ISSN 2042-2695

CENTRE *for* ECONOMIC PERFORMANCE

CEP Discussion Paper No 1495

August 2017

Can Rising Instructional Time Crowd out Student Pro-Social Behaviour? Unintended Consequences of a German High School Reform

Christian Krekel

THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

Abstract

We study whether raising instructional time can crowd out student pro-social behaviour. To this end, we exploit a large educational reform in Germany that has raised weekly instructional hours for high school students by 12.5% as a quasi-natural experiment. Using a difference-in-differences design, we find that this rise has a negative and sizeable effect on volunteering, both at the intensive and at the extensive margin. It also affects political interest. There is no similar crowding out of scholastic involvement, but no substitution either. Impacts seem to be driven by a reduction in available leisure time as opposed to a rise in intensity of instruction, and to be temporary only. Robustness checks, including placebo tests and triple differencing, confirm our results.

Keywords: instructional time, student pro-social behaviour, volunteering, scholastic involvement, political interest, quasi-natural experiment, “G8” reform, SOEP
JEL codes: I21; I28; D01

This paper was produced as part of the Centre’s Wellbeing Programme. The Centre for Economic Performance is financed by the Economic and Social Research Council.

Acknowledgements

I would like to thank Claudia Senik, Andrew Clark, Katrin Rehdanz, Stephen Gibbons, Gert Wagner, Nicolas Ziebarth, Heike Solga, Sarah Dahmann, Mathias Hübener, Christopher Wratil, and especially Jürgen Schupp for valuable comments and suggestions. Financing by the German Ministry for Education and Science (FKZ: NIMOERT2/Kassenzeichen: 8103036999784/\#30857, Projekt: "Nicht-monetäre Erträge von Bildung in den Bereichen Gesundheit, nicht-kognitive Fähigkeiten sowie gesellschaftliche und politische Partizipation") is gratefully acknowledged.

Christian Krekel, Centre for Economic Performance, London School of Economics.

Published by
Centre for Economic Performance
London School of Economics and Political Science
Houghton Street
London WC2A 2AE

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means without the prior permission in writing of the publisher nor be issued to the public or circulated in any form other than that in which it is published.

Requests for permission to reproduce any article or part of the Working Paper should be sent to the editor at the above address.

© C. Krekel, submitted 2017.

1. Introduction

A growing body of empirical literature documents the importance of instructional time for student learning and performance (Patall et al., 2010). Raising instructional time – the allocated number of hours per year that students spend in formal classroom settings – is often found to have positive effects on cognitive skills such as maths and language ability (Bellei, 2009; Cortes and Goodman, 2014; Taylor, 2014), as well as standardised maths, reading, and scientific literacy test scores (Andrietti, 2016; Cattaneo et al., 2016; Huebener et al., 2017).[1] Differences in instructional time between countries are also found to account for some of the observed international gaps in student achievement (Lavy, 2015; Woessmann, 2003).[2] Thus, despite being a relative costly input into the educational production function, raising instructional time features high on the policy agenda in many countries (OECD, 2016). Yet, outcomes other than student learning and achievement have not been broadly studied (Patall et al., 2010), and particularly little is known about how changes in instructional time might influence student leisure activities and behaviours. Can raising instructional time have hidden costs by – unintentionally – crowding out student leisure activities and behaviours that parents, educators, and policy-makers alike would otherwise consider worth promoting?

In this paper, we are interested in a particular type of student behaviour: pro-social behaviour, defined as voluntary behaviour intended to benefit one or more individuals other than oneself (Eisenberg et al., 2013). This type of behaviour can cover a broad range of actions such as helping, sharing, and other forms of cooperation (Batson and Powell, 2003), and is distinct from altruism in that it is not purely motivated by increasing another individual's welfare, but can be motivated by, for example, empathy, reciprocity, or self-image (Evren and Minardi, 2017). Pro-social behaviour, and in particular volunteering, is linked to various positive outcomes: at the societal level, it can help build social capital through fostering cooperation and trust and through promoting citizenship (Putnam, 2000), and social capital is linked to higher levels of subjective well-being in societies (Helliwell et al., 2011); at the individual level, it is found to nurture important cognitive and non-cognitive skills that can improve individual

[1]There is growing evidence that the effect of raising instructional time on student learning and performance is heterogeneous, and in particular, that higher-performing students tend to benefit relatively more (Cattaneo et al., 2016; Huebener et al., 2017).

[2]The importance of instructional time for student achievement varies between educational systems, and in particular, between developed and developing countries (Woessmann, 2016), pointing towards potentially important complementarities in educational production, for example, between instructional time and teacher quality or effective classroom management techniques (Rivkin and Schiman, 2015).

labour market outcomes, to have positive physical and mental health benefits, and to raise subjective well-being over and beyond other benefits (Wilson and Musick, 2012), as confirmed in both observational (Binder and Freytag, 2013; Borgnonovi, 2008; Meier and Stutzer, 2008) and experimental studies (Dunn et al., 2008; Aknin et al., 2008). Specifically for youth, there is evidence that volunteering from an early age on enhances psychological development by raising self-esteem and self-confidence and by discouraging risky behaviours (Hart et al., 2007; Wilson and Musick, 2012).

To study the effect of raising instructional time on student pro-social behaviour, we exploit a large educational reform in Germany as a quasi-natural experiment: starting with school cohorts in the early 2000s, the number of school years required to obtain the university entrance qualification has been reduced from 13 to 12. In Germany, secondary education, which is compulsory until the age of 16, is tripartite: after (mostly) four years of joint primary education, students are tracked into different school types according to their abilities: lower, intermediate, or upper track schools.[3] Only the upper track leads to the university entrance qualification, and the reform has reduced the number of school years in this particular track – hereafter simply referred to as *high school* – from nine to eight. It aimed at reducing the graduation age of high school students, which was high in international comparison, to enable an earlier entry into the labour market. This, in turn, aimed at counteracting demographic change, especially an eroding contributor base and a shortage of skilled labour.

This reform – commonly referred to as ”G8”, where ”G” refers to high schools (*Gymnasien*) and ”8” to the new, reduced number of school years (as opposed to the old ”G9” system) – has two features that make it particularly interesting for us: first, the overall curriculum and thus total instructional time required to obtain the university entrance qualification has not changed, which, in turn, has lead to a 12.5% rise in weekly instructional hours across all subjects plus a rise in accompanying coursework.[4] Importantly, there have been no changes to the taught curriculum that target pro-social behaviour. In terms of an educational production function, this means that only instructional time has changed, whereas other inputs such as class size, in-

[3] Some federal states offer schools that combine the lower and intermediate track, or comprehensive schools or alternative school types that postpone tracking.

[4] Starting from the fifth grade (the first year of secondary education), students generally have to complete at least 265 year-week hours before being allowed to take the university entrance qualification exam (Standing Conference of the Ministers of Education and Cultural Affairs, 2016). Thus, average instructional hours per year increased from 1,051 to 1,184, compared to 950 in upper secondary education in England and 1,038 in the US (OECD, 2014). The rise in weekly instructional hours can be calculated as $(((265/(8))/(265/(9))) - 1) \times 100$.

structional materials, and teacher quality have not been changed as a result of the reform. This allows us to estimate the "pure" effect of raising instructional time on student pro-social behaviour, excluding potentially confounding changes to the educational system that are typically accompanied by similar reforms. Second, since education in Germany is the responsibility of the 16 federal states, there has been a staggered implementation of the reform: while some federal states implemented it as early as 2001 (*Saarland*), others waited until 2007 (*Schleswig-Holstein*); yet others have never fully implemented it (*Rhineland-Palatine*), or as in case of *Saxony* and *Thuringia*, have always required only 12 school years to obtain the university entrance qualification (Standing Conference of the Ministers of Education and Cultural Affairs, 2016). This allows the estimation of the causal effect of raising instructional time on student pro-social behaviour by exploiting variation in the implementation of the reform.

Figure A.1 shows this variation across federal states and school cohorts (highlighted in different shades of gold). It also shows the share of students in the different tracks (highlighted in different shades of red): in school year 2013/14, of 5,187,960 students in total, 2,329,990 (45%) were in the upper track; with few exceptions, they made up the largest share of students in each federal state.

Using survey data on youth and adolescents from the German Socio-Economic Panel Study (SOEP) and a difference-in-differences design, we find that the 12.5% rise in weekly instructional hours significantly crowds out student pro-social behaviour: it has a negative and sizeable effect on volunteering, decreasing the likelihood to volunteer at least once a month by about six percentage points. Given that almost 34 percent of students report to volunteer at least monthly, this amounts to a decrease of about 19 percent in this share. In other words, the rise in instructional time leads almost every fifth student to change her behaviour from volunteering at least monthly to volunteering less often or not at all. This change is primarily driven by students who report to volunteer on a weekly basis, and it affects both the intensive and extensive margin of volunteering: while two thirds of the students cut back on their activities, the other third give them up completely. Students from disadvantaged backgrounds are more likely to cut back on their activities. We find no similar crowding out of scholastic involvement, but no substitution either. Interestingly, we find that the rise in instructional time has a differential impact on student political interest: it leads to a depolarisation at both ends of the spectrum, decreasing the share of students who report to be at least fairly interested in politics while at the same time decreasing the share that report to be not interested at all. The size of these changes is very strong: every third student switches category. Impacts seem to be driven by a reduction in available leisure time as opposed to a rise in intensity of instruction, and to be temporary only. The results are robust to a different model specification, time trends, and seasonal variation; selection and implementation; triple differencing, and potentially confounding other reforms that are implemented during the observation period. They also withstand a series of placebo tests.

This finding is significant for several reasons: first, in the given context, it is significant because of the sheer number of students affected. In Germany, in school year 2013/14 alone, the reform affected 2,329,990 high school students, about half of the entire student population in secondary education (Federal Statistical Office, 2016b). Second, it is significant because of the important role pro-social behaviour, and in particular volunteering, plays, both for individuals, as described above, and for society at large: the OECD estimates the economic value of volunteering for Germany in 2013 to be around USD 117.6 billion or 3.3% of real GDP (OECD, 2015), roughly comparable to the UK and the US. From a time use perspective, the average German citizen spends about seven minutes per day volunteering, compared to about eight minutes per day for the average US citizen (OECD, 2011). Finally, to the extent that students from disadvantaged backgrounds are disproportionally affected, the decrease in volunteering for these groups might further increase educational inequalities, and thus inequalities in later life outcomes.

We contribute to two strands of literature: first, we contribute to the economic literature on the external, non-monetary effect of education on civic engagement, which focuses on the effect of years of education on predominantly political interest, information, and participation (Dee, 2004; Dhillon and Peralta, 2002; Milligan et al., 2004; Pelkonen, 2012; Siedler, 2010), as well as reciprocity (Fehr and Gachter, 2000; Kosse et al., 2014).[5] Here, the study most closely related to ours is Gibson (2001): the author uses a sample of twins to hold unobservable family characteristics constant, showing that more years of education are associated with a lower probability of volunteering and supply of volunteer hours. We complement this study by focusing on instructional time rather than amount of instruction.[6] Second, we contribute

[5]See Lochner (2011) and Oreopoulos and Salvanes (2011) for reviews.

[6]Next to this literature in economics stands a large body of literature in political science on the relationship

to the literature on instructional time (Bellei, 2009; Cortes and Goodman, 2014; Cortes et al., 2015; Herrmann and Rockoff, 2012; Taylor, 2014), and in particular, to the stream that exploits the "G8" reform as a source of exogenous variation: since the first data became available, the reform has been used – due to its features – as a laboratory for empirical research in educational economics. The more sophisticated studies use difference-in-differences designs that exploit variation in its implementation across federal states and school cohorts; they examine its effects on graduation age, grade repetition, and graduation rates (Huebener and Marcus, 2015), post-secondary educational choices (Meyer et al., 2015), performance (Andrietti, 2016; Homuth, 2012; Huebener et al., 2017), health (Quis and Reif, 2017), and well-being (Quis, 2015).[7] Here, the studies that are methodologically most closely related to ours are Dahmann and Anger (2014) and Dahmann (2017): we use the same dataset and a similar specification as these authors, who show that the reform affects personality traits, and to some extent, cognitive skills. So far, the potentially negative effects on leisure activities of youth and adolescents have played only a minor role relative to educational outcomes, although this point has sparked considerable controversy amongst students, parents, and educators alike (see, for example, Süddeutsche Zeitung (2010) for a feature), and continues to do so today. The only academic study so far that has looked at leisure activities of youth and adolescents is Meyer and Thomsen (2015): the authors use self-collected cross-section data on students from the double graduation cohort (which may be subject to implementation effects) in the federal state of *Saxony-Anhalt* two years after graduation, showing that students in this cohort and state indeed felt more pressured and tended to spend less time on leisure activities such as jobbing or volunteering. More generally, the impact of instructional time on student pro-social behaviour has received little attention so far.

The rest of this paper is organised as follows: Section 2 describes the data used in the empirical analysis, Section 3 the empirical model and identification strategy. The results, including robustness checks, are presented in Section 4. Section 5 discusses them against the background of recent trends in the educational sector, and gives policy implications.

between education and political participation, especially voter turnout. See, for example, Henderson and Chatfield (2011), Hillygus (2005), Persson (2014), and Sondheimer and Green (2010), to name just a few.

[7] Early studies relied mostly on double graduation cohorts, often analysed within a single federal state (see Büttner and Thomsen (2015) or (Thiel et al., 2014), for example). These cohorts, however, may be subject to implementation effects.

2. Data

2.1. German Socio-Economic Panel Study

The German Socio-Economic Panel Study (SOEP)[8] is a representative panel of private households in Germany. It has been conducted annually since 1984, and includes almost 30,000 individuals in more than 11,000 households in its current wave. The SOEP provides rich information on all household members, covering Germans living in the old and new federal states, foreigners, and recent immigrants (Wagner et al., 2007, 2008). Most importantly, it provides information on the volunteering, scholastic involvement, and political interest of youth and adolescents, as well as on their demographic, educational, and parental household characteristics.

During fieldwork, typically, two types of questionnaires are administered: an individual questionnaire is filled out by each household member aged 18 and above; a separate household questionnaire is filled out by the household head. The former covers personal characteristics such as education, leisure activities, and attitudes, the latter household and neighbourhood characteristics that apply to all household members equally. Moreover, since 2000, a separate youth questionnaire including both prospective and retrospective items on childhood and schooling is administered to youth in the year in which they turn 17. This is when individuals enter the SOEP at the earliest. If they enter at a later point in time, they are administered – in addition to the individual questionnaire – a supplementary biography questionnaire that includes most of the items of the youth questionnaire in order to complement missing information.

The youth questionnaire is our main data source: it includes items on the volunteering, scholastic involvement, and political interest of youth annually from 2006 onwards. To increase sample size, we complement these data with data on adolescents from the individual questionnaire, which includes the same items on volunteering biannually from 2001 and on political interest biannually from 2000 onwards (with few exceptions). The supplementary biography questionnaire complements items on scholastic involvement.[9] The SOEP also provides readily usable, generated items on educational trajectories of respondents, including the year and federal state in which they started school, the type of school they are currently attending, and in

[8]Socio-Economic Panel (SOEP), data for years 1984-2015, version 32, SOEP, 2016, DOI: 10.5684/soep.v32

[9]In unreported robustness checks, we account for between-survey differences at any point in time by including a dummy variable for the respective survey: the results remain robust, and are available upon request. Moreover, we account for within-survey differences over time by routinely controlling for school cohorts.

case they have already graduated, the year and federal state in which they have graduated, as well as the degree they have obtained. In case the year or state of school enrolment is missing, we impute it using their date of birth or state of residence, respectively.[10] If we have multiple observations of the same individual, we only include the observation at the youngest age.

We restrict our sample to the years 2000 to 2015, and to individuals aged 17 to 20 in order to create a homogeneous age group and avoid confounding effects associated with entrance into tertiary education. We focus only on high school students and graduates, as only those have been affected by the reform. In doing so, we omit students from comprehensive schools: as we cannot clearly identify whether these students are attending or graduated from the academic track, we take a conservative approach and omit them altogether. Moreover, we omit all individuals from federal states where the reform has never been implemented, or during years in which it has been implemented only partially. Finally, we omit all individuals with missing observations on outcomes and covariates. Depending on how many observations on outcomes are available, this gives us a sample of 2,240 students for volunteering, 1,958 for scholastic involvement, and 2,536 for political interest.[11]

Outcomes

Volunteering. We select *volunteering* as our main outcome for pro-social behaviour. The indicator is obtained from a single-item five-point Likert scale that asks respondents "How often do you do volunteer work in clubs or social services during free time?". Possible answers include "daily" (about 6% of respondents), "every week" (16%), "every month" (12%), "less often" (29%), and "never" (37%). We create a binary indicator that equals one if respondents volunteer at least once a month, that is, if they volunteer daily, weekly, or monthly, and zero else. About 34% of respondents do so.[12]

[10]When benchmarking the imputed values with the original ones, we find that they match in about 99% of cases for the state and 66% of cases for the year of school enrolment. Obviously, for the latter, there is some discretion on side of parents (we account for differences in cut-off dates for school enrolment across states and over time): if we assume that parents have a tendency to redshirt, that is, to strategically postpone school enrolment in order to provide their children with educational advantages due to relative and absolute maturity (Bedard and Dhuey, 2006; Black et al., 2011), their children are correctly allocated to the treatment group. Enrolling in school prematurely, on the contrary, is very rare in Germany.

[11]If not stated otherwise, descriptive statistics are given on the sample for volunteering.

