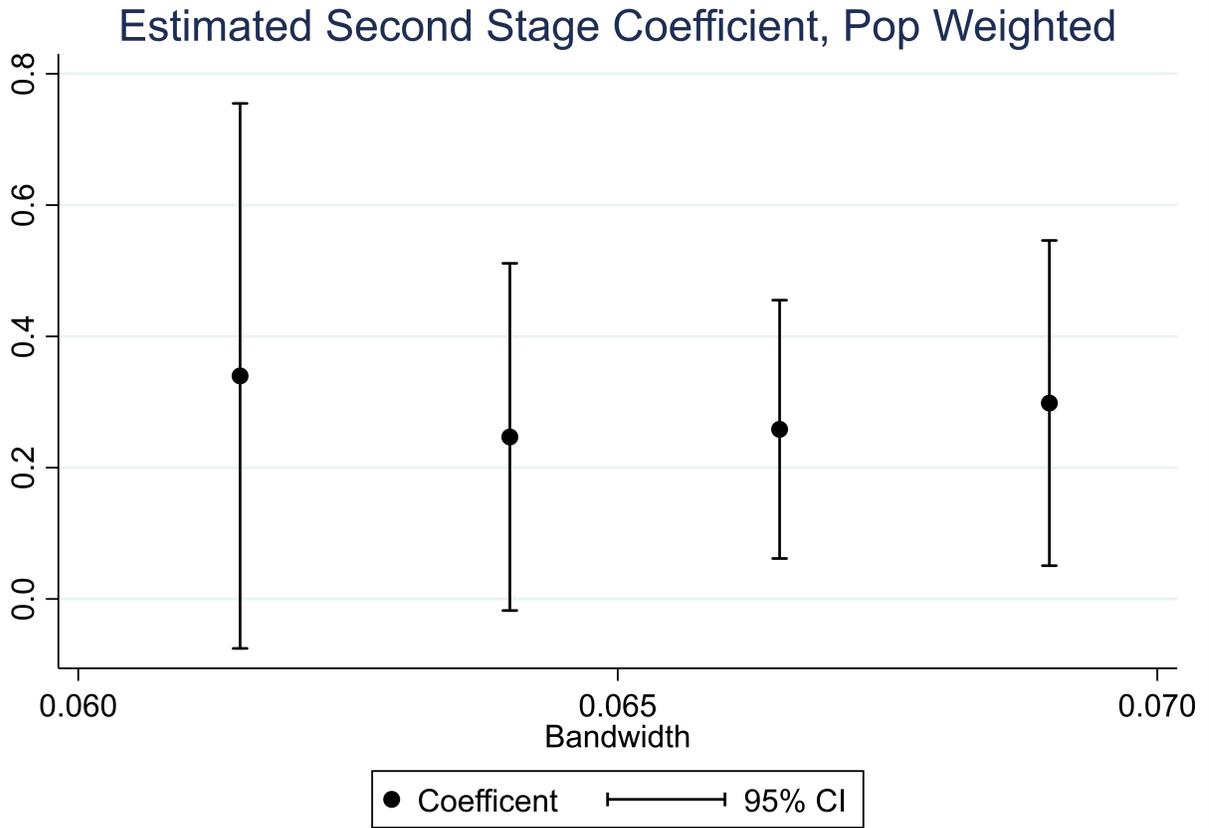


Table N.1: Triple Diff Within FRK Bandwidth

	(1)	(2)	(3)
	Literacy Required	Literacy Required	Literacy Required
Years of Treatment	0.00257** (0.00112)	0.00146** (0.000723)	0.00139* (0.000727)
Observations	1,854,485	1,854,484	1,837,820
Controls	NO	NO	YES
Parish Fixed Effects	NO	YES	YES
Arrival-Age Fixed Effects	YES	YES	YES
Age Trend Effects	YES	YES	YES
Arrival Trend Effects	YES	YES	YES

Robust standard errors in parentheses, clustered at the parish level. Controls (all interacted with a 1901 dummy): 1871 parish population; 1871 parish population squared; distance from London; distance from London squared; 1871 school supply per capita; 1851 Church of England attendees per capita 1851 Catholic attendees per capita; 1851 other religious service attendees per capita. *** p<0.01, ** p<0.05, * p<0.1