[12]In our baseline specification, we use a binary indicator because it splits the share of students who volunteer at least once a month, less often, and never into approximately equal shares. In extended specifications, at a later point, we will also use binary indicators for each category (see Figure A.8): as we shall see, these turn out insignificant, as the reform shifts the entire volunteering distribution.

Figure A.2 shows the development of this outcome over the observation period. The x-axis denotes the interview year, and the y-axis the fitted annual mean, covariate adjusted for observables described in Sub-Section 2.1.

We can see that, over the past decade, there has been an initial rise in the share of students who volunteer at least once a month, up until the year 2005, whereafter this share started to decline until it reached again its initial value in the year 2015.[13]

As this indicator is framed in such a way as to refer to activities outside school, we select various ways of *scholastic involvement* as additional outcomes to cover activities inside, in line with a broad definition of pro-social behaviour. The respective indicators are obtained from a battery of binary items that asks respondents "Besides normal classes, there are also other ways to get involved in school. Have you ever – before or right now – been involved in one or more of the following ways?" Possible answers include "student representative" (about 3% of respondents), "class representative" (42%), "school magazine" (10%), "drama or dance group" (20%), "choir or orchestra" (33%), "sports group" (27%), "other voluntary group" (37%), and "none" (19%). We create a binary indicator for each activity that equals one if respondents have ever been engaged in it, and zero else.

Political Interest. Pro-social behaviour is voluntary behaviour intended to benefit one or more individuals other than oneself (Eisenberg et al., 2013), and is distinct from altruism in that it is not purely motivated by increasing another individual's welfare, but can be motivated by, for example, empathy, reciprocity, or self-image. We adopt a broad definition of pro-sociality here, and are also interested in political behaviour and participation, which can (but, of course, does not necessarily have to) be motivated by the willingness to benefit specific groups of society or society as a whole.

Apart from items on voting intentions in federal elections, the SOEP does not include specific items on political behaviour, for example, on membership in political parties or participation in youth organisations.[14] However, it regularly asks respondents about their degree of

[13] Figure W.1 in the Web Appendix shows the development of volunteering for students in the lower and intermediate track: compared to those in the upper track, these students tend to volunteer less. The rise in the share that volunteers at least once a month prolonged much longer, up until the year 2009, whereafter it started to decline.

[14] As federal elections (normally) happen only once every four years, the sample size is not large enough to analyse these items. The SOEP also asks respondents whether they lean towards a specific party, and if so,

interest in politics more generally. As interest has long been seen as a necessary condition for subsequent behaviour (Fishbein and Ajzen, 1975), we select *political interest* as outcome to proxy for political behaviour. The indicator is obtained from a single-item four-point Likert scale that asks respondents “Generally speaking, how much are you interested in politics?”. Possible answers include “very much” (about 6% of respondents), “much” (26%), ”not so much” (52%), and “not at all” (15%). We create a binary indicator for each of these categories.

Subjective Well-Being. As our final outcome, we select the self-reported *life satisfaction* of students, an evaluative measure of subjective well-being. To be clear, we are not interested in the impact of the reform on student subjective well-being *per se*: this has been researched elsewhere, and our focus is pro-social behaviour. We will, however, exploit this indicator to shed some more light on whether the impact of the reform on pro-social behaviour is driven by a reduction in available leisure time, or alternatively, by a rise in intensity of instruction. The indicator is obtained from an eleven-point single-item Likert scale that asks respondents “How satisfied are you currently, all in all, with your life?”. Possible answers range from “0” (completely dissatisfied) to “10” (completely satisfied), with a mean of 7.8, suggesting that students are quite happy with their lives.

Covariates

We routinely control for age and whether a student has graduated in all our regressions. The mean age of students is 17.5, and only 4% of them have already graduated. Moreover, in our preferred specification, we control for a rich set of other demographic and parental household characteristics. These include gender (about 53% of students are female), migration background (about 20% have a migration background, either direct or indirect), and place of residence (about 13% live in East Germany and 28% live in rural areas). When it comes to their parents, about 52% of students have at least one parent with a tertiary degree, 12% have a parent that is a blue-collar worker, and 64% have a parent that works full time. Finally, about 4% of students are risen by a single parent, and about 17% are the only child. The average number of children in the household is 2.4. See Table W.1 in the Web Appendix for more descriptive statistics.

towards which party they lean and to what extent. As there is no *a priori* reason to believe that an increase in instructional time changes political orientation, we do not analyse these items.

3. Empirical Strategy

To investigate whether raising instructional time can crowd out pro-social behaviour, we exploit the recent educational reform in Germany that reduced the number of school years required to obtain the university entrance qualification as a quasi-natural experiment. Specifically, we set up a difference-in-differences design that exploits variation in the implementation of the reform across federal states and school cohorts: students are allocated to the treatment group if they belong to a school cohort in a federal state which was affected by the reform (or, in other words, if they enrolled in the year in which the reform was implemented or any year thereafter in the respective federal state), and to the control group else. Thus, students in the treatment group are exposed to more weekly instructional hours of, on average, 12.5% plus accompanying coursework than those in the control group. For both groups, however, the taught curriculum is the same. From 2,240 students in our sample on volunteering, 994 are in the treatment and 1,246 are in the control group; for scholastic involvement, these are 936 and 1,022 out of 1,958 students; for political interest 1,000 and 1,536 out of 2,536; and for subjective well-being 1,011 and 1,533 out of 2,544. Table W.2 in the Web Appendix exemplary shows the distribution of students by age in treatment and control group for volunteering over time.

3.1. Regression Equation

We employ linear probability models, which are estimated using ordinary least squares with robust standard errors clustered at the federal state level.[15] More specifically, following Dahmann and Anger (2014) and Dahmann (2017), we use the following specification:

$$y_{isc,(17-20)} = \beta_0 + \beta_1 Reform_{sc} + \beta_2' \mathbf{X_{isc,(17-20)}} + \\ + \sum_{s=1}^{16} \gamma_s State_s + \sum_{c=1}^{14} \delta_c Cohort_c + \varepsilon_{isc,(17-20)} \tag{1}$$

where y is the pro-social behaviour of student i in federal state s and school cohort c, measured at age 17 to 20; *Reform* is a dummy variable that equals one if the student belongs to a

[15]In our preferred specification, less than 1% of predicted values lie outside the [0;1] interval. Moreover, the results are similar when using a probit model, as shown in Table A.6. Out-of-sample prediction, therefore, seems to be less of an issue. Finally, the results remain the same when using weighted regressions and bootstrapped standard errors. See Table W.6 in the Web Appendix for these results.

school cohort in a federal state which was affected by the reform, and zero else; and X is a vector of controls, including demographic, educational, and parental household characteristics. We routinely include a full set of federal state and school cohort dummy variables.[16] Our regressor of interest is β_1, which captures the reform effect. It can be interpreted as the average treatment effect on the treated, and is causal if the identifying assumptions described in Sub-Section 3.2 hold.

This difference-in-differences design has two features. First, it is generalised in the sense that treatment can occur at different points in time for different individuals. In fact, at any point in time over the observation period, we compare students who are affected by the reform with those who are not (yet) affected.[17] Thus, towards the beginning of the observation period, the treatment group is relatively small, and as the reform gradually fades in, it increases as more and more observations on affected students become available, and *vice versa* for the control group. Second, this difference-in-differences design is pseudo in the sense that we only observe each student once. This is due to the fact that individuals enter the SOEP in the year in which they turn 17 at the earliest.[18] In other words, at the point of the first interview, students are near school completion, or even shortly thereafter. As a consequence, we cannot observe their pre-treatment outcomes, which would have had to be recorded prior to enrolment.[19]

This difference-in-differences design imposes stronger identifying assumptions than a conventional one. For example, as we do not observe the same individuals over time but compare different individuals at the same points in time, we cannot readily net out unobserved heterogeneity amongst individuals by including individual fixed effects; rather, in case there is unobserved heterogeneity, we have to assume that there is a balance in unobservable characteristics between treatment and control group, and that this balance remains constant over time (this is sometime referred to as *bias stability*) (Heckman et al., 2009). In Sub-Section 3.2 we provide

[16] We also routinely include controls for sub-samples, as the SOEP consists of 16 random samples, which partly focus on different population strata.

[17] As a robustness check, in Sub-Section 4.2, we implement a triple differencing design which exploits the fact that the reform affected only students in the upper track by employing students in the other tracks as a clean control group.

[18] The SOEP also includes several mother-child questionnaires, which have been administered since 2003. However, these questionnaires, which are highly age-specific and cover the age span from birth to 10, are completed by the mother and do not include the items that are relevant for this study. A separate student questionnaire, covering ages 11 and 12, has been administered since 2014 only (and does not include these items either).

[19] Strictly speaking, even if we would observe their pre-treatment outcomes, it is questionable whether we could use them effectively: the kind of pro-social behaviour we are interested in plays a relatively minor role prior to age 12.

evidence that, although our identifying assumptions are stronger, they are likely to hold.

3.2. Identification

Our main identifying assumption is that, in the absence of treatment, the pro-social behaviour of students in the treatment group would have followed the same time trend as that of students in the control group. Although this *common trend assumption* is not formally testable as the counterfactual is not observable, in Sub-Sections 3.2 and 3.2, we provide evidence that it is likely to hold.[20]

Balancing on Observables

The first piece of evidence comes from Table A.1: it shows the means of all covariates, overall and separately for treatment and control group, along with their scale-free normalised differences. Here, covariate imbalance between treatment and control group could indicate a deviation from a common time trend.

Imbens and Wooldridge (2009) suggest that a normalised difference above 0.25 indicates covariate imbalance. This is not the case for most of our covariates: only the age is above the threshold, and whether a student has graduated comes close. This is no surprise, though, given that the reform explicitly aimed at reducing the number of school years, thus indirectly reducing the graduation age. In fact, Huebener and Marcus (2015) estimate that the reform decreased the graduation age by about 10 months. Thus, we conclude that the sample is well-balanced on observables, and therefore most likely on unobservables as well. Finally, we routinely control

[20]Implicitly, we also require *ignorability* and the *stable unit treatment value assumption* to hold: the former implies that treatment assignment is independent of the outcome, the latter that whether a student is treated or not should not depend on the outcome of another student. Both are likely to be true: the rise in instructional time for a student does not depend on the amount of volunteering of that student, neither does it depend on the amount of volunteering of another student. Moreover, there should be no variation in treatment intensity between students. Again, this is likely to be true as the reform aimed at reducing the number of school years only while holding everything else constant. For the vast majority of students in the first school cohorts affected, the resulting rise in instructional time was present from the point of enrolment onwards. Only students in the federal states of *Saxony-Anhalt* and *Mecklenburg-West Pomerania* had already started school when the reform was implemented. In fact, these students were in grades seven to nine, in which some schools allocated a disproportionally higher share of the overall rise in instructional time, potentially yielding a different treatment intensity for these students. In Sub-Section 4.2, we explore this possibility in more detail.

for age and whether a student has graduated in all our regressions in order to rule out any age and graduation effects.[21]

Graphical Evidence

Next, we take a closer look at how volunteering, our main variable of interest, evolves over time. Figure A.3 is constructed similarly as Figure A.2: it shows the development of volunteering over the observation period. The x-axis denotes the interview year. Different from Figure A.2, there are now two y-axes: the left y-axis denotes the fitted annual mean, covariate adjusted for observables, whereas the right y-axis denotes the percentage of the interviewed who were treated. The vertical line marks the interview year before the first observations of the treated become available.

It is clearly visible that the vertical line marks a structural break, dividing the observation period into two: using local-mean polynomial smoothing, we can see that there is a clear upwards trend in volunteering in the first half of the observation period, whereas in the second, this trend is reversed. Moreover, the trend reversal coincides with an increasing share of the treated amongst the interviewed.

Figure A.4 takes Figure A.3 one step further: it decomposes, in the second half of the observation period, the overall mean into that of the treatment and control group, respectively. It also plots – in addition to that of the overall mean – the polynomial fit of the control group mean.[22]

We can make three observations. First, when focusing on the control group mean only, it becomes clear that part of the trend reversal in volunteering probably would have come about in the absence of the reform: the polynomial fit of the control group mean tilts downwards irrespective of whether the share of the treated amongst the interviewed increases.[23] Second,

[21]Note that covariance imbalance between treatment and control group would not necessarily be a threat to our identification strategy: we control for a rich set of time-varying observables in our preferred specification. Moreover, including federal state and school cohort dummy variables nets out systematic differences in both time-invariant observables and unobservables between federal states and school cohorts, respectively.

[22]See Figure W.2 in the Web Appendix for a similar illustration of political interest.

[23]This also raises the question to what extent the identified reform effect is driven by time trends. In Sub-Section 4.2, we explore this possibility in more detail.

the treatment group mean is systematically lower than the control group mean, and as the share of the treated amongst the interviewed increases, the difference between the polynomial fit of the overall and that of the control group mean increases as well. This is already suggestive that part of the trend reversal in volunteering is indeed driven by the reform; in our regressions, we are measuring the mean difference between the control group and the treatment group mean in the second half of the observation period. Finally, important for identification, the treatment group mean, when fading in, evolves in parallel to the control group mean, when fading out. This is suggestive of a common trend between treatment and control group.

To illustrate this common trend in more detail, we plot the overall mean for different federal states that implemented the reform quite late during the observation period. Figures A.5 and A.6 are constructed similarly to Figure A.3: the former, Figure A.5, shows the overall mean for two groups of states in which the first observations of the treated become available in the same interview year, pooled together (the first group, labelled "Gr. 1", includes the states of *Baden-Wuerttemberg*, *Bavaria*, *Bremen*, *Hesse*, and *Lower Saxony*; the second group, labelled "Gr. 2", includes the states of *Berlin*, *Brandenburg*, and *Schleswig-Holstein*). The latter, Figure A.6, shows them separately for two large area states in which this is not the case (the first state, labelled "St. 1", is the state of *Schleswig-Holstein*; the second state, labelled "St. 2", is the state of *North Rhine-Westphalia*).[24]

Again, we can make three observations. First, irrespective of whether we plot the overall mean for groups of states pooled together or separately for single states, there is a common trend between these states before the first observations of the treated become available. Second, the interview year before the first observations of the treated become available marks a structural break. Finally, after this structural break, these states once again exhibit common trend behaviour.[25]

Taken together, the balancing properties of observables and the graphical evidence is clearly supportive of a common trend between treatment and control group. Moreover, in case there is unobserved heterogeneity, there seems to be a balance in unobservable characteristics between them that remains constant over time.

[24] Again, see Figure W.3 in the Web Appendix for a similar illustration of political interest.

[25] The latter point is also suggestive evidence that the *stable unit treatment value assumption* is likely to hold: common trend behaviour post-treatment implies that treatment intensity is likely to be the same across states.

4. Results

4.1. Baseline Results

We now turn to our baseline results in Table A.2: column (1) includes only the reform dummy variable, our regressor of interest; columns (2) and (3) then successively add age and a graduation dummy variable in order to account for age and graduation effects. Finally, column (4) includes all of the above, along with a rich set of other demographic and parental household characteristics; it is our preferred specification, and the regression equivalent to Figure A.4.[26]

Volunteering

Table A.2 shows that the reform has a negative and sizeable effect on volunteering across the board, which is significant at the 1% level: in our preferred specification, it decreases the likelihood to volunteer at least once a month by about six percentage points. The size of this effect is also economically significant: given that almost 34 percent of all students in the sample report to volunteer at least monthly, it amounts to a decrease of about 19 percent in this share. In other words, the reform led almost every fifth student to change her behaviour from volunteering at least monthly to volunteering less often or not at all. The fact that the sign, size, and significance level is similar across all models reinforces the notion of a quasi-natural experiment.[27]

Another way to look at this result is through the lens of an event study. Figure A.8 plots the fitted annual mean of volunteering alongside 99% confidence intervals in the years just before and the years just after the implementation of the reform (which varies by federal state), whereby the year of implementation is normalised to be at $t = 0$. As can be seen, in the years running up to the reform, the share of students who volunteer at least once a month follows a rather stable path, averaging between 41% and 38%. In the year of implementation, at $t = 0$, this share falls sharply to about 34%, and remains stable at a lower endline level in the ensuing years.

[26] See Table W.3 in the Web Appendix for the full set of controls.

[27] This is suggestive evidence that *ignorability* is likely to hold, even unconditionally: the fact that our estimates vary so little depending on covariates implies that treatment is likely to be exogenous.

As our outcome is a binary indicator constructed from a categorical variable, it would be interesting to see how the overall frequency distribution of volunteering changes due to the reform. Figure A.7 illustrates this: it compares the means of the different frequencies of volunteering before and after the reform.

We can make three observations. First, the reform affects the entire frequency distribution of volunteering, as all categories are affected, although to different degrees. Second, the driving force behind the decrease in the share of students who volunteer at least once a month are students who volunteer weekly, followed by those that volunteer monthly: the share of the former drops by about 25%, the share of the latter by about 14%. On the contrary, the share of students who report the highest frequency of volunteering sees almost no reduction (about 10%). This, however, is only a small fraction: about 5% report to volunteer daily, as opposed to about 16% and 12% reporting to volunteer weekly and monthly, respectively. Second, these reductions are met with rises by about 13% and 7%, respectively, in the share of students who volunteer less often and the share of students who volunteer never; the difference between these flows is significant. This implies that the reform affected both the intensive and the extensive margin of volunteering: while some students (about two thirds) cut back on their activities, others (about one third) gave them up completely, which is broadly in line with the results obtained for the double graduation cohort in the federal state of *Saxony-Anhalt* by Meyer and Thomsen (2015). At the same time, this might point towards potential effect heterogeneities, and indeed, although we find little evidence that the effects vary much by student demographics and achievement, we find, in line with evidence from other OECD countries (OECD, 2015), that students from disadvantaged backgrounds – defined here as students with at least one parent being a blue-collar worker – are slightly more likely to cut back on their activities (see Table W.4 for this result).

The question arises whether there is a similar crowding out of scholastic involvement as for volunteering. Alternatively, one could ask whether the crowding out of volunteering is matched by an increase in scholastic involvement. In other words, is there a substitution of activities outside school with activities inside? Table A.3 shows that neither is the case: it takes our preferred specification, column (4) in Table A.2, and uses the likelihood of various ways of scholastic involvement as outcomes. Clearly, the reform has no significant effect on any of them, and neither is there a clear pattern in terms of sign. To get a sense of whether the reform affects the extensive margin of scholastic involvement, we also tested an alternative outcome: a binary indicator that equals one if respondents have ever been engaged in any of the activities in columns (a) to (g), and zero else. Again, the reform has no significant effect on this alternative outcome (not shown).[28] A potential caveat of this analysis is that we have slightly fewer observations for scholastic involvement than for volunteering: the sample size decreases from 2,240 to 1,958 students. This decrease, however, is mostly driven by students in the control group: 936 are now in the treatment and 1,022 are in the control group.

Political Interest

Finally, we ask how the reform affects political interest, which we take as a proxy for political behaviour. Table A.4 sheds light on this question. Once again, we take our preferred specification, column (4) in Table A.2, and use the likelihood of being interested in politics with a particular strength, including strongly, fairly, weakly, or not at all, as outcomes. We also combine the first two categories to form a new one, namely being moderately interested in politics.[29]

Interestingly, we find that the reform has a differential impact on political interest: it has a significantly positive effect on being weakly interested at the 5% level. At the same time, however, it has a significantly negative effect on not being interested at all as well as on being moderately interested at the 5% level.[30] In other words, there is a depolarisation at both ends of the spectrum: the reform decreases the share of students who report to be moderately, that is, at least fairly, interested in politics by about 11 percentage points while at the same time decreasing the share that report to be not interested at all by about six percentage points. Taken

[28]In another specification, we excluded the activities in columns (a) and (b). Arguably, these activities should react inelastically to changes in instructional time: by German school law, there has to be a student and a class representative. The result, however, remains the same.

[29]In this analysis, we have slightly more observations: the sample size increases from 2,240 to 2,536, 1,000 of which are now in the treatment and 1,536 are in the control group.

[30]See Table W.5 in the Web Appendix for the result on being moderately interested in politics.

together, this equals the incremental 17 percentage points of those being weakly interested. These migration flows are very strong: every third student switches from the higher category to the lower, and *vice versa*.[31] We find no effect heterogeneities with respect to student socio-demographics and achievement.

A potential explanation for this differential impact on political interest is that the reform crowds out political interest on one side of the spectrum, namely for those already interested in politics, while at the same time encouraging others on the other side to become politically active, especially those who have not been so previously, for example, by joining a protest group or party that opposes the reform. In fact, the reform has sparked considerable controversy amongst students, parents, and educators alike (some in anticipation of the adverse effects presented in this study), and continues to do so today. This has led some federal states to announce its revocation, and others like the federal state of *Rhineland-Palatine* not to implement it in the first place.

Channel

The reform has reduced the number of school years in high school from nine to eight, while leaving the taught curriculum and all other inputs into the educational production function constant. There are two potential ways to implement such a reform: the first is to leave the total number of weekly instructional hours constant and go through the curriculum at a faster pace, the second to append more weekly instructional hours and go through the curriculum at the same pace. The question now is whether the identified effects are driven by an increase in the intensity of instruction (the former) or a decrease in the availability of leisure time (the latter). The guidelines for the implementation of the reform and the available evidence so far suggests that the number of weekly instructional hours has indeed increased, thus decreasing the number of weekly leisure hours (Homuth, 2017). A reduction in the number of weekly leisure hours has then, potentially, decreased the time available for volunteering activities. How can we provide more evidence for this channel?

One way to do this is to standardise our primary outcome – volunteering at least monthly – by the available leisure time per month: 24×7 week hours in total minus about 5×6 week hours in school under the old regime equals 138 hours in available leisure time per week or

[31] In unreported robustness checks, we include dummy variables for state, federal, and European elections, either individually or jointly: the results remain robust, and are available upon request.

552 hours per month; 24×7 week hours in total minus about 5×7 week hours in school under the new regime equals 133 hours in available leisure time per week or 532 hours per month.[32] The first column of Table A.5 uses our preferred specification, column (4) in Table A.2, with volunteering at least monthly standardised by the available leisure time per month as outcome: as expected, the coefficient is significant but greatly attenuated, suggesting indeed that the identified effects are driven by a decrease in the availability of leisure time.[33] This is also corroborated by the fact that, in our preferred specification, the likelihood to volunteer at least monthly decreases by about 19 percent as the number of weekly instructional hours increases by about 13 percent – an almost linear relationship.

Another piece of evidence comes from subjective well-being data: the second column of Table A.5 uses our preferred specification, column (4) in Table A.2, with self-reported life satisfaction as outcome. As can be seen, the reform has no significant impact on this evaluative measure of subjective well-being. Arguably, an increase in the intensity of instruction may raise stress levels, everything else held constant, which may lead to lower levels of subjective well-being. The non-finding may suggest that, although the reform has reduced the available leisure time of students, in order to remain on the same welfare level, students may have responded by reducing their commitment to leisure activities, which, in turn, may have alleviated their stress levels.

4.2. Robustness Checks

In the following, we conduct a number of robustness checks to confirm the robustness of our baseline results. Specifically, we test whether they remain robust to a different model specification, time trends, and seasonal variation; selection and implementation; triple differencing; and potentially confounding other reforms that are implemented during the observation period. We also conduct a series of placebo tests. All robustness checks build on our preferred specification, column (4) in Table A.2. For the sake of brevity, we focus on volunteering, our main

[32] Implicitly, this accounts for sleeping time as there could, hypothetically, be substitution effects with sleep.

[33] Note that standardisation by available leisure time introduces measurement error: if a respondent indicates to volunteer weekly, this could mean that she volunteers once a week up to, theoretically, six times a week, given that the next higher category is to volunteer daily (at, on average, more than 3.5 volunteering events per week the respondent might indicate to volunteer daily). The same holds true for volunteering monthly. At the same time, there is only a very small share of students who volunteer daily (they are also the least affected), which would be the most precise measure.

variable of interest.

Model Specification, Time Trends, and Seasonal Variation

First, we turn to a different model specification. In column (1) of Table A.6, we use a probit instead of a linear model. As can be seen, the reform still has a negative effect on volunteering, which is significant at the 1% level. The size of the coefficient, however, is slightly larger.

Figure A.4 suggests that some of the decline in volunteering during the observation period probably would have come about in the absence of the reform, which raises the question to what extent the identified reform effect is driven by time trends. To be clear, this is not a threat to our identification strategy as long as time trends do not affect treatment and control group differentially, and time trends are not correlated with the outcome. To explore this possibility nevertheless, in columns (2) and (3) of Table A.6, we include a linear and quadratic time trend, respectively. Then, in column (4), we include both of them at the same time. As can be seen, the reform still has a negative effect on volunteering, which is significant at the 1% level, across all models, and the size of the coefficients is very similar. We go even one step further: in column (5), we include both state-specific linear and quadratic time trends, counting up the years for each state individually, and in column (6), we include both treatment-specific linear and quadratic time trends, counting up the years for each state individually starting from the interview year in which the first observations of the treated amongst the interviewed become available. Arguably, both specifications are very restrictive in the sense that they take out much variation in the data, which is in part reflected by lower significance levels. The point estimates remain, nevertheless, quite robust.

Finally, we turn to seasonal variation. Again, this is not a threat to our identification strategy as long as treatment and control group are not systematically interviewed at different dates, and interview dates are not correlated with the outcome. To explore this possibility nevertheless, in columns (7) and (8) of Table A.6, we include quarterly and monthly dummy variables, respectively. As expected, the sign, size, and significance level of the reform effect in both models is very similar to that in our preferred specification.[34]

[34]One might argue that, at the time of interview, students in the treatment group are relatively closer to their high school finals than those in the control group, which might, in turn, partially or even fully account for the

Selection and Implementation

Next, we turn to selection, which may come in two flavours: within-sample and out-of-sample selection. First, students may self-select from the treatment into the control group within the sample, for example, by moving from one federal state to another in order to avoid the reform.[35] Alternatively, students may self-select out of the sample altogether, for example, by dropping out of high school. To be clear, this is not a threat to our identification strategy as long as self-selection is not correlated with the outcome. Assuming that students who move or drop out are those who are most adversely affected by the reform, our estimates are downward biased and can be interpreted as a lower bound.

We believe that within-sample selection is unlikely to be an issue: moving from one federal state to another is associated with high monetary and non-monetary costs for both students and parents. Besides, geographic mobility in Germany is traditionally low: in a given year, only about 6% of respondents in the SOEP move. This is even more so the case in a selective sample like ours, comprising families with children that attend high school: in a given year, only about 3% of them move. Nevertheless, in column (1) of Table A.7, we evaluate how movers affect our estimates: here, we exclude all students who move during the observation period. As it turns out, this does not change our estimates much: the reform still has a negative effect on volunteering, which is significant at the 1% level; the size of the effect is somewhat reduced.[36] A more serious problem arises, however, for students living close to a state border: rather than move to avoid the reform, they may transfer to a school in a neighbouring state that has not yet implemented it, and commute. In column (2) of Table A.7, we exclude all students who live within a 10km radius to a state border (about 27%).[37] As it turns out, the size of the effect becomes larger, presumably since some of these students are allocated to the treatment group

identified reform effect. To rule out this non-random measurement error, we followed the approach by Dahmann and Anger (2014), restricting our sample to students aged 17 and interacting our main effect with monthly dummy variables. We do not find a clear pattern in terms of sign, size, and significance level for these interactions; the point estimate of the main effect remains robust, but its significance is greatly reduced, most likely due to loss of observations (about a quarter of our sample). We take this as evidence that non-random measurement error due to time of interview is, if anything, a minor issue.

[35] Implicitly, we assume that students self-select from the treatment into the control group, as they have a preference to avoid the reform. To be more precise, it is unlikely that students themselves *self*-select; rather, it is their parents who – probably after joint decision-making with their children – decide on taking this action. For simplicity, we refer to students throughout.

[36] In column (a) of Table W.7 in the Web Appendix, we regress the probability of moving on the reform: the effect is small and insignificant. We take this as evidence that the reform has no effect on moving behaviour *per se*.

[37] Similar results are obtained when using a 20km or a 30km radius.

although, in fact, they should be allocated to the control group.[38]

Rather than geographically sorting between schools, students may also sort within them, for example, by skipping a grade in order to avoid the reform. Unfortunately, we do not have information on whether a student skipped a grade. We argue, however, that sorting within schools is more of a theoretical problem for three reasons: first, skipping a grade is not entirely discretionary to students, and requires considerable effort in terms of previous academic achievement. Second, those students who are allowed to skip a grade are presumably those that are the least affected by the reform, and thus have the lowest incentive to avoid it. Finally, skipping a grade leads students to graduate in the same cohort as their former peers, which – in terms of time to graduation – has no advantage. Moreover, as we argue below, this double cohort has certain features that render grade-skipping to avoid the reform an unattractive strategy. Related, students may also sort within schools by repeating a grade. Although this is not a feasible strategy to avoid the reform, it could nevertheless affect our estimates, as students could switch from the control to the treatment group. Assuming that students who must repeat a grade under the old regime are likely to struggle even more under the new one, omitting them would bias our estimates downwards. Again, this issue applies only to a small subset of students, namely those that are in the last pre-treatment cohorts preceding the first treatment ones. Nevertheless, in column (3) of Table A.7, we dig deeper into this issue: here, we exclude all students who repeat a grade (about 7%). We find that the reform still has a negative effect on volunteering, which is significant at the 1% level. As expected, the size of the effect is somewhat reduced.[39]

Finally, we turn to out-of-sample selection: clearly, if dropping out of high school were a deliberate strategy to avoid the reform, it would be the one with the highest opportunity costs, as students would effectively forego their university entrance qualification. In column (4) of Table A.7, we evaluate how drop-outs affect our estimates: here, we exclude all students who drop out of high school (about 8%). As it turns out, the sign, size, and significance level of the

[38]Related, a staggered self-selection of federal states is also thinkable: first, they decide on whether to implement the reform or not; then, they decide on when to implement it. Again, as long as self-selection is not correlated with the outcome, this does not threaten our identification strategy. Moreover, Dahmann and Anger (2014) convincingly show that federal states which implement the reform early do not systematically differ from those that do so late regarding their proportion of high school students, governing party, next election date, and GDP per capita.

[39]As with moving, in column (b) of Table W.7 in the Web Appendix, we regress the probability of repeating a grade on the reform: the effect is small and insignificant.

effect is very similar to that in our preferred specification.[40]

Although the reform has been swiftly integrated into the German secondary education landscape, there may have been various implementation effects – confounding one-off effects arising from the implementation of the reform into regular school business. This is particularly true for students in double, first treatment, and last pre-treatment cohorts, across all federal states.[41] Moreover, in the federal states of *Saxony-Anhalt* and *Mecklenburg-West Pomerania*, students in the first treatment cohorts had already started school when the reform was implemented. For example, for students in the double cohort, such implementation effects may be due to increased competition for secondary and post-secondary educational resources; for students in the first treatment cohort, they may be due to inexperience of teachers, or insecurity on side of students; and for students in the last pre-treatment cohort, they may be due to increased motivation not to repeat a grade, and be affected by the reform. On the other hand, teachers may treat students in these cohorts in a more easy way. Although it is unlikely that such implementation effects are the driving force behind the aggregate effect, they can still affect our estimates.

In columns (5) to (8) of Table A.7, we explore this possibility in more detail: here, we include state-specific controls individually for students in double, first treatment, and last pre-treatment cohorts, as well as for students in the first treatment cohorts in the federal states of *Saxony-Anhalt* and *Mecklenburg-West Pomerania*. If anything, we find that controlling for cohorts that might suffer from implementation effects slightly increases the aggregate effect in our preferred specification. Confounding implementation effects, therefore, seem to be less of an issue.[42]

Triple Differencing

Recall that our difference-in-differences design is both generalised and pseudo: generalised in the sense that treatment can occur at different points in time for different students, pseudo in the sense that each student is observed only once. One implication of this design is that students

[40] Once again, in column (c) of Table W.7 in the Web Appendix, we regress the probability of dropping out on the reform: the effect is small and insignificant. We take this as evidence that the reform has no effect on drop-out behaviour. This is in line with Huebener and Marcus (2015) who find that the reform does not affect drop-out rates.

[41] We define the first treatment cohorts as the cohorts succeeding the double cohorts in order to avoid mixing up implementation effects.

[42] In column (3) of Table W.6 in the Web Appendix, we go even one step further and control for all cohorts that might suffer from implementation effects at the same time: the results remain the same.

who are affected by the reform are observed towards the end and students who are not affected towards the beginning of the observation period, each only once. This means that, at each point in time, students who are treated earlier are compared with students who are treated later. At the end, all students are treated.

Although there is evidence that federal states which implement the reform early do not systematically differ from those that do so late (see Dahmann and Anger (2014), for example) and that time trends play only a minor role, students who are interviewed towards the beginning may still systematically differ from those who are interviewed towards the end of the observation period. To alleviate such concerns, we follow Quis and Reif (2017) and use a triple differencing design as a robustness check. More specifically, we exploit the fact that the reform affected only students in the upper track by employing students in the other tracks as a clean control group. This boils down to the following specification:

$$\begin{aligned} y_{isc,(17-20)} = \beta_0 + \beta_1 (Reform_{sc} \times G) + \beta_2' \mathbf{X_{isc,(17-20)}} + \beta_3 G + \\ + (\sum_{s=1}^{16} \gamma_s State_s \times G) + (\sum_{c=1}^{14} \delta_c Cohort_c \times G) + \varepsilon_{isc,(17-20)} \end{aligned} \tag{2}$$

where G is a dummy variable that equals one if the student belongs to the upper track, and zero else. The remaining coefficients are the same as before.

Table A.8 takes our preferred specification, column (4) in Table A.2, and adds this additional stage of differencing: as can be seen, the reform still has a significant, negative impact on volunteering. The size of the effect is larger, and significance somewhat reduced.

Other Reforms

Over the past two decades, there have been various other reforms in the German secondary education landscape, some of which fall into the observation period, and could potentially be confounding.[43] For example, having long been standard in the majority of states, the remainder has only recently moved towards state-wide harmonised high school finals by introducing central exit examinations. Others, trying to open up the traditionally less permeable and rigid

[43] See Huebener and Marcus (2015) for a detailed overview of these reforms.

German education system, introduced changes to the grade at which tracking takes place, or reduced tracking altogether by combining the lower and intermediate tracks into a single one. Yet others have introduced changes to the choice of subjects available to high school seniors. Probably the biggest change in recent decades, however, has been the abolishment of mandatory military or civil service right after finishing secondary education: in 2011, it was replaced with the (non-mandatory) Federal Volunteer Service.