Table N.2: **Difference-in-differences for those over 35**

VARIABLES	(1) Literacy Required	(2) Literacy Required
(Treatment)*(1901 Year Dummy)	-0.00177 (0.00358)	0.000399 (0.00374)
Observations	1,220,922	1,211,168
Controls	NO	YES
Parish Fixed Effects	YES	YES
Arrival-Age Fixed Effects	YES	YES
Age Trend Effects	YES	YES

Robust standard errors in parentheses, clustered at the parish level. Controls (all interacted with a 1901 dummy): 1871 parish population; 1871 parish population squared; distance from London; distance from London squared; 1871 school supply per capita; 1851 Church of England attendees per capita 1851 Catholic attendees per capita; 1851 other religious service attendees per capita.
 *** p<0.01, ** p<0.05, * p<0.1

O Appendix: Census Linking Procedure

The method used for both the 1861-1881 and 1881-1901 census linkages is based on that conducted in a series of papers by Abramitzky, Boustan, and Eriksson (2012, 2014, 2019a). The main area of departure is how I use birthplace variables. The Abramitzky et al. method utilizes only a single birthplace variable. In the UK censuses, however, there are three possible birthplace variables to use, each with their own pros and cons. Likely the most accurate is the county of birth. However, this is also the least specific, and thus results in fewer unique matches. A second variable is parish of birth as reported. Obviously this is much more specific than county. However, parishes very often had several acceptable names, and therefore if one name was used in the initial census and another in the later census, this match would be missed. To deal with this, the Integrated Census Microdata project, where the data was received from, created a “standardized” parish of birth variable. However, such a standardization consists of some guesswork, and thus undoubtedly contains errors.⁵³ To mitigate the individual shortcomings of each of these variables, I make use of all three, as described below in the step-by-step procedure used to link the 1861-1881 censuses and the 1881-1901 censuses:

Begin with universe of men aged 5-25 in the early census (either 1861 or 1881, depending on linkage), and the universe of men aged 25-45 in the late census (either 1881 or 1901).

1. Use code provided by Abramitzky et al. to clean names. This removes non-alphabetic characters and also accounts for common nicknames so that, for instance, ‘Ben’ and ‘Benjamin’ would match.
2. For each birth place variable (county, parish as reported, and standardized parish), run the following steps:
 - (a) For each record in the early census, create a group of possible matches with records from the late census. Possible matches are defined as any records that share the same birthplace, first and last initials, and whose birth years are within two years of each other.
 - (b) For each possible match, calculate Jaro-Winkler similarity scores for first and last names, separately. This is a measure of similarity between two strings, with more weight placed on characters at the beginning of the strings. It is equal to 1 if the strings are identical and 0 if there is no similarity. Keep only those matches for which first name and last name Jaro-Winkler scores are both 0.9 or above.
3. Run the following steps on the surviving matches determined using county birthplace:
 - (i) For each record in the early census, determine the smallest birth year difference

⁵³See “Integrated Census Microdata (I-CeM) Guide”, pg. 117,207

between it and its potential matches in the late census.⁵⁴ Similarly, for each record in the late census, determine the smallest birth year difference between it and its potential matches in the early census.

- (ii) Keep only those potential matches for which the birth year difference is equal to both the smallest birth year difference for the early side of the match and the smallest birth year difference for the late side of the match.
 - (iii) For each remaining potential match, keep only if unique for both the early and late sides of the match.
4. Combine the county matches remaining after Step 3 with the lists of matches determined using standardized parish birthplace and parish birthplace as recorded. Drop duplicate matches. For each county match, keep only when both the early and late side of the match do not appear among any potential parish matches.
 5. Rerun Steps (i)-(iii) on the remaining potential matches.

⁵⁴For example, if an individual was born in year x and any of its remaining potential matches were born in x , its smallest birth year difference would be 0; if, however, none of its remaining potential matches were born in x but some were born in either $x-1$ or $x+1$, its smallest birth year difference would be 1.