To be clear, it is unlikely that any of these reforms systematically biases our estimates for two reasons: first, it would have to be correlated with the outcome. More importantly, however, it would have to affect treatment and control group differentially. This would be the case if reforms were correlated, for example, if reducing the number of years required to obtain a high school degree went hand in hand with restricting the subject choice available to high school seniors. Alternatively, one could argue that states which are more prone to reform may be the first to reduce the number of high school years, and may also be inclined to introduce other reforms shortly after, or the other way around.

To rule out this possibility, in columns (1) to (5) of Table A.9, we include state-time-specific controls for these potentially confounding other reforms. As expected, the sign, size, and significance level of the coefficients is very similar to that in our preferred specification. Likewise, excluding students who have already graduated, and who might thus be participating in the Federal Volunteer Service, leaves results unchanged (not shown). Confounding other reforms, therefore, seem to be less of an issue.[44]

Placebo Tests

Finally, as a last exercise, we conduct placebo tests: in columns (1) and (2) of Table A.10, we lag the first treatment cohort by one and two, respectively; in columns (3) and (4), we randomly allocate treatment status to school cohorts and federal states, respectively, keeping the other constant. Finally, in column (5), we completely perturb both school cohorts and federal states, and then randomly allocate treatment status. As can be seen, none of the coefficients is significant at any conventional level. For the first two columns, we can see that the coefficients

[44] In column (4) of Table W.6 in the Web Appendix, we go even one step further and control for all potentially confounding reforms at the same time: the results remain the same.

are negative, pointing towards the overall trend reversal in volunteering we see during the observation period; the fact that the coefficient of the second column is slightly larger than that of the first one suggests that there are no *ex-ante* behavioural changes due to anticipation effects (Ashenfelter's dip). Note that in both of these columns, we lose observations that fall out of the observation period window. For the last three columns, we cannot observe a clear pattern of coefficients.

5. Discussion and Policy Implications

We find robust empirical evidence that raising instructional time has the potential to significantly affect student leisure activities and behaviours, and in particular, to significantly crowd out student pro-social behaviour, a behaviour that is linked to various positive outcomes – both at the societal and individual level – and that parents, educators, and policy-makers alike would otherwise consider worth promoting. In the given context, a rise in weekly instructional hours of about 13% had a negative and sizeable effect on volunteering, reducing the share of students who volunteer at least once a month by about 19 percent. In other words, it led almost every fifth student to change her behaviour from volunteering at least monthly to volunteering less often or not at all. Students who volunteer on a regular basis are most adversely affected, and there is some evidence that those from disadvantaged backgrounds are more likely to disengage. While two thirds of students cut back on their activities, one third give them up completely. We find no similar crowding out of involvement in activities within school, but no substitution either. Finally, there is evidence that raising instructional time also has the potential to affect political interest, which we take as a proxy for political behaviour.

Probably the most important remaining question, especially from a policy perspective, is whether the identified effects are permanent or temporary. The fact that impacts seem to be driven by a reduction in available leisure time as opposed to a rise in intensity of instruction may already point into the direction of temporary effects: once students graduate and more leisure time becomes available, we would expect that they – at least temporarily – return to their baseline level of volunteering, or even exceed it. To put this to a more formal test, we compare the raw mean levels of volunteering, our main outcome, for students in the treatment group with those in the control group, respectively, before and after graduation: as expected, for students in the control group, we observe only a small change in volunteering levels after graduating, from a share of about 36% to 39%; for students in the treatment group, however, we observe a pronounced change, from a share of about 30% to 40%, or in other words, from a lower baseline level to about the same level as the control group. One caveat of this tabulation is that sample size reduces quite drastically for post-graduation raw means: for students in the treatment group, there are only ten observations left (for students in the control group, there are 76). Extending our tabulation to include individuals not used in our regressions (i.e. individuals with missings on observables) leads to similar raw mean levels of volunteering for students in the control group (37% pre-graduation and 39% post-graduation), but quite different ones for students in the treatment group: here, the share of students who volunteer at least monthly increases from about 32% pre-graduation to 48% post-graduation, suggesting that students may more than offset a lack of volunteering during school.[45] This is in line with Meyer and Thomsen (2015) who find that, for students affected by the reform, university enrolment in the year of graduation decreases while participation in voluntary service or staying abroad increases. Taken together, these results suggest that the identified effects are only temporary.

Why are our findings important? First of all, in the given context, they are important because of the large number of students affected. In Germany, in school year 2014/15 alone, of 2,329,990 high school students in total (Federal Statistical Office, 2016b), about 785,000 volunteered at least monthly. We estimate that the rise in instructional time decreases this share by about 149,000: 99,000 cut back on their activities, and 50,000 give them up completely. It is difficult to measure the economic value of volunteering for society: there exist various definitions of volunteering, and at least as many ways to measure it, for example through national accounts, labour force surveys, or social and time use surveys. It is clear, however, that this value is substantial.[46] Through time use surveys, the OECD estimates the economic value of volunteering for Germany in 2013 to be around USD 117.6 billion or 3.3% of real GDP (OECD, 2015).[47] We can calculate back-of-the-envelope that losing between 50,000 and 149,000 volunteers is equal to losing volunteer work worth between USD 71.5 million and USD 213.2 million.[48] These figures are likely to be lower bounds for two reasons: first, volunteering in the

[45] In this tabulation, sample size for students in the post-graduation treatment group increases from ten to 202.

[46] See The Economist (2014) for a recent feature.

[47] This figure is roughly comparable to the UK (2.5%) and to the US (3.7%).

[48] There were 82.2 million people living in Germany in 2015 (Federal Statistical Office, 2016a). Thus, assuming the distribution of activities in the general population is similar to that in the population under scrutiny, the loss in volunteer work can be calculated as $(117,600,000,000 \times 50,000)/82,200,000$ and $(117,600,000,000 \times$

general population is less prevalent than in the population under scrutiny.[49] Second, to the extent that volunteering during youth and adolescence contributes to habit formation (Hart et al., 2007) and has positive peer effects (Wilson and Musick, 1997), impacts may be permanent rather than temporary. The available, tentative evidence so far, however, suggests that this is rather not the case, but this point warrants further investigation.

Besides negative effects for society *per se*, the decrease in volunteer work can also have negative micro implications: a growing body of evidence documents the importance of volunteering for individual labour market outcomes. For example, in a recent correspondence testing study, Baert and Vujić (2016) show that job seekers who indicate volunteering on their resumes receive one third more interview invitations, and that this volunteering premium is higher for women. A leading professional social network, LinkedIn (2016), using data on members, estimates that one in five managers hire someone because of their volunteering experience. Sauer (2015), using a structural model and longitudinal data for the US, estimates that an extra year of pro-social engagement increases wage offers in future full-time (part-time) work by 2.6% (8.5%) for women aged 25 to 55, in line with Freeman (1997) who estimates that volunteering raises paid work hours by between 3% and 7%. Remarkably similar, using a Mincerian earnings regression with Heckman's two-step procedure to correct for self-selection into employment, we regress monthly net individual income at age 24 to 33 on volunteering at age 17, and find that volunteering in youth is associated with higher earnings in adulthood of about 7% (results available upon request).[50] There is evidence that being engaged from an early age

$149,000)/82,200,000$, respectively, for the 50,000 students giving up and for the 149,000 students cutting back *and* giving up their activities.

[49] In the general population, only about 23% of individuals report to volunteer at least once a month, according to the OECD. In the SOEP, this share is even lower: 20%. Again, both figures are roughly comparable to the UK (18%) and to the US (30%).

[50] We estimate the following Mincerian earnings regression:

$$ln(y_{i,(24-33)}) = \beta_0 + \beta_1 Volunteering_{i,17} + \beta_2' \mathbf{X_{i,(24-33)}} + \mu \hat{\lambda_{i,(24-33)}} + \varepsilon_{i,(24-33)}$$

where y is the log monthly net individual income of individual i, measured at age 24 to 33, and X is a vector of controls, including dummy variables for age, gender, marital status (single, partnered, or married), educational attainment (secondary or tertiary degree), employment status (in training, full-time employed, part-time employed, irregularly employed, on parental leave, or unemployed). The age interval is determined by two factors: the upper bound is determined by data availability (volunteering at age 17 is only available for a limited number of individuals), the lower bound by the average age of entry into the labour force of students with a first professional qualification. $\hat{\lambda_{i,(24-33)}} = \frac{\phi(z_i'\gamma)}{\Phi(z_i'\gamma)}$ is the inverse Mill's ratio obtained from an auxiliary probit regression that estimates whether an individual works or not. Our regressor of interest is β_1, which measures the percentage change in adult earnings due to youth volunteering.

on enhances psychological development by raising self-esteem and self-confidence and by discouraging risky behaviours (Hart et al., 2007; Wilson and Musick, 2012). The physical and mental health benefits of volunteering (Wilson and Musick, 2012), as well as its subjective well-being returns are well documented (Binder and Freytag, 2013; Borgnonovi, 2008; Meier and Stutzer, 2008). Finally, to the extent that students from disadvantaged backgrounds are disproportionally affected, the role that volunteering can play in the production process of skills, for example, through generating early life skills that complement other skills later on (Cunha and Heckman, 2007; Fuchs, 2016), or in the selection process for further education, as is for example the case in the German scholarship system or for admissions to US colleges, the decrease in volunteering for these groups might further increase educational inequalities.[51]

Raising instructional time is often found to have positive impacts on student learning and performance, especially when the additional time is used effectively, and there surely is an optimal amount of weekly instructional hours that balances student learning with student leisure activities and behaviours. For a complete cost-benefit account of raising instructional time, however, its impacts on student leisure activities and behaviours, in particular on beneficial behaviours such as volunteering, should be taken into account. In view of our findings, education policy could consider lowering access barriers to volunteering. In particular, given that we do not find a similar crowding out of involvement in activities within school as for volunteering outside school, it could provide alternative activities for volunteering such as high school community service within schools, or encourage it through the curriculum, for example, by introducing volunteering days.

There are many limitations to this study, which is only a cautious exploration into the relationship between instructional time and student pro-social behaviour. The most obvious is that we can only present tentative results on how persistent the identified effects are. The fact that raw mean levels of volunteering of affected students converge to, or even exceed, those of unaffected students, however, suggests that they are only temporary. Once more data become available, it would be interesting to test this more formally. Moreover, in unreported regressions, we do not find that raising instructional time crowds out other student leisure activities and behaviours such as playing music or doing sports, and an obvious question would be why

[51] See The Behavioural Insights Team (2016) for a recent impact evaluation of programmes that promote social action: it shows that such programmes can nurture skills such as empathy or grit that are critical for educational success.

volunteering, in particular, is affected. Access barriers might play a role here. Finally, external validity is an issue. The fact that the UK and the US exhibit similar profiles regarding instructional time and volunteering demographics than Germany (Bureau of Labor Statistics, 2015; OECD, 2015), however, might point towards the fact that findings are rather transferable.

Acknowledgements

I would like to thank Claudia Senik, Andrew Clark, Katrin Rehdanz, Stephen Gibbons, Gert Wagner, Nicolas Ziebarth, Heike Solga, Sarah Dahmann, Mathias Hübener, Christopher Wratil, and especially Jürgen Schupp for valuable comments and suggestions. I am also thankful to Jan Göbel for continuous support with the SOEP. Falk Voit provided excellent research assistance. All remaining shortcomings and errors are my own.

Project Funding

German Ministry for Education and Science
FKZ: NIMOERT2/Kassenzeichen: 8103036999784/#30857
Projekt: ”Nicht-monetäre Erträge von Bildung in den Bereichen Gesundheit, nicht-kognitive Fähigkeiten sowie gesellschaftliche und politische Partizipation”

Aknin, L. B., C. P. Barrington-Leigh, E. W. Dunn, J. F. Helliwell, J. Burns, R. Biswas-Diener, I. Kemeza, P. Nyende, and C. E. Ashton-James (2008). Prosocial Spending and Well-Being: Cross-Cultural Evidence for a Psychological Universal. *Science 319*(5870), 1687–1688.

Andrietti, V. (2016). The Causal Effects of an Intensified Curriculum on Cognitive Skills: Evidence from a Natural Experiment. *Universidad Carlos III de Madrid Working Paper, Economic Series 16-06*.

Baert, S. and S. Vujić (2016). Does it Pay to Care? Pro-Social Engagement and Employment Opportunities. *IZA Discussion Paper 9649*.

Batson, C. D. and A. A. Powell (2003). Altruism and Prosocial Behavior. In I. B. Weiner (Ed.), *Handbook of Psychology*, Volume 5. London: Wiley.

Bedard, K. and E. Dhuey (2006). The Persistence of Early Childhood Maturity: International Evidence of Long-Run Age Effects. *Quarterly Journal of Economics 121*(4), 1437–1472.

Bellei, C. (2009). Does Lengthening the School Day Increase Students' Academic Achievement? Results from a Natural Experiment in Chile. *Economics of Education Review 28*(5), 629–640.

Binder, M. and A. Freytag (2013). Volunteering, Subjective Well-Being, and Public Policy. *Journal of Economic Psychology 34*, 97–119.

Black, S. E., P. J. Devereux, and K. G. Salvanes (2011). Too Young to Leave the Nest? The Effects of School Starting Age. *Review of Economics and Statistics 93*(2), 455–467.

Borgnonovi, F. (2008). Doing well by doing good. The relationship between formal volunteering and self-reported health and happiness. *Social Science & Medicine 66*, 2321–2334.

Büttner, B. and S. L. Thomsen (2015). Are We Spending Too Many Years in School? Causal Evidence of the Impact of Shortening Secondary School Duration. *German Economic Review) 16*(1), 65–86.

Bureau of Labor Statistics (2015). Volunteering in the United States, 2015. `http://www.bls.gov/news.release/volun.nr0.htm`.

Cattaneo, A., C. Oggenfuss, and S. C. Wolter (2016). The More, The Better? The Impact of Instructional Time on Student Performance. *CESifo Working Paper 5813*.

Cortes, K. E. and J. S. Goodman (2014). Ability-Tracking, Instructional Time, and Better Pedagogy: The Effect of Double-Dose Algebra on Student Achievement. *American Economic Review 104*(5), 400–405.

Cortes, K. E., J. S. Goodman, and T. Nomi (2015). Intensive Math Instruction and Educational Attainment: Long-Run Impacts of Double-Dose Algebra. *Journal of Human Resources 50*(1), 108–158.

Cunha, F. and J. Heckman (2007). The Technology of Skill Formation. *American Economic Review 97*(2), 31–47.

Dahmann, S. (2017). How Does Education Improve Cognitive Skills? Instructional Time versus Timing of Instruction. *Labour Economics*.

Dahmann, S. and S. Anger (2014). The Impact of Education on Personality: Evidence from a German High School Reform. *IZA Discussion Paper 8139*.

Dee, T. S. (2004). Are There Civic Returns to Education? *Journal of Public Economics 88*(9–10), 1697–1720.

Dhillon, A. and S. Peralta (2002). Economic Theories of Voter Turnout. *Economic Journal 112*(480), F332–F352.

Dunn, E. W., L. B. Aknin, and M. Norton (2008). Spending Money on Others Promotes Happiness. *Science 319*(5870), 1687–1688.

Eisenberg, N., T. L. Spinrad, and A. S. Morris (2013). Prosocial Development. In P. D. Zelazo (Ed.), *Oxford Handbook of Developmental Psychology*, Volume 2. Oxford: Oxford University Press.

Evren, . and S. Minardi (2017). Warm-glow Giving and Freedom to be Selfish. *Economic Journal 127*(603), 1381–1409.

Federal Agency for Cartography and Geodesy (2016). Administrative Areas 1:2,500,000. `http://www.geodatenzentrum.de`.

Federal Statistical Office (2016a). Genesis Online, Table 12411-0001 - Population: Germany, Effective Date December 31, 2013. `http://www.destatis.de`.

Federal Statistical Office (2016b). Genesis Online, Table 21111-0003 - Students: Federal States, School Year 2013/14, Gender, School Type. `http://www.destatis.de`.

Fehr, E. and S. Gachter (2000). Fairness and Retaliation: The Economics of Reciprocity. *Journal of Economic Perspectives 14*(3), 159–181.

Fishbein, M. and I. Ajzen (1975). *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research*. Addison-Wesley.

Freeman, R. B. (1997). Working for Nothing: The Supply of Volunteer Labor. *Journal of Labor Economics 15*(1), S140–S166.

Fuchs, B. (2016). The Effect of Teenage Employment on Character Skills, Expectations, and Occupational Choice Strategies. *Hohenheim Discussion Papers in Business, Economics, and Social Sciences 14-2016*.

Gibson, J. (2001). Unobservable Family Effects and the Apparent External Benefits of Education. *Economics of Education Review 20*(3), 225–233.

Hart, D., T. M. Donnelly, J. Youniss, and R. Atkins (2007). High School Community Service as a Predictor of Adult Voting and Volunteering. *American Educational Research Journal 44*(1), 197–219.

Heckman, J. J., R. J. LaLonde, and J. A. Smith (2009). The Economics and Econometrics of Active Labor Market Programs. *Handbook of Labor Economics 3*.

Helliwell, J. F., H. Huang, and S. Wang (2011). New Evidence on Trust and Well-being. *National Bureau of Economic Research Working Paper* (22450).

Henderson, J. and S. Chatfield (2011). Who Matches? Propensity Scores and Bias in the Causal Effects of Education on Participation. *Journal of Politics 73*(3), 646–658.

Herrmann, M. A. and J. E. Rockoff (2012). Worker Absence and Productivity: Evidence from Teaching. *Journal of Labor Economics 30*(4), 749–782.