P Appendix: Linkage Summary Statistics

Table P.1: Linkage Summary Statistics: Early Periods

1861-1881 Linkage Summary Statistics: 1861				
	(1)	(2)	(3)	(4)
VARIABLES	Full	Matched to School Data	Linked & Matched to School Data	Linked & Resides with Father & Matched to School Data
Live in treated (in future) parish		0.298	0.284	0.284
Age	14.25	13.99	13.61	12.02
Population of parish	12,107	4,868	4,282	4,306
Resides with Father	0.594	0.612	0.664	
Father Class Dummy	0.356	0.279	0.284	
Linked to 1881	0.371	0.409		
Matched to School Data	0.510			
Observations	4,102,162	2,093,087	855,591	568,107
1881-1901 Linkage Summary Statistics: 1881				
	(1)	(2)	(3)	(4)
VARIABLES	Full	Matched to School Data	Linked & Matched to School Data	Linked & Resides with Father & Matched to School Data
Live in treated parish		0.334	0.318	0.320
Age	14.19	13.99	13.83	12.32
Population of parish	21,665	9,789	8,552	8,485
Resides with Father	0.643	0.661	0.707	
Father Class Dummy	0.396	0.312	0.312	
Linked to 1901	0.422	0.468		
Matched to School Data	0.521			
Observations	5,591,747	2,915,615	1,365,184	964,760

Note: The top panel describes the summary statistics of males aged 5-25 in the 1861 census. The bottom panel describes males aged 5-25 in the 1881 census. Column 1 provides summary statistics for the universe of men in England and Wales, as described in the full-count censuses. Columns 2 and 3 compare the subset of the population matched to school data with the corresponding linked sample. Column 4 shows summary statistics of the subset of linked individuals matched to school data who resided with their father in the earlier census, enabling measurement of childhood class (using father's occupation).

Table P.2: **Linkage Summary Statistics: Late Periods**

1861-1881 Linkage Summary Statistics: 1881				
	(1)	(2)	(3)	(4)
VARIABLES	Full	Linked	Linked & Matched to School Data	Linked & Resided with Father & Matched to School Data
Linked to 1861	0.382			
Age	33.54	33.60	33.48	31.89
Class Dummy	0.415	0.421	0.366	0.367
Resided with Father Matched to School Data		0.655	0.664	
Observations	4,043,487	1,522,047	855,591	568,107
1881-1901 Linkage Summary Statistics: 1901				
	(1)	(2)	(3)	(4)
VARIABLES	Full	Linked	Linked & Matched to School Data	Linked & Resided with Father & Matched to School Data
Linked to 1881	0.531			
Age	34.06	33.87	33.74	32.24
Class Dummy	0.436	0.438	0.385	0.383
Resided with Father Matched to School Data		0.692	0.707	
Observations	4,444,271	2,357,948	1,365,184	964,760

Note: The top panel describes the summary statistics of males aged 25-45 in the 1881 census. Columns 1 and 2 provide the summary statistics of the universe of men recorded in the census and the subset successfully linked across censuses, respectively. Column 3 examines the subset successfully linked and matched to school data. Column 4 shows summary statistics of the subset of linked individuals matched to school data who resided with their father in the earlier census, enabling measurement of childhood class (using father's occupation).

Q Appendix: Middle Initial Test for False Positives

To test the rate of false positives, I make use of middle initials. Given that middle initial was not used in the matching process (as it is only available for a subset of individuals), a middle initial shared by both sides of a match is a strong indicator that it is indeed a true match. While the majority of records do not provide middle names, it is provided by both the early and late side of 205,476 of the 1861-1881 matches (12.9%) and 469,071 of the 1881-1901 matches (19%). Among these, 79.2% of the 1861-1881 matches and 90.2% of the 1881-1901 matches have the same middle initial on both sides of the match.

It is likely of course that these numbers do not perfectly reflect the distribution of true matches in the linkages. On the one hand, even among matches with identical middle initials there are no doubt some false matches. Assuming the middle initials on either side of a false match follow the distribution of middle initials found in the population, the estimated probability of two individuals incorrectly matched sharing the same middle initial, or $\hat{P}(\text{MI Match} \mid \text{False Match})$, is 8.6% in the 1861-1881 linkage and 9.4% in the 1881-1901 linkage.

On the other hand, it is likely that among many true matches, transcription error resulted in mismatching middle initials. Take, for example, the initial “K”. Among matches where one side of the match has initial “K” while the other side has a different letter, the other side’s initial is a similarly shaped letter “H” or “R” 32.1% and 31.1% of the time, respectively, while given the population distribution of middle initials one would expect them to appear only 17.1% and 4.6% of the time, respectively. To pick out which letter combinations were likely common transcription errors, I compared each combinations (a) frequency in the data with (b) how often one would expect it to occur if each letter were randomly drawn from the population distribution of middle initials. Assuming the true probability of the letter combination occurring was (b), if the probability of observing frequency (a) in the data is less than 0.000001 then I treat the occurrences in excess of what was expected as transcription errors.

Table Q.1 shows the letter combinations meeting this criteria.⁵⁵ The included combinations rarely if ever surprise; in general they are what one would expect as a result of transcription error.

I can now estimate $\hat{P}(\text{False Link} \mid \text{No MI Match})$, taking the the number of mismatched middle initials less the estimated number of transcription errors, and dividing it by the total number of mismatched middle initials. Doing this, it is estimated that only 56.5% of links

⁵⁵Note that this criteria was applied separately to the 1861-1881 and the 1881-1901 linkage. Due to the larger sample size in the later linkage, there are a number of letter combinations that only meet the criteria there.

with mismatched middle initials in the 1861-1881 linkage are actually false, and 60.4% in the 1881-1901 census.

Using $\hat{P}(\text{MI Match} \mid \text{False Match})$ and $\hat{P}(\text{False Link} \mid \text{No MI Match})$, along with the proportion of mismatching middle initials $\hat{P}(\text{MI Match})$, I can estimate the true match rate:

$$\begin{aligned}
 P(\text{True Link}) &= 1 - P(\text{False Link}) \\
 &= 1 - \frac{P(\text{False Link} \cap \text{No MI Match})}{P(\text{No MI Match})} \\
 &= 1 - \frac{P(\text{False Link} \mid \text{No MI Match})P(\text{No MI Match})}{1 - P(\text{MI Match} \mid \text{False Link})}
 \end{aligned} \tag{18}$$

The results, reported in Table Q.2, suggest that 87.2% of the 1861-1881 links and 93.5% of the 1881-1901 links are true. The literature suggests that these are very good: Bailey et al. (2019) state that common match procedures lead to PPVs (Positive Prediction Value, as the ratio of true matches to total matches is known in the machine learning literature) between 66% and 90%, while Abramitzky et al. (2019b) are slightly more optimistic, arguing that modern procedures lead to PPVs between 70% and 95%. The high PPVs found here are likely driven by the use of full-count censuses. As pointed out in Bailey et al. (2019), a match that appears unique when linking a subsample may not be when linking the full population. Such a match, which has a high probability of being a false positive, would be included in the subsample linkage but excluded when linking full-count samples, leading to lower false positive rates in full-count linkages.

Table Q.1: **Overrepresented Combinations among Differing Middle Initials**

A	D, H, O, R	I	F*, J, S*, T	R	A, B, K, N, P
B	D*, P*, R	J	F, G, I, L, S, T	S	G*, I*, J, L, P, T
C	E, G, L, O	K	H, R, W [†]	T	F, I, J, L, P*, S, Y*
D	A, B*, L*, O	L	C, D*, F*, J, S, T	U	N*, V, W
E	C, G	M	H*, N, W	V	N*, O, P*, U
F	H*, I*, J, L*, P*, T	N	H, M, R, U*, V*, W	W	H, K [†] , M, N, U
G	C, E, J, S*, Y	O	A, C, D, V	Y	G, T*
H	A, F*, K, M*, N, W	P	B*, F*, R, S, T*, V*		

[†] Only identified in and applied to 1861-1881 linkage

* Only identified in and applied to 1881-1901 linkage

Table Q.2: Estimated Linking Characteristics

	$\hat{P}(\text{No MI Match})$	$\hat{P}(\text{MI Match} \mid \text{False Link})$	$\hat{P}(\text{False Link} \mid \text{No MI Match})$	$\hat{\mathbf{P}}(\text{True Link})$
1861-1881 Linkages				
Standard	0.208	0.086	0.565	0.872
Unique in 5-Year Band, Age Exact	0.171	0.084	0.451	0.916
Unique in 5-Year Band, Age & Name Exact	0.165	0.084	0.447	0.919
1881-1901 Linkages				
Standard	0.098	0.094	0.604	0.935
Unique in 5-Year Band, Age Exact	0.071	0.091	0.442	0.965
Unique in 5-Year Band, Age & Name Exact	0.069	0.092	0.428	0.968