Hillygus, D. S. (2005). The Missing Link: Exploring the Relationship Between Higher Education and Political Behavior. *Political Behavior 27*(1), 25–47.

Homuth, C. (2012). Der Einfluss des achtjaehrigen Gymnasiums auf den Kompetenzerwerb. *Bamberg Graduate School of Social Sciences*.

Homuth, C. (2017). *Die G8-Reform in Deutschland*. Springer.

Huebener, M., S. Kuger, and J. Marcus (2017). Increased instruction hours and the widening gap in student performance. *Labour Economics*.

Huebener, M. and J. Marcus (2015). Moving up a Gear: The Impact of Compressing Instructional Time Into Fewer Years of Schooling. *DIW Berlin Discussion Papers 1450*.

Imbens, G. W. and J. M. Wooldridge (2009). Recent Developments in the Econometrics of Program Evaluation. *Journal of Economic Literature 47*(1), 5–86.

Kosse, F., T. Deckers, H. Schildberg-Hörisch, and A. Falk (2014). Formation of Human Pro-Sociality: Causal Evidence on the Role of Social Environment. *mimo*.

Lavy, V. (2015). Do Differences in Schools' Instruction Time Explain International Achievement Gaps? Evidence from Developed and Developing Countries. *Economic Journal 125*(588), F397–F424.

LinkedIn (2016). LinkedIn. `http://www.linkedin.com`.

Lochner, L. (2011). Non-Production Benefits of Education: Crime, Health, and Good Citizenship. In E. Hanushek, S. Machin, and L. Woessmann (Eds.), *Handbook of the Economics of Education*, Volume 4. Amsterdam: Elsevier.

Meier, S. and A. Stutzer (2008). Is Volunteering Rewarding in Itself? *Economica 75*(297), 39–59.

Meyer, T. and S. L. Thomsen (2015). Schneller fertig, aber weniger Freizeit? Eine Evaluation der Wirkungen der verkürzten Gymnasialzeit auf die außerschulischen Aktivitäten der Schülerinnen und Schüler. *Schmollers Jahrbuch 135*, 249–278.

Meyer, T., S. L. Thomsen, and H. Schneider (2015). New Evidence on the Effects of the Shortened School Duration in the German States: An Evaluation of Post-Secondary Education Decisions. *IZA Discussion Paper 9507*.

Milligan, K., E. Moretti, and P. Oreopoulos (2004). Does Education Improve Citizenship? Evidence from the United States and the United Kingdom. *Journal of Public Economics 88*(9–10), 1667–1695.

OECD (2011). How's Life? Measuring Well-Being. *OECD Publishing*.

OECD (2014). Education at a Glance 2014. *OECD Publishing*.

OECD (2015). How's Life? Measuring Well-Being. *OECD Publishing*.

OECD (2016). Student Learning Time: A Literature Review. *OECD Education Working Paper* (127).

Oreopoulos, P. and K. G. Salvanes (2011). Priceless: The Non-Pecuniary Benefits of Schooling. *Journal of Economic Perspectives 25*(1), 159–184.

Patall, E. A., H. Cooper, and A. B. Allen (2010). Extending the School Day or School Year: A Systematic Review of Research (1985–2009). *Review of Educational Research 80*(3), 401–436.

Pelkonen, P. (2012). Length of Compulsory Education and Voter Turnout: Evidence from a Staggered Reform. *Public Choice 150*(1), 51–75.

Persson, M. (2014). Testing the Relationship Between Education and Political Participation Using the 1970 British Cohort Study. *Political Behavior 36*(4), 877–897.

Putnam, R. D. (2000). *Bowling Alone: The Collapse and Revival of American Community*. "New York": Simon and Schuster.

Quis, J. S. (2015). Does higher learning Intensity affect student well-being? Evidence from the National Educational Panel Study. *BERG Working Paper Series 94*.

Quis, J. S. and S. Reif (2017). Health Effects of Instruction Intensity – Evidence from a Natural Experiment in German High-Schools. *BERG Working Paper Series 123*.

Rivkin, S. G. and J. C. Schiman (2015). Instruction Time, Classroom Quality, and Academic Achievement. *Economic Journal 125*(588), F425–F448.

Sauer, R. M. (2015). Does it Pay for Women to Volunteer? *International Economic Review 56*(2), 537–564.

Süddeutsche Zeitung (2010). Unterricht, der krank macht. `http://www.sueddeutsche.de/karriere/stress-durch-ganztagsschulen-unterricht-der-krank-macht-1.942372`.

Siedler, T. (2010). Schooling and Citizenship in a Young Democracy: Evidence from Postwar Germany. *Scandinavian Journal of Economics 112*(2), 315–338.

Sondheimer, R. M. and D. P. Green (2010). Using Experiments to Estimate the Effects of Education on Voter Turnout. *American Journal of Political Science 54*(1), 174–189.

Standing Conference of the Ministers of Education and Cultural Affairs (2016). High School (German only). `https://www.kmk.org/themen/allgemeinbildende-schulen/bildungswege-und-abschluesse/sekundarstufe-ii-gymnasiale-oberstufe-und-abitur.html`.

Taylor, E. (2014). Spending More of the School Day in Math Class: Evidence from a Regression Discontinuity in Middle School. *Journal of Public Economics 117*, 162–181.

The Behavioural Insights Team (2016). Evaluating Youth Social Action: Does Participating in Social Action Boost the Skills Young People Need to Succeed in Adult Life? *Final Report*.

The Economist (2014). The Economics of Volunteering: Hiding in Plain Sight. http://www.economist.com/blogs/freeexchange/2014/09/economics-volunteering.

Thiel, H., S. L. Thomsen, and B. Büttner (2014). Variation of learning intensity in late adolescence and the effect on personality traits. *Journal of the Royal Statistical Society: Series A (Statistics in Society) 177*(4), 861–892.

Wagner, G. G., J. R. Frick, and J. Schupp (2007). The German Socio-Economic Panel Study (SOEP) – Scope, Evolution, and Enhancements. *Schmollers Jahrbuch 127*(1), 139–169.

Wagner, G. G., J. Goebel, P. Krause, R. Pischner, and I. Sieber (2008). Das Sozio-Oekonomische Panel (SOEP): Multidisziplinaeres Haushaltspanel und Kohortenstudie fuer Deutschland – Eine Einfuehrung (fuer neue Datennutzer) mit einem Ausblick (fuer erfahrene Anwender). *AStA Wirtschafts- und Sozialstatistisches Archiv 2*(4), 301–328.

Wilson, J. and M. Musick (1997). Who Cares? Toward an Integrated Theory of Volunteer Work. *American Sociological Review 62*(5), 694–713.

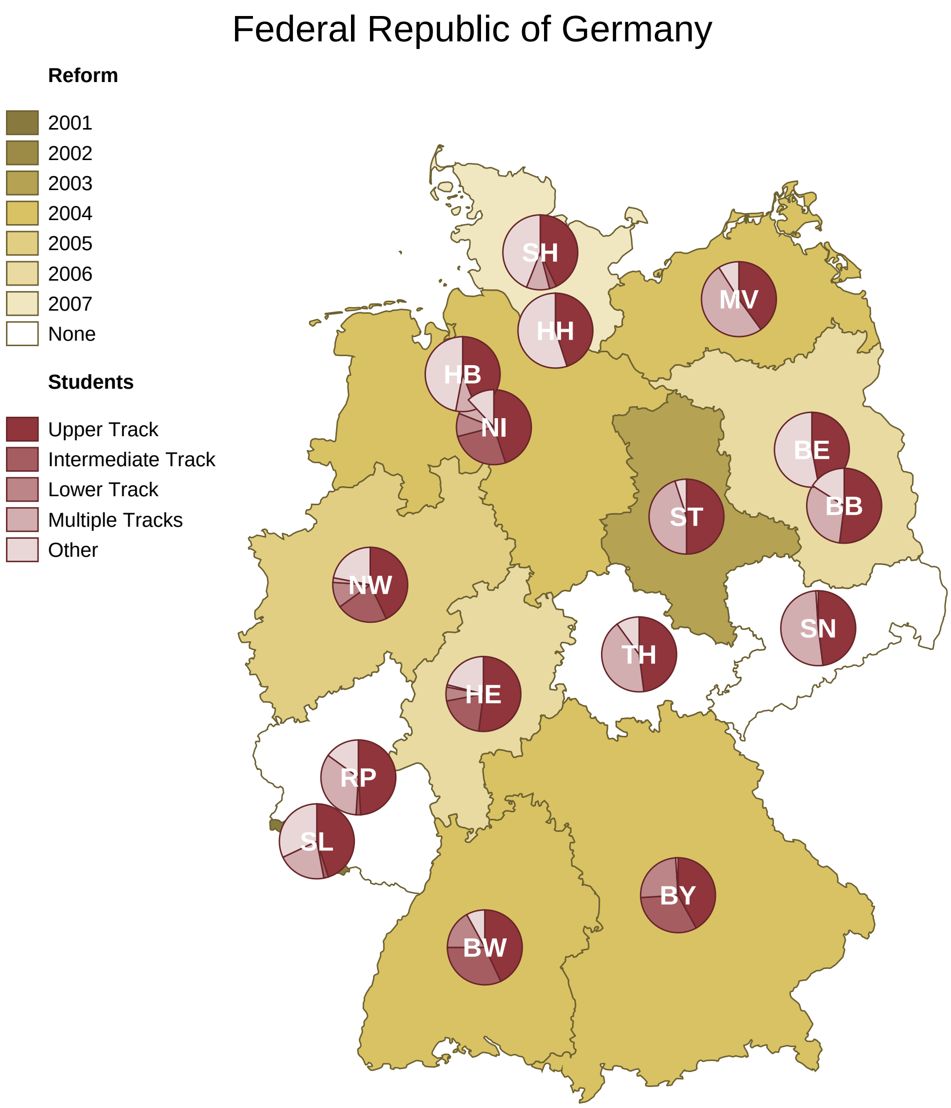
Wilson, J. and M. Musick (2012). The Effects of Volunteering on the Volunteer. *Law and Contemporary Problems 62*(4), 141–168.

Woessmann, L. (2003). Schooling Resources, Educational Institutions, and Student Performance: The International Evidence. *Oxford Bulletin of Economics and Statistics 65*(2), 117–170.

Woessmann, L. (2016). The Importance of School Systems: Evidence from International Differences in Student Achievement. *IZA Discussion Paper 10001*.

Appendix

Figure A.1: Implementation of Reform, Variation Across States and Over Time



Note: The figure shows variation in the implementation of the reform across states and over time. It also reports the shares of students in the different tracks for each state, as of school year 2014/15. The category *multiple tracks* includes students in schools combining the *intermediate* and *lower track*; *other* includes students in comprehensive and Waldorf schools.

The states are *Baden-Wuerttemberg* (BW), *Bavaria* (BY), *Berlin* (BE), *Brandenburg* (BB), *Bremen* (HB), *Hamburg* (HH), *Hesse* (HE), *Lower Saxony* (NI), *Mecklenburg-West Pomerania* (MV), *North Rhine-Westphalia* (NW), *Rhineland-Palatinate* (RP), *Saarland* (SL), *Saxony* (SN), *Saxony-Anhalt* (ST), *Schleswig-Holstein* (SH), and *Thuringia* (TH).

Sources: Federal Agency for Cartography and Geodesy (2016), Federal Statistical Office (2016b), Standing Conference of the Ministers of Education and Cultural Affairs (2016), own calculations

Pro-Social Behaviour

Figure A.2: Pro-Social Behaviour, *Outside of School*, Over Time

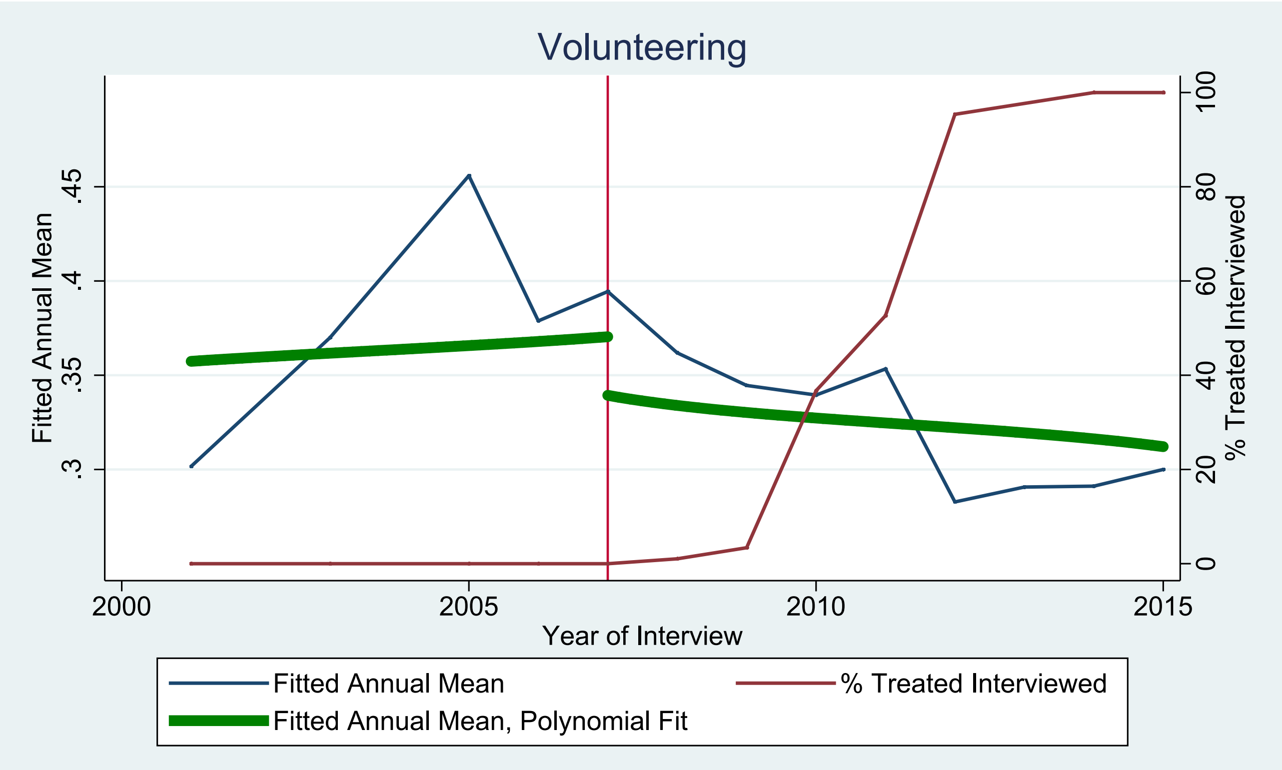
Note: The figure shows the fitted annual mean of volunteering, covariate-adjusted for demographic, educational, and parental household characteristics, as well as for sub-samples and a full set of federal state and school cohort dummy variables.

See Section 2.1 for a description of the variables used.

Source: SOEP, 2001-2015, students (upper track) aged 17 to 20, own calculations

Graphical Evidence

Figure A.3: Graphical Evidence - Pro-Social Behaviour, *Outside of School*, Over Time, 1 of 2

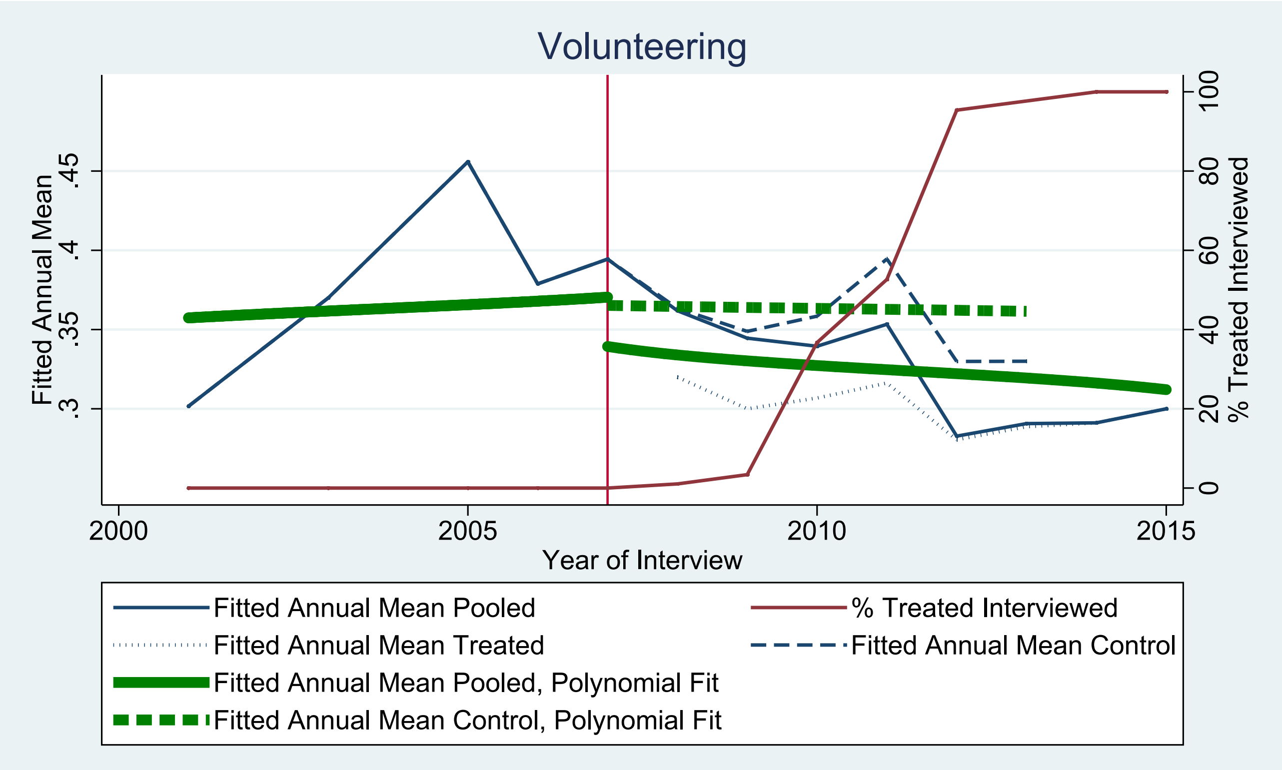


Note: The figure shows the fitted annual mean of volunteering, covariate-adjusted for demographic, educational, and parental household characteristics, as well as for sub-samples and a full set of federal state and school cohort dummy variables.

See Section 2.1 for a description of the variables used.

Source: SOEP, 2001-2015, students (upper track) aged 17 to 20, own calculations

Figure A.4: Graphical Evidence - Pro-Social Behaviour, *Outside of School*, Over Time, 2 of 2

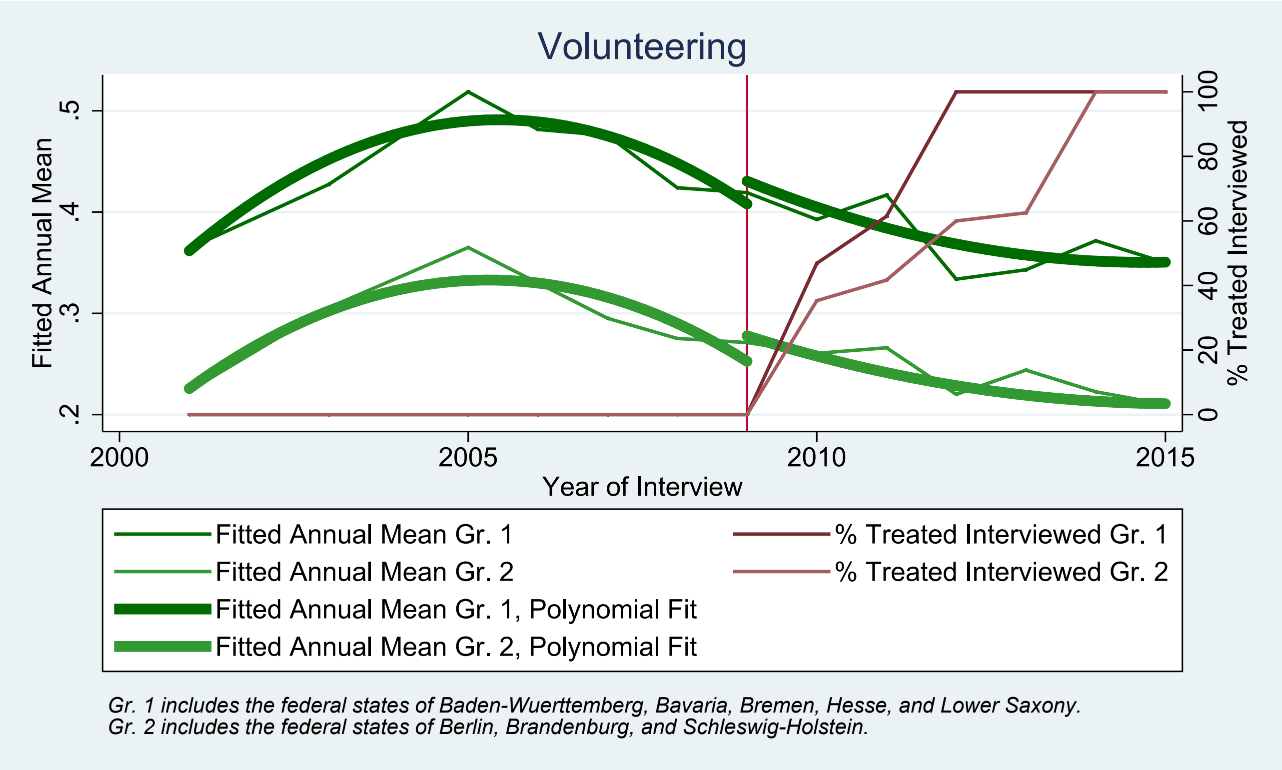


Note: The figure shows the fitted annual mean of volunteering, covariate-adjusted for demographic, educational, and parental household characteristics, as well as for sub-samples and a full set of federal state and school cohort dummy variables.

See Section 2.1 for a description of the variables used.

Source: SOEP, 2001-2015, students (upper track) aged 17 to 20, own calculations

Figure A.5: Graphical Evidence - Pro-Social Behaviour, *Outside of School*, Common Trend, 1 of 2

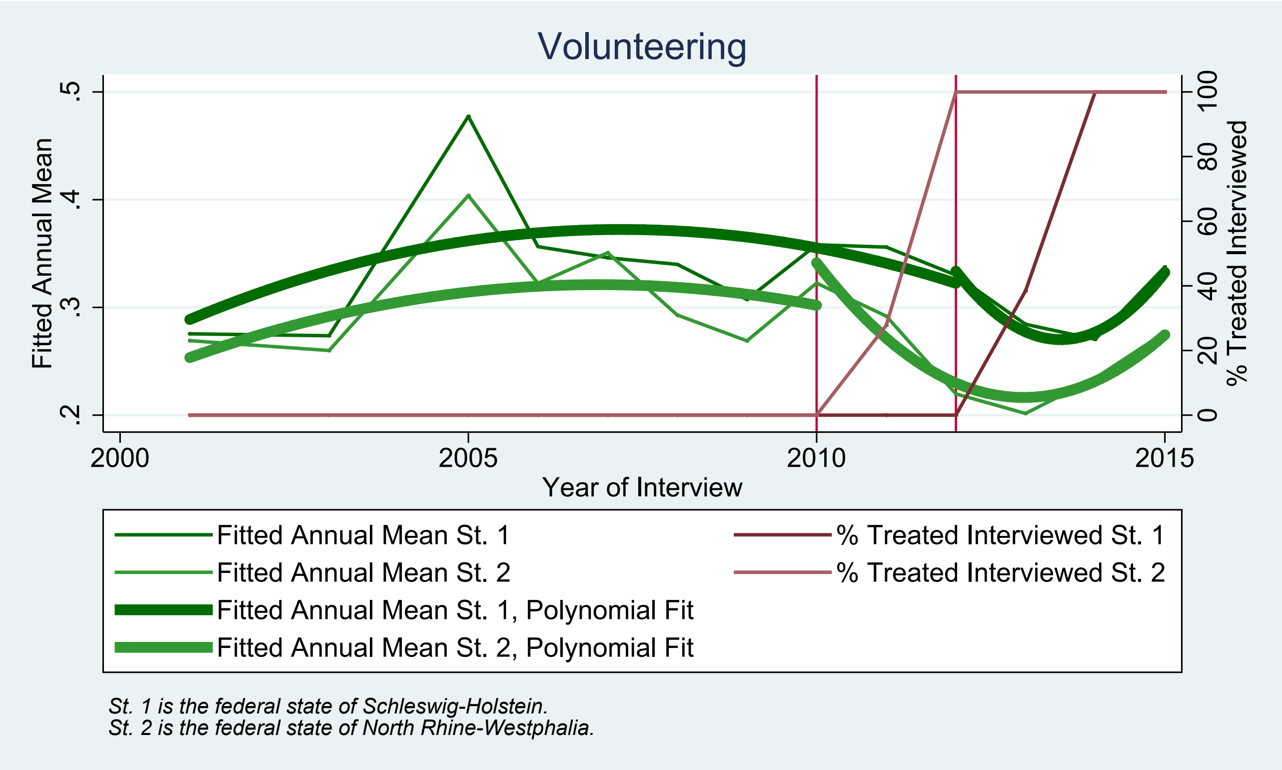


Note: The figure shows the fitted annual mean of volunteering, covariate-adjusted for demographic, educational, and parental household characteristics, as well as for sub-samples and a full set of federal state and school cohort dummy variables.

See Section 2.1 for a description of the variables used.

Source: SOEP, 2001-2015, students (upper track) aged 17 to 20, own calculations

Figure A.6: Graphical Evidence - Pro-Social Behaviour, *Outside of School*, Common Trend, 2 of 2

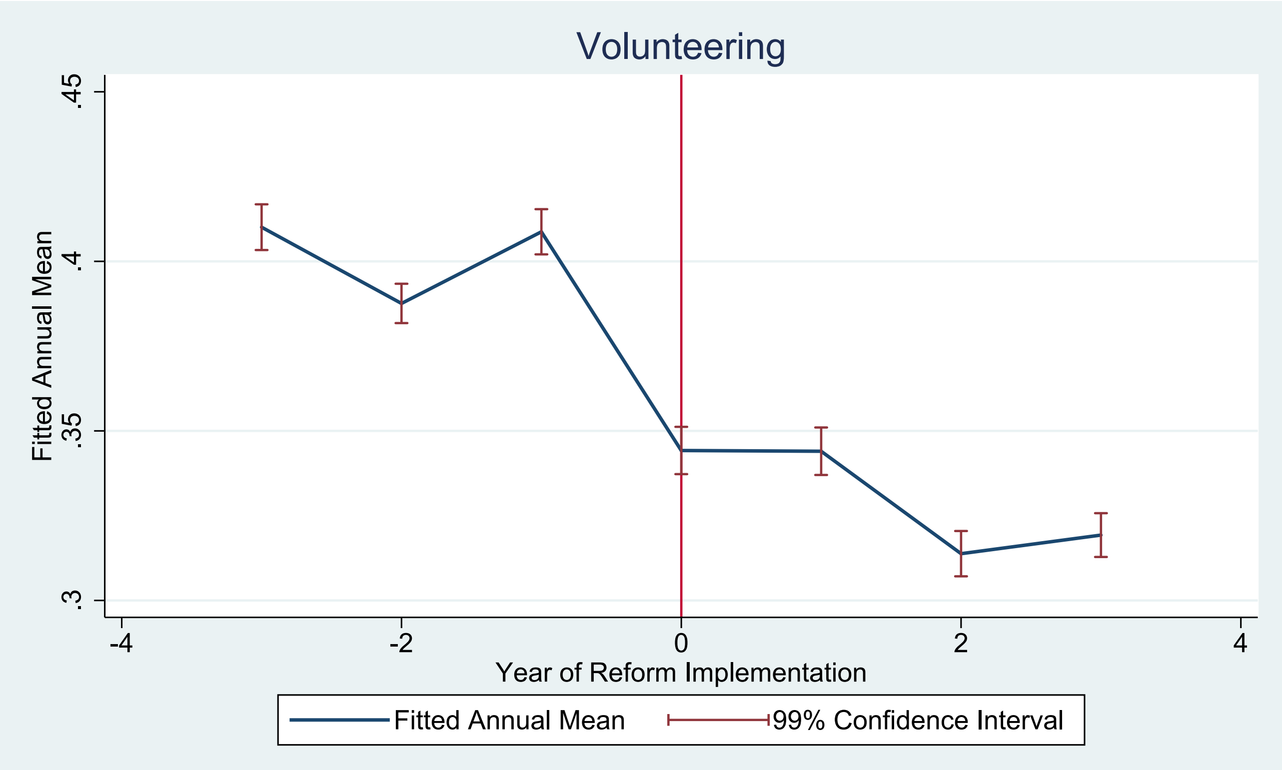


Note: The figure shows the fitted annual mean of volunteering, covariate-adjusted for demographic, educational, and parental household characteristics, as well as for sub-samples and a full set of federal state and school cohort dummy variables.

See Section 2.1 for a description of the variables used.

Source: SOEP, 2001-2015, students (upper track) aged 17 to 20, own calculations

Figure A.7: Baseline Results - Pro-Social Behaviour, *Outside of School*, Event Study

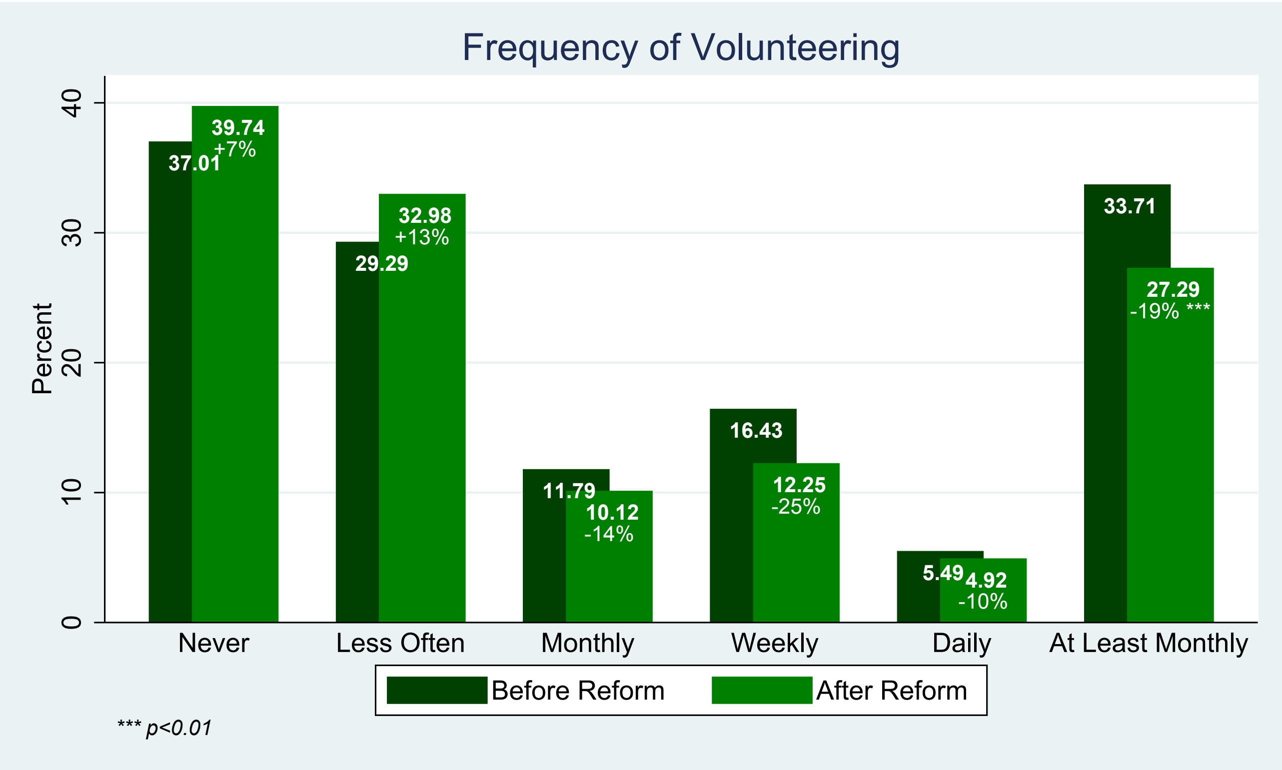


Note: The figure shows the fitted annual mean of volunteering, covariate-adjusted for demographic, educational, and parental household characteristics, as well as for sub-samples and a full set of federal state and school cohort dummy variables.

See Section 2.1 for a description of the variables used.

Source: SOEP, 2001-2015, students (upper track) aged 17 to 20, own calculations

Figure A.8: Baseline Results - Pro-Social Behaviour, *Outside of School*, Change in Distribution



Note: The figure shows the change in the frequency distribution of volunteering due to the reform. The respective change is estimated from a separate regression in which the respective frequency of volunteering serves as the outcome. The regressions routinely control for demographic, educational, and parental household characteristics, as well as for sub-samples and for federal states and school cohorts.

See Section 2.1 for a description of the variables used.

Source: SOEP, 2001-2015, students (upper track) aged 17 to 20, own calculations

Balancing on Observables

Table A.1: Descriptive Statistics

		Mean		
Variables	Mean	Treatment Group	Control Group	Normalised Difference
Age	17.4911	17.1187	17.7881	0.5759
Has Graduated	0.0384	0.0101	0.0610	0.1964
Is Female	0.5335	0.5473	0.5225	0.0352
Has Migration Background	0.2040	0.2384	0.1766	0.1081
Lives in East	0.1254	0.1127	0.1356	0.0492
Lives in Countryside	0.2835	0.2998	0.2705	0.0459
Parent Has Tertiary Degree	0.5214	0.4738	0.5594	0.1214
Parent is Blue-Collar Worker	0.1246	0.1066	0.1388	0.0694
Parent is Full-Time Employed	0.6438	0.5674	0.7047	0.2037
Parent is Single	0.0429	0.0604	0.0291	0.1074
Is Only Child	0.1701	0.1751	0.1661	0.0168
Number of Children in Household	2.4277	2.5412	2.3371	0.1308
Number of Observations	2,240	994	1,246	-

Note: The last column shows the normalised difference, which is calculated as $\triangle x = (\bar{x}_t - \bar{x}_c) \div \sqrt{\sigma_t^2 + \sigma_c^2}$, where $\bar{x}_t$ and $\bar{x}_c$ is the sample mean of the covariate for the treatment and control group, respectively. σ^2 denotes the variance. As a rule of thumb, a normalised difference greater than 0.25 indicates a non-balanced covariate, which might lead to sensitive results (Imbens and Wooldridge, 2009).

See Section 2.1 for a description of the variables used.