R Appendix: Comparison With Notable Linkages in the Literature

Table R.1: Comparison of Linkages

Paper	Sources	Match Rate	Number Linked
Milner (2019)	1861 England & Wales Census (Full, Men 5-25) to 1881 England & Wales Census (Full, Men 25-45)	37.1%	1,522,047
Milner (2019)	1881 England & Wales Census (Full, Men 5-25) to 1901 England & Wales Census (Full, Men 25-45)	42.2%	2,357,948
Long & Ferrie (2013)	1881 England & Wales Census (2% Sample, Men 0-25) to 1881 England & Wales Census (Full, Men)	20.3%	14,191
Feigenbaum (2015)	1915 Iowa Census (Golden & Katz (2000, 2008) Sample, Men 3-17) to 1940 US Census (Full, Men)	57.4%	4,349
Abramitzky et al. (2012)	1865 Norwegian Census (Full, Men 3-15) to 1900 Norwegian Census (Full, Men) or 1900 Roster of Norwegians Immigrants in US (Full, Men)	7.3%	20,446
Abramitzky et al. (2014)	1900 US Census (Subsample of white native & European born men 18-35) to 1910 US Census (Full, Men) and 1920 US Census (Full, Men)	Native Born: 16.5% Immigrant: 8.2%	1,650 20,218
Baker et al. (2018)	1940 US Census (Full, Men born in South 23-58) to 1900, 1910, or 1920 US Census (in each case Full, Men 3-18)	White: 27.5% Black: 18.6%	432,235 170,923

S Appendix: More Conservative Matching Procedures

Evidence presented in Appendix R strongly suggests that the method used to link censuses in this paper (described in detail in Appendix O) generated very few false positives. Nonetheless, to ensure that these few false positives are not driving results, the mobility regressions are rerun on linkages constructed using more conservative methods. Both of the conservative methods used create subsets of the baseline linkage. These methods, while likely omitting many of the baseline method’s true matches, should lead to fewer false matches.

The first new method matches more strictly on age, while the second matches more strictly on both age and name, creating a subset of the first. Both methods require that each match’s early and the late records have identical ages and have no other possible matches within the five year window created by adding or subtracting two from age. In contrast, the baseline match simply requires each match’s early and the late records be within two years of each other and have no other possible matches equal to or closer in age. The strictest method also demands that the early and late records within each match have identical first and last names, while the baseline and other conservative method only require both the first and last name Jaro-Winkler similarity scores to be 0.9 or above.

As reported in Table Q.2, it is estimated that the baseline method results in true match rates of 87.2% and 93.5% for the 1861-1881 and 1881-1901 linkages, respectively. In contrast, the estimated true match rates using the age-conservative method are 91.6% and 96.5%, and using the name-and-age-conservative method are 91.9% and 96.8%.

Table S.1 shows the results of rerunning the mobility regressions using the more conservative linkages. In all regressions, both public school access’s absolute effect on lower class children’s outcomes and the difference between this and its effect on higher class children are larger than those found using the baseline linkage (shown in Table 7). While it should be borne in mind that the more conservative linkages may be less representative of the total population than the baseline model, these results nonetheless suggest that if anything the results reported in the paper should be taken as a lower bound.

Table S.1: **Social Mobility Triple Difference, Matches Unique Within 5-year Band**

	Perfect Matches on Age		Perfect Matches on Name & Age	
	(1)	(2)	(3)	(4)
Years of Treatment	0.00234** (0.00102)	0.00494** (0.00221)	0.00227** (0.00111)	0.00596** (0.00256)
(Years of Treatment)*(Father Class Dummy)	-0.00407** (0.00192)	-0.00353 (0.00411)	-0.00557*** (0.00208)	-0.00532 (0.00457)
Observations	768,988	301,498	655,729	247,857
Controls Included	YES	N.A.	YES	N.A.
Birth Parish-Class Fixed Effects Included	YES	N.A.	YES	N.A.
Arrival-Age-Class Fixed Effects Included	YES	YES	YES	YES
Age-Class Trend Effects Included	YES	YES	YES	YES
Arrival-Class Trend Effects Included	YES	N.A.	YES	N.A.
Brother Fixed Effects Included	NO	YES	NO	YES

Robust standard errors in parentheses, clustered at parish level. Controls (all interacted with a 1901 dummy, and including interactions with father's class dummy): 1871 parish population; 1871 parish population squared; distance from London; distance from London squared; 1871 school supply per capita; 1851 Church of England attendees per capita 1851 Catholic attendees per capita; 1851 other religious service attendees per capita. *** p<0.01, ** p<0.05, * p<0.1

T Appendix: Alternative Mobility Measures

In the paper, occupations that make use of literacy are used as a proxy for good, high paying work in the mobility regressions. This is made necessary by the fact that wage data from the period is scarce. However, other income proxies do exist.