Source: SOEP, 2001-2015, students (upper track) aged 17 to 20, own calculations

Baseline Results

Table A.2: Baseline Results - Pro-Social Behaviour, *Outside of School*

	Volunteering			
Regressors	(1)	(2)	(3)	(4)
Reform	-0.0728***	-0.0658***	-0.0657***	-0.0642***
	(0.0135)	(0.0136)	(0.0136)	(0.0168)
Age 18		0.0386	0.0382	0.0368
		(0.0328)	(0.0327)	(0.0341)
Age 19		0.0885*	0.0846*	0.0858*
		(0.0420)	(0.0425)	(0.0427)
Age 20		0.0593	0.0410	0.0413
		(0.0495)	(0.0530)	(0.0468)
Has Graduated			-0.0437	-0.0500*
			(0.0254)	(0.0267)
Other Demographic Characteristics	No	No	No	Yes
Parental Characteristics	No	No	No	Yes
Household Characteristics	No	No	No	Yes
Number of Observations	2,240	2,240	2,240	2,240
R^2	0.0534	0.0555	0.0557	0.0776
Adjusted R^2	0.0327	0.0335	0.0333	0.0513

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

Note: All models include a constant, controls for sub-samples, and a full set of federal state and school cohort dummy variables. All figures are rounded to four decimal places.

See Section 2.1 for a description of the variables used.

Source: SOEP, 2001-2015, students (upper track) aged 17 to 20, own calculations

Table A.3: Baseline Results - Pro-Social Behaviour, *Inside of School*

	Scholastic Involvement							
Regressors	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
Reform	0.0049	0.0249	-0.0162	-0.0075	0.0426	0.0377	-0.0091	0.0016
	(0.0115)	(0.0329)	(0.0357)	(0.0462)	(0.0616)	(0.0415)	(0.0480)	(0.0425)
Demographic Characteristics [a]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Household Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	1,958	1,958	1,958	1,958	1,958	1,958	1,958	1,958
R^2	0.0362	0.0351	0.0332	0.1018	0.0881	0.0376	0.0428	0.0390
Adjusted R^2	0.0063	0.0052	0.0031	0.0739	0.0597	0.0077	0.0131	0.0091

[a] Including *Age 18*, *Age 19*, *Age 20*, and *Has Graduated*

(a) Student Representative, (b) Class Representative, (c) School Magazine, (d) Drama or Dance Group,
(e) Choir or Orchestra, (f) Sports Group, (g) Other Voluntary Group, (h) None

Robust standard errors clustered at the federal state level in parentheses
**** $p<0.01$, ** $p<0.05$, * $p<0.1$*

Note: All models include a constant, controls for sub-samples, and a full set of federal state and school cohort dummy variables. All figures are rounded to four decimal places.

See Section 2.1 for a description of the variables used.

Source: SOEP, 2001-2015, students (upper track) aged 17 to 20, own calculations

Table A.4: Baseline Results - Political Interest

Regressors	Political Interest (a)	(b)	(c)	(d)
Reform	-0.0336	-0.0733	0.1651**	-0.0582**
	(0.0311)	(0.0475)	(0.0597)	(0.0251)
Demographic Characteristics [a]	Yes	Yes	Yes	Yes
Parental Characteristics	Yes	Yes	Yes	Yes
Household Characteristics	Yes	Yes	Yes	Yes
Number of Observations	2,536	2,536	2,536	2,536
R^2	0.0518	0.0635	0.0526	0.0605
Adjusted R^2	0.0272	0.0393	0.0281	0.0361

[a] Including *Age 18*, *Age 19*, *Age 20*, and *Has Graduated*

(a) Strong, (b) Fair, (c) Weak, (d) None

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

Note: All models include a constant, controls for sub-samples, and a full set of federal state and school cohort dummy variables. All figures are rounded to four decimal places.

See Section 2.1 for a description of the variables used.

Source: SOEP, 2000-2015, students (upper track) aged 17 to 20, own calculations

Table A.5: Baseline Results - Channel

Regressors	Volunteering, Standardised [b]	Life Satisfaction
Reform	-0.0001***	0.2618
	(0.0000)	(0.1515)
Demographic Characteristics [a]	Yes	Yes
Parental Characteristics	Yes	Yes
Household Characteristics	Yes	Yes
Number of Observations	2,240	2,544
R^2	0.0761	0.0422
Adjusted R^2	0.0498	0.0174

[a] Including *Age 18*, *Age 19*, *Age 20*, and *Has Graduated*

[b] Standardisation by available leisure time per month

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

Note: All models include a constant, controls for sub-samples, and a full set of federal state and school cohort dummy variables. All figures are rounded to four decimal places.

See Section 2.1 for a description of the variables used.

Source: SOEP, 2001-2015, students (upper track) aged 17 to 20, own calculations

Robustness Checks

Table A.6: Robustness Checks 1 of 5 (Model Specification/Time Trends/Seasonal Variation) - Pro-Social Behaviour, *Outside of School*

	Volunteering							
Regressors	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Reform	-0.0711***	-0.0628***	-0.0637***	-0.0639***	-0.0626**	-0.0680**	-0.0637***	-0.0702***
	(0.0178)	(0.0173)	(0.0172)	(0.0172)	(0.0228)	(0.0269)	(0.0171)	(0.0183)
Demographic Characteristics [a]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Household Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	2,240	2,240	2,240	2,240	2,240	2,240	2,240	2,240
(Pseudo) R^2	0.0637	0.0786	0.0793	0.0793	0.0809	0.0884	0.0785	0.0834
Adjusted R^2		0.0519	0.0526	0.0522	0.0486	0.0511	0.0509	0.0525

[a] Including *Age 18*, *Age 19*, *Age 20*, and *Has Graduated*

(1) Probit Model (Marginal Effect), (2) Adds Linear Trend, (3) Adds Quadratic Trend, (4) Adds Linear and Quadratic Trends, (5) Adds State-Specific Linear and Quadratic Trends, (6) Adds Treatment-Specific Linear and Quadratic Trends, (7) Adds Quarterly Dummy Variables, (8) Adds Monthly Dummy Variables

Robust standard errors clustered at the federal state level in parentheses
**** $p<0.01$, ** $p<0.05$, * $p<0.1$*

Note: All models include a constant, controls for sub-samples, and a full set of federal state and school cohort dummy variables. All figures are rounded to four decimal places.

See Section 2.1 for a description of the variables used.

Source: SOEP, 2001-2015, students (upper track) aged 17 to 20, own calculations

Table A.7: Robustness Checks 2 of 5 (Selection/Implementation) - Pro-Social Behaviour, *Outside of School*

	Volunteering							
Regressors	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Reform	-0.0529***	-0.0960**	-0.0477**	-0.0624***	-0.0632***	-0.0868**	-0.0660***	-0.0715**
	(0.0170)	(0.0365)	(0.0168)	(0.0170)	(0.0174)	(0.0346)	(0.0177)	(0.0297)
Demographic Characteristics [a]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Household Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	2,162	1,700	2,075	2,068	2,240	2,240	2,240	2,240
R^2	0.0746	0.0929	0.0831	0.0743	0.0777	0.0778	0.0776	0.0776
Adjusted R^2	0.0473	0.0585	0.0549	0.0457	0.0510	0.0511	0.0509	0.0509

[a] Including *Age 18*, *Age 19*, *Age 20*, and *Has Graduated*

(1) Excludes Individuals Who Move, (2) Excludes Individuals Who Live Within 10km to State Border,
(3) Excludes Individuals Who Repeat Grade, (4) Excludes Individuals Who Drop Out,
(5) Includes Dummy Variable for Double Cohorts, (6) Includes Dummy Variable for First Treatment Cohorts,
(7) Includes Dummy Variable for Special Treatment Cohorts in Saxony-Anhalt and Mecklenburg-West Pomerania,
(8) Includes Dummy Variable for Last Pre-Treatment Cohorts

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

Note: All models include a constant, controls for sub-samples, and a full set of federal state and school cohort dummy variables. All figures are rounded to four decimal places.

See Section 2.1 for a description of the variables used.

Source: SOEP, 2001-2015, students (upper track) aged 17 to 20, own calculations

Table A.8: Robustness Checks 3 of 5 (Triple Differencing) - Pro-Social Behaviour, *Outside of School*

Regressors	Volunteering
Reform	-0.1120**
	(0.0418)
Demographic Characteristics [a]	Yes
Parental Characteristics	Yes
Household Characteristics	Yes
Number of Observations	4,716
R^2	0.0739
Adjusted R^2	0.0545

[a] Including *Age 18*, *Age 19*, *Age 20*, and *Has Graduated*

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

Note: All models include a constant, controls for sub-samples, and a full set of federal state and school cohort dummy variables, both interacted with a dummy variable for academic-track students and non-interacted. All figures are rounded to four decimal places.

See Section 2.1 for a description of the variables used.

Source: SOEP, 2001-2015, students (different tracks) aged 17 to 20, own calculations

Table A.9: Robustness Checks 4 of 5 (Other Reforms) - Pro-Social Behaviour, *Outside of School*

	Volunteering				
Regressors	(1)	(2)	(3)	(4)	(5)
Reform	-0.0617***	-0.0609**	-0.0634***	-0.0641***	-0.0604***
	(0.0169)	(0.0212)	(0.0167)	(0.0157)	(0.0191)
Demographic Characteristics [a]	Yes	Yes	Yes	Yes	Yes
Parental Characteristics	Yes	Yes	Yes	Yes	Yes
Household Characteristics	Yes	Yes	Yes	Yes	Yes
Number of Observations	2,240	2,240	2,240	2,240	2,240
R^2	0.0777	0.0777	0.0777	0.0776	0.0784
Adjusted R^2	0.0510	0.0510	0.0510	0.0509	0.0518

[a] Including *Age 18*, *Age 19*, *Age 20*, and *Has Graduated*

(1) Includes Dummy Variable for Changes in Central Exit Examinations,
(2) Includes Dummy Variable for Changes in Tracking at Grade Seven,
(3) Includes Dummy Variable for Changes in Two-Tier System,
(4) Includes Dummy Variable for Changes in Subject Choice,
(5) Includes Dummy Variable for Federal Volunteer Service

Robust standard errors clustered at the federal state level in parentheses
**** $p<0.01$, ** $p<0.05$, * $p<0.1$*

Note: All models include a constant, controls for sub-samples, and a full set of federal state and school cohort dummy variables. All figures are rounded to four decimal places.

See Section 2.1 for a description of the variables used.

Source: SOEP, 2001-2015, students (upper track) aged 17 to 20, own calculations

Table A.10: Robustness Checks 5 of 5 (Placebo Tests) - Pro-Social Behaviour, *Outside of School*

Regressors	Volunteering (1)	(2)	(3)	(4)	(5)
Reform	-0.0147	-0.0301	-0.0421	0.0122	0.0275
	(0.0317)	(0.0394)	(0.0298)	(0.0225)	(0.0296)
Demographic Characteristics [a]	Yes	Yes	Yes	Yes	Yes
Parental Characteristics	Yes	Yes	Yes	Yes	Yes
Household Characteristics	Yes	Yes	Yes	Yes	Yes
Number of Observations	2,217	2,204	2,240	2,240	2,240
R^2	0.0792	0.0788	0.0783	0.0780	0.0776
Adjusted R^2	0.0527	0.0521	0.0520	0.0510	0.0513

[a] Including *Age 18*, *Age 19*, *Age 20*, and *Has Graduated*

(1) Placebo School Cohorts (c-1), (2) Placebo School Cohorts (c-2),
(3) Placebo School Cohorts (Random), (4) Placebo Federal States (Random),
(5) Placebo School Cohorts and Federal States (Random)

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

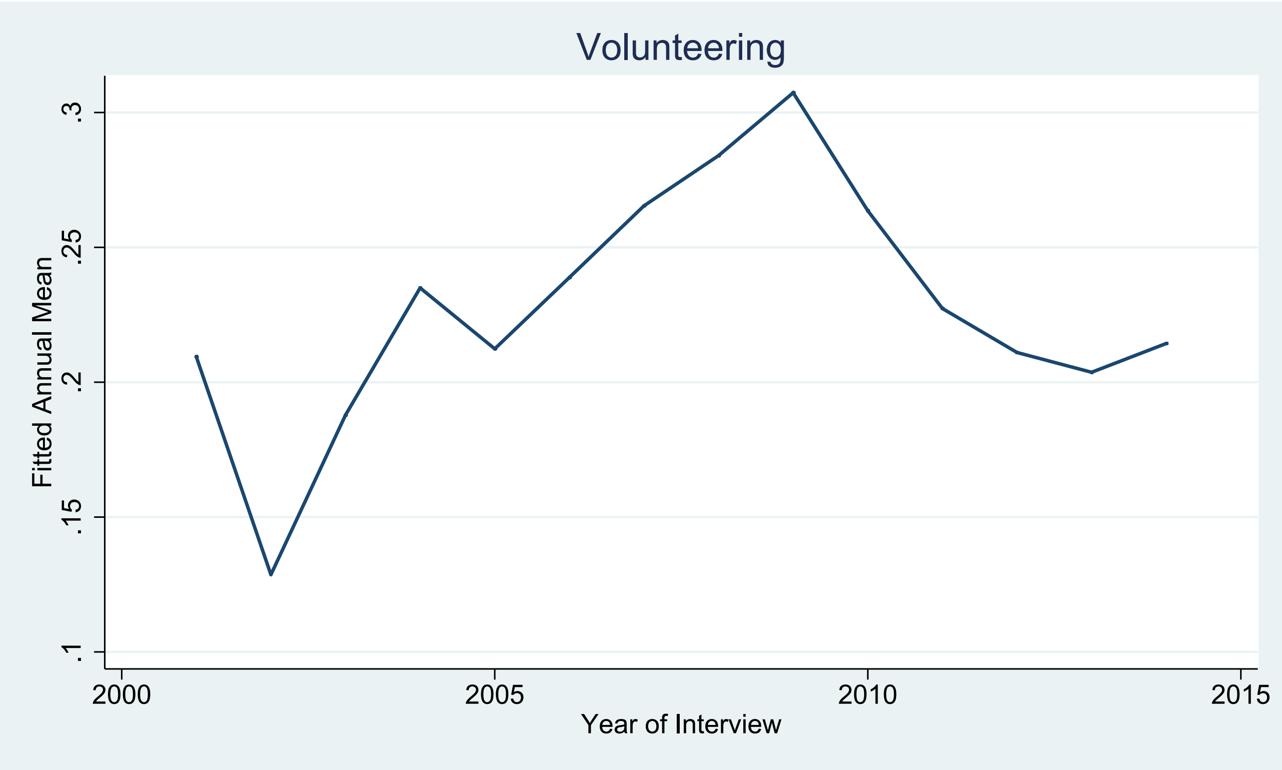
Note: All models include a constant, controls for sub-samples, and a full set of federal state and school cohort dummy variables. All figures are rounded to four decimal places.

See Section 2.1 for a description of the variables used.

Source: SOEP, 2001-2015, students (upper track) aged 17 to 20, own calculations

Web Appendix

Figure W.1: Graphical Evidence - Volunteering, Over Time

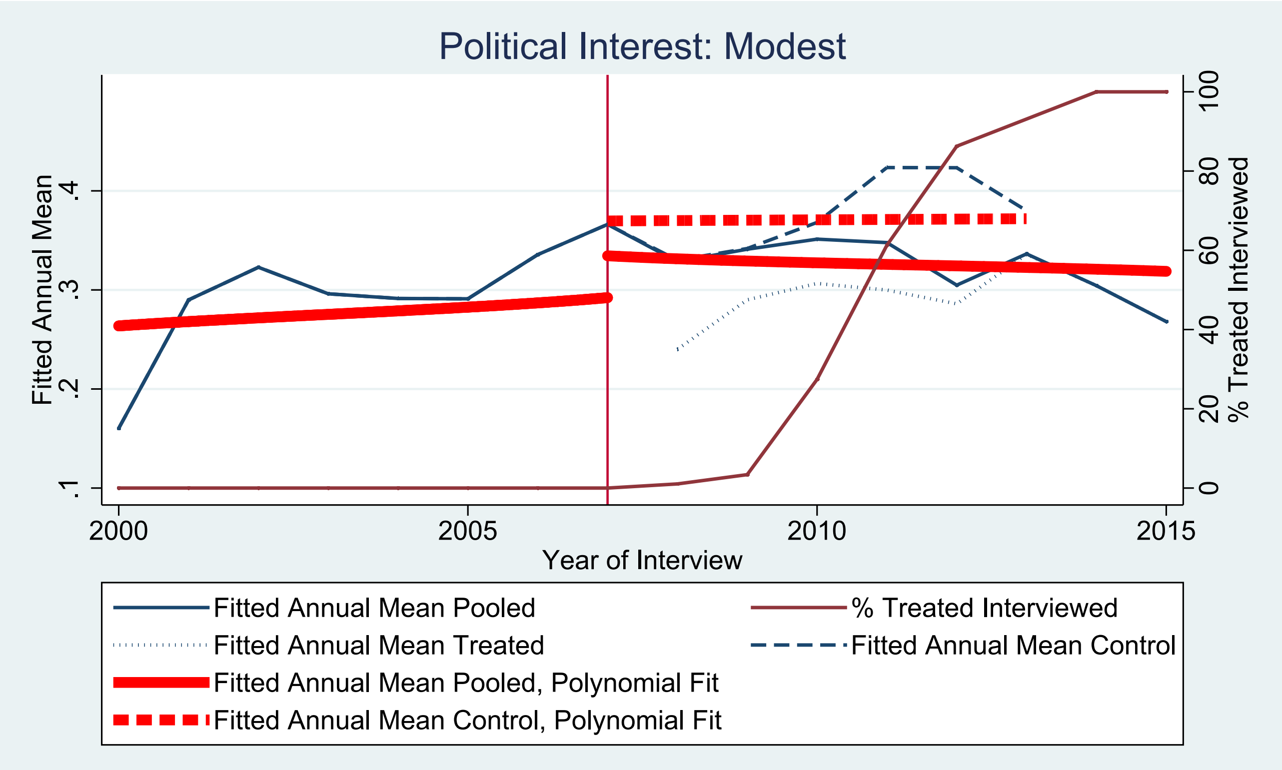


Note: The figure shows the fitted annual mean of volunteering, covariate-adjusted for demographic, educational, and parental household characteristics, as well as for sub-samples and a full set of federal state and school cohort dummy variables.

See Section 2.1 for a description of the variables used.

Source: SOEP, 2001-2015, students (lower and intermediate track) aged 17 to 20, own calculations

Figure W.2: Graphical Evidence - Political Interest, Over Time

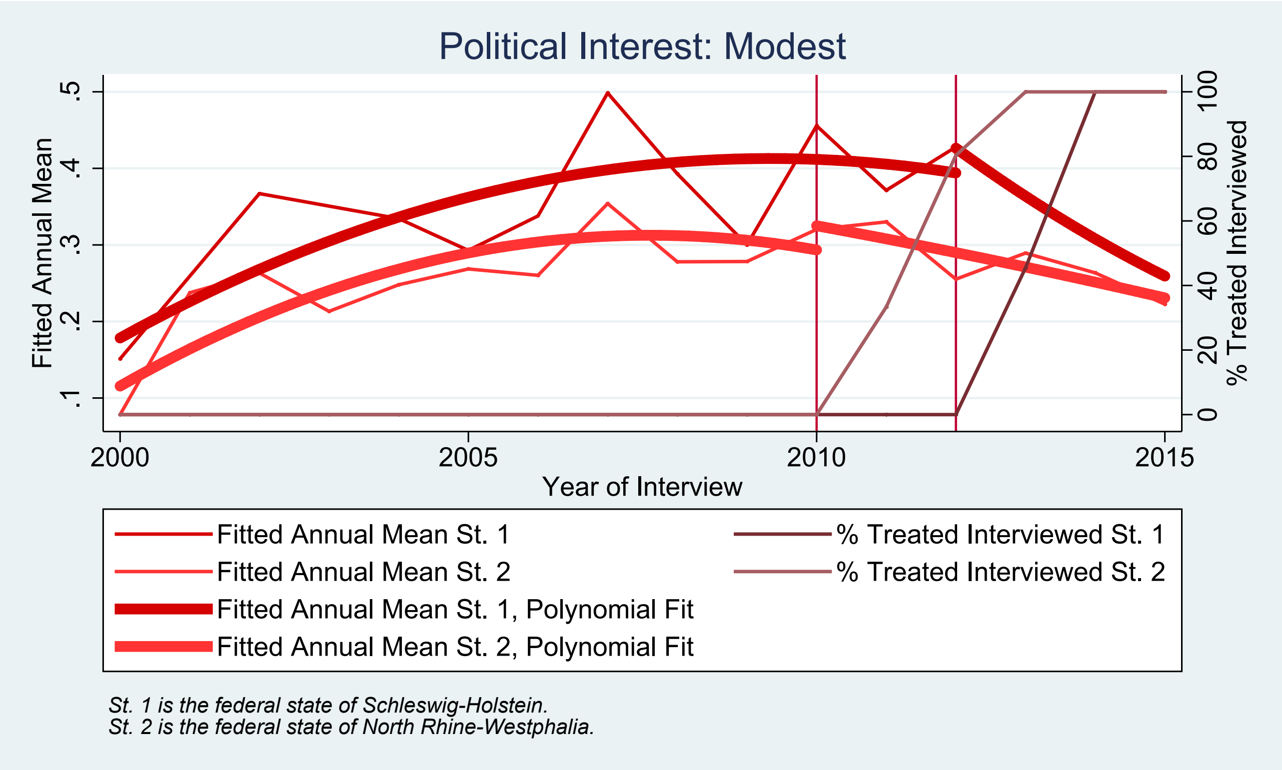


Note: The figure shows the fitted annual mean of political interest, covariate-adjusted for demographic, educational, and parental household characteristics, as well as for sub-samples and a full set of federal state and school cohort dummy variables.

See Section 2.1 for a description of the variables used.

Source: SOEP, 2001-2015, students (upper track) aged 17 to 20, own calculations

Figure W.3: Graphical Evidence - Political Interest, Common Trend



Note: The figure shows the fitted annual mean of political interest, covariate-adjusted for demographic, educational, and parental household characteristics, as well as for sub-samples and a full set of federal state and school cohort dummy variables.

See Section 2.1 for a description of the variables used.

Source: SOEP, 2001-2015, students (upper track) aged 17 to 20, own calculations

Data

Table W.1: Descriptive Statistics

Variables	Mean	Standard Deviation	Minimum	Maximum	Number of Observations
Dependent Variables					
Pro-Social Behaviour, Outside of School					
Volunteering	0.3371	0.4728	0	1	2,240
Pro-Social Behaviour, Inside of School					
Scholastic Involvement: Student Representative	0.0276	0.1638	0	1	1,958
Scholastic Involvement: Class Representative	0.4157	0.4930	0	1	1,958
Scholastic Involvement: School Magazine	0.1021	0.3029	0	1	1,958
Scholastic Involvement: Drama or Dance Group	0.1966	0.3976	0	1	1,958
Scholastic Involvement: Choir or Orchestra	0.3274	0.4694	0	1	1,958
Scholastic Involvement: Sports Group	0.2748	0.4465	0	1	1,958
Scholastic Involvement: Other Voluntary Group	0.3682	0.4824	0	1	1,958
Scholastic Involvement: None	0.1881	0.3909	0	1	1,958
Political Interest					
Political Interest: Strong	0.0623	0.2418	0	1	2,536
Political Interest: Fair	0.2606	0.4391	0	1	2,536
Political Interest: Weak	0.5229	0.4996	0	1	2,536
Political Interest: None	0.1542	0.3612	0	1	2,536
Subjective Well-Being					
Life Satisfaction	7.7649	1.4496	0	10	2,544
Independent Variables					
Reform	0.4438	0.4969	0	1	2,240

Continued on next page

Continued from previous page

Variables	Mean	Standard Deviation	Minimum	Maximum	Number of Observations
Age	17.4911	0.4969	17	20	2,240
Has Graduated	0.0384	0.1922	0	1	2,240
Is Female	0.5335	0.4990	0	1	2,240
Has Migration Background	0.2040	0.4031	0	1	2,240
Lives in East	0.1254	0.3313	0	1	2,240
Lives in Countryside	0.2835	0.4508	0	1	2,240
Parent Has Tertiary Degree	0.5214	0.4997	0	1	2,240
Parent is Blue-Collar Worker	0.1246	0.3303	0	1	2,240
Parent is Full-Time Employed	0.6438	0.4790	0	1	2,240
Parent is Single	0.0429	0.2027	0	1	2,240
Is Only Child	0.1701	0.3758	0	1	2,240
Number of Children in Household	2.4277	1.1055	1	12	2,240

See Section 2.1 for a description of the variables used.

Source: SOEP, 2001-2015, students (upper track) aged 17 to 20, own calculations

Table W.2: Distribution of Students by Age in Treatment and Control Group for Volunteering Over Time

Year	Treatment Group ($n = 994$)				Control Group ($n = 1,246$)				Total
	Age 17	Age 18	Age 19	Age 20	Age 17	Age 18	Age 19	Age 20	
2001	0	0	0	0	71	89	78	72	310
2002	0	0	0	0	0	0	0	0	0
2003	0	0	0	0	25	20	34	25	104
2004	0	0	0	0	0	0	0	0	0
2005	0	0	0	0	102	10	4	4	120
2006	0	0	0	0	95	0	0	0	95
2007	0	0	0	0	130	15	17	5	167
2008	1	0	0	0	84	8	3	0	96
2009	4	0	0	0	96	13	4	1	118
2010	47	0	0	0	81	0	0	0	128
2011	134	13	2	0	36	24	51	23	283
2012	165	0	0	0	8	0	0	0	173
2013	181	8	7	1	5	0	8	5	215
2014	200	0	0	0	0	0	0	0	200
2015	189	16	18	8	0	0	0	0	231
Total	921	37	27	9	733	179	199	135	2,240

Source: SOEP, 2001-2015, students (upper track) aged 17 to 20, own calculations

Baseline Results

Table W.3: Baseline Results - Pro-Social Behaviour, *Outside of School*

	Volunteering			
Regressors	(1)	(2)	(3)	(4)
Reform	-0.0728***	-0.0658***	-0.0657***	-0.0642***
	(0.0135)	(0.0136)	(0.0136)	(0.0168)
Age 18		0.0386	0.0382	0.0368
		(0.0328)	(0.0327)	(0.0341)
Age 19		0.0885*	0.0846*	0.0858*
		(0.0420)	(0.0425)	(0.0427)
Age 20		0.0593	0.0410	0.0413
		(0.0495)	(0.0530)	(0.0468)
Has Graduated			-0.0437	-0.0500*
			(0.0254)	(0.0267)
Is Female				-0.0221
				(0.0219)
Has Migration Background				-0.0833*
				(0.0395)
Lives in East				-0.1885***
				(0.0323)
Lives in Countryside				0.0007
				(0.0235)
Parent Has Tertiary Degree				0.0562**
				(0.0220)
Parent is Blue-Collar Worker				-0.0753**
				(0.0343)
Parent is Full-Time Employed				0.0078
				(0.0168)
Parent is Single				-0.0673***
				(0.0201)
Is Only Child				-0.0096
				(0.0336)
Number of Children in Household				0.0223
				(0.0130)
Number of Observations	2,240	2,240	2,240	2,240
R^2	0.0534	0.0555	0.0557	0.0776
Adjusted R^2	0.0327	0.0335	0.0333	0.0513

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

Note: All models include a constant, controls for sub-samples, and a full set of federal state and school cohort dummy variables. All figures are rounded to four decimal places.

See Section 2.1 for a description of the variables used.

Source: SOEP, 2001-2015, students (upper track) aged 17 to 20, own calculations

Table W.4: Heterogeneous Results - Pro-Social Behaviour, *Outside of School*

	Volunteering			
Regressors	(a)	(b)	(c)	(d)
Reform × Column	0.1105*	-0.0298	-0.0195***	0.0371
	(0.0520)	(0.0925)	(0.0597)	(0.0486)
Reform	-0.0844***	-0.0634***	-0.0817***	-0.0704*
	(0.0192)	(0.0131)	(0.0152)	(0.0366)
Column	-0.1387**	-0.0168	0.0138***	-0.1030**
	(0.0489)	(0.0765)	(0.0421)	(0.0391)
Demographic Characteristics [a]	Yes	Yes	Yes	Yes
Parental Characteristics	Yes	Yes	Yes	Yes
Household Characteristics	Yes	Yes	Yes	Yes
Number of Observations	2,240	2,240	2,240	1,709
R^2	0.0795	0.0765	0.0797	0.0853
Adjusted R^2	0.0528	0.0498	0.0589	0.0502

[a] Including *Age 18*, *Age 19*, *Age 20*, and *Has Graduated*

(a) Has Migration Background, (b) Parent Has Lower Than Secondary Degree,
(c) Parent is Blue-Collar Worker, (d) Was not Recommended Upper Track

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

Note: All models include a constant, controls for sub-samples, and a full set of federal state and school cohort dummy variables. All figures are rounded to four decimal places.

See Section 2.1 for a description of the variables used.

Source: SOEP, 2001-2015, students (upper track) aged 17 to 20, own calculations

Table W.5: Baseline Results - Political Interest

Regressors	Political Interest Modest
Reform	-0.1069**
	(0.0492)
Demographic Characteristics [a]	Yes
Parental Characteristics	Yes
Household Characteristics	Yes
Number of Observations	2,536
R^2	0.1015
Adjusted R^2	0.0783

[a] Including *Age 18*, *Age 19*, *Age 20*, and *Has Graduated*

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

Note: All models include a constant, controls for sub-samples, and a full set of federal state and school cohort dummy variables. All figures are rounded to four decimal places.

See Section 2.1 for a description of the variables used.

Source: SOEP, 2000-2015, students (upper track) aged 17 to 20, own calculations

Robustness Checks

Table W.6: Robustness Checks (1/2) - Pro-Social Behaviour, *Outside of School*

	Volunteering			
Regressors	(1)	(2)	(3)	(4)
Reform	-0.0720***	-0.0642***	-0.1614**	-0.0530**
	(0.0211)	(0.0122)	(0.0696)	(0.0241)
Demographic Characteristics [a]	Yes	Yes	Yes	Yes
Parental Characteristics	Yes	Yes	Yes	Yes
Household Characteristics	Yes	Yes	Yes	Yes
Number of Observations	2,240	2,240	2,240	2,240
R^2	0.0785	0.0776	0.0785	0.0786
Adjusted R^2	0.0522	0.0513	0.0505	0.0502

[a] Including *Age 18*, *Age 19*, *Age 20*, and *Has Graduated*

(1) Uses Weights,
(2) Uses Bootstrapped Standard Errors,
(3) Includes All Dummy Variables From 5 to 8 in Table A.7,
(4) Includes All Dummy Variables From 1 to 5 in Table A.9

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

Note: All models include a constant, controls for sub-samples, and a full set of federal state and school cohort dummy variables. All figures are rounded to four decimal places.

See Section 2.1 for a description of the variables used.

Source: SOEP, 2001-2015, students (upper track) aged 17 to 20, own calculations

Table W.7: Robustness Checks (2/2) - Pro-Social Behaviour, *Outside of School*

Regressors	Probability (a)	(b)	(c)
Reform	0.0222	0.0372	-0.0081
	(0.0138)	(0.0308)	(0.0204)
Demographic Characteristics [a]	Yes	Yes	Yes
Parental Characteristics	Yes	Yes	Yes
Household Characteristics	Yes	Yes	Yes
Number of Observations	2,240	2,240	2,240
R^2	0.0526	0.0838	0.2913
Adjusted R^2	0.0256	0.0578	0.2711

[a] Including *Age 18*, *Age 19*, *Age 20*, and *Has Graduated*

(a) Moving, (b) Repeating Grade, (c) Dropping Out

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

Note: All models include a constant, controls for sub-samples, and a full set of federal state and cohort dummy variables. All figures are rounded to four decimal places.

See Section 2.1 for a description of the variables used.

Source: SOEP, 2001-2015, students (upper track) aged 17 to 20, own calculations

CENTRE FOR ECONOMIC PERFORMANCE
Recent Discussion Papers

1494	Lorenzo Caliendo Luca David Opromolla Fernando Parro Alessandro Sforza	Goods and Factor Market Integration: A Quantitative Assessment of the EU Enlargement
1493	Andrew E. Clark Sarah Flèche Warn N. Lekfuangfu	The Long-Lasting Effects of Family and Childhood on Adult Wellbeing: Evidence from British Cohort Data
1492	Daniel Paravisini Veronica Rappoport Philipp Schnabl	Specialization in Bank Lending: Evidence from Exporting Firms
1491	M.A. Clemens J. Hunt	The Labor Market Effects of Refugee Waves: Reconciling Conflicting Results
1490	V. Bhaskar Robin Linacre Stephen Machin	The Economic Functioning of Online Drug Markets
1489	Abel Brodeur Warn N. Lekfuangfu Yanos Zylberberg	War, Migration and the Origins of the Thai Sex Industry
1488	Giuseppe Berlingieri Patrick Blanchenay Chiara Criscuolo	The Great Divergence(s)
1487	Swati Dhingra Gianmarco Ottaviano Veronica Rappoport Thomas Sampson Catherine Thomas	UK Trade and FDI: A Post-Brexit Perspective
1486	Christos Genakos Tommaso Valletti Frank Verboven	The Compact City in Empirical Research: A Quantitative Literature Review

1485	Andrew E. Clark Sarah Flèche Richard Layard Nattavudh Powdthavee George Ward	The Key Determinants of Happiness and Misery
1484	Matthew Skellern	The Hospital as a Multi-Product Firm: The Effect of Hospital Competition on Value-Added Indicators of Clinical Quality
1483	Davide Cantoni Jeremiah Dittmar Noam Yuchtman	Reallocation and Secularization: The Economic Consequences of the Protestant Reformation
1482	David Autor David Dorn Lawrence F. Katz Christina Patterson John Van Reenen	The Fall of the Labor Share and the Rise of Superstar Firms
1481	Irene Sanchez Arjona Ester Faia Gianmarco Ottaviano	International Expansion and Riskiness of Banks
1480	Sascha O. Becker Thiemo Fetzer Dennis Novy	Who voted for Brexit? A Comprehensive District-Level Analysis
1479	Philippe Aghion Nicholas Bloom Brian Lucking Raffaella Sadun John Van Reenen	Turbulence, Firm Decentralization and Growth in Bad Times
1478	Swati Dhingra Hanwei Huang Gianmarco I. P. Ottaviano João Paulo Pessoa Thomas Sampson John Van Reenen	The Costs and Benefits of Leaving the EU: Trade Effects

The Centre for Economic Performance Publications Unit
Tel: +44 (0)20 7955 7673 Email info@cep.lse.ac.uk
Website: http://cep.lse.ac.uk Twitter: @CEP_LSE