One such proxy is OCCSCORE. This records the median total income by occupation in the 1950 US Census, and is commonly used as an proxy for US wages in earlier periods. Admittedly, the difference in both time and place in this case suggest a somewhat weak relationship with true income. But its prevalence in the literature merits its usage, if only as a source of comparison. As demonstrated in the first two columns of Table T.1, the mobility results presented in the paper are robust to replacing the literate occupation dummy with $\ln(\text{OCCSCORE})$ or percentile as determined using OCCSCORE.

The other proxy tested is a ranking of occupations based on the prospects of practitioners' sons. The core assumption used in its construction is that sons of better employed fathers are more likely to end up better employed themselves. To create this ranking I used the following recursive procedure:

1. For each occupation, determine the proportion of practitioners' sons who in adulthood are practitioners of an occupation in one of the top two categories of the seven category HISCLASS variable.
2. Assign each occupation a percentile based on these proportions.
3. Using these percentiles, determine for each occupation the expected percentile of a practitioner's son's adult occupation.
4. Redetermine each occupation's percentile based on these expected percentiles.
5. Repeat Steps 3-4 until the ranking of occupations stabilizes.

Using the percentiles generated by this procedure, I again estimate the effect of public school access on lower and higher class children. Once again, the results affirm those shown in the paper.

Table T.1: Social Mobility Triple Difference, Alternative Measures

VARIABLES	(1) ln(OCCSCORE)	(2) OCCSCORE Percentile	(3) Recursive Percentile
Years of Treatment	0.00181** (0.000828)	0.104* (0.0539)	0.0820** (0.0398)
(Years of Treatment)*(Father Class Dummy)	-0.00288** (0.00123)	-0.152* (0.0867)	-0.0363 (0.0731)
Observations	1,096,626	1,096,970	1,411,146
Controls Included	YES	YES	YES
Birth Parish-Class Fixed Effects Included	YES	YES	YES
Arrival-Age-Class Fixed Effects Included	YES	YES	YES
Age-Class Trend Effects Included	YES	YES	YES
Arrival-Class Trend Effects Included	YES	YES	YES

Robust standard errors in parentheses, clustered at parish level. Controls (all interacted with a 1901 dummy, and including interactions with father's class dummy): 1871 parish population; 1871 parish population squared; distance from London; distance from London squared; 1871 school supply per capita; 1851 Church of England attendees per capita 1851 Catholic attendees per capita; 1851 other religious service attendees per capita. *** p<0.01, ** p<0.05, * p<0.1

U Appendix: Advantage of Higher Class Children

To determine the advantage of higher class children in 1901 absent the treatment effect, I regress the literate occupation dummy (Y) on father's class, controlling for age and parish of birth and using only those in the 1881-1901 linked sample who did not receive any treatment. Note that this includes individuals of all ages in untreated parishes as well as those in treated parishes too old to receive treatment. The baseline regression equation looks as follows:

$$Y_{iapc} = \beta(FC_i) + \alpha_a + \eta_p + u_{iapc} \quad (19)$$

where α_a represents age fixed effects and η_p parish fixed effects.

To verify that age effects do not vary by parish treatment status, I replace age fixed effects with arrival year-age fixed effects:

$$Y_{iapc} = \beta(FC_i) + \alpha_{ar} + \eta_p + u_{iapc} \quad (20)$$

Finally, if high class children concentrate in parishes with high opportunity for literate jobs, including parish fixed effects would bias the effect downward. Thus I also run the baseline regression excluding parish fixed effects:

$$Y_{iapc} = \beta(FC_i) + \alpha_a + u_{iapc} \quad (21)$$

The results of regression on Equations 19 - 21 are shown in columns (1)-(3) of Table U.1, respectively.

Table U.1: **Advantage of Untreated High Class Children, 1901**

VARIABLES	(1) Literacy Required	(2) Literacy Required	(3) Literacy Required
Father Class Dummy	0.19976*** (0.00212)	0.19974*** (0.00212)	0.23850*** (0.00259)
Observations	740,118	740,118	740,461
Parish Fixed Effects	YES	YES	NO
Age Fixed Effects	YES	NO	YES
Arrival-Age Fixed Effects	NO	YES	NO

Robust standard errors in parentheses, clustered at the parish level. *** p<0.01, ** p<0.05, * p<0.1