Essays on the Economics of Education and Industrial Relations

Vom Fachbereich IV der Universität Trier zur Verleihung
des akademischen Grades Dr. rer. pol. genehmigte Dissertation

von
Gabriel Schultze-Romahn

2023

Tag der Abgabe: 01.02.2023

Tag der mündlichen Prüfung: 12.07.2023

Promotionsausschuss:

Vorsitzende: Univ.-Prof. Dr. Xenia Matschke

Erstgutachter: Univ.-Prof. Dr. Laszlo Goerke

Zweitgutachter: Univ.-Prof. Dr. Uwe Jirjahn

Abstract

This thesis comprises of four research papers on the economics of education and industrial relations, which contribute to the field of empirical economic research. All of the corresponding papers focus on analysing how much time individuals spend on specific activities. The allocation of available time resources is a decision that individuals make throughout their lifetime. In this thesis, we consider individuals at different stages of their lives - students at school, university students, and dependent employees at the workplace.

Part I includes two research studies on student's behaviour in secondary and tertiary education.

Chapter 2 explores whether students who are relatively younger or older within the school year exhibit differential time allocation. Building on previous findings showing that relatively younger students perform worse in school, the study shows that relatively younger students are aware of their poor performance in school and feel more strain as a result. Nevertheless, there are no clear differences to be found in terms of time spent on homework, while relatively younger students spend more time watching television and less time on sports activities. Thus, the results suggest that the lower learning outcomes are not associated with different time allocations between school-related activities and non-school-related activities.

Chapter 3 analyses how individual ability and labour market prospects affect study behaviour. The theoretical modelling predicts that both determinants increase study effort. The empirical investigation is based on cross-sectional data from the National Educational Panel Study (NEPS) and includes thousands of students in Germany. The analyses show that more gifted students exhibit lower subjective effort levels and invest less time in self-study. In contrast, very good labour market prospects lead to more effort exerted by the student, both qualitatively and quantitatively. The potential endogeneity problem is taken into account by using regional unemployment data as an instrumental variable.

Part II includes two labour economic studies on determinants of overtime. Both studies belong to the field of industrial relations, as they focus on union membership on the one hand and the interplay of works councils and collective bargaining coverage on the other.

Chapter 4 shows that union members work less overtime than non-members do. The econometric approach takes the problem of unobserved heterogeneity into account; but provides no evidence that this issue affects the results. Different channels that could lead to this relationship are analysed by examining relevant subgroups separately. For example, this effect of union membership can also be observed in establishments with works councils and for workers who are very likely to be covered by collective bargaining agreements. The study concludes that the observed effect is due to the fact that union membership can protect workers from corresponding increased working time demands by employers.

Chapter 5 builds on previous studies showing a negative effect of works councils on overtime. In addition to co-determination by works councils at the firm level, collective bargaining coverage is an important factor in the German industrial relations system. Corresponding data was not available in the SOEP for quite some time. Therefore, the study uses recent SOEP data, which also contains information on collective bargaining coverage. A cross-sectional analysis is conducted to examine the effects of works councils in establishments with and without collective bargaining coverage. Similar to studies analysing other outcome variables, the results show that the effect of works councils exists only for employees covered by a collective bargaining agreement.

Deutsche Kurzfassung (German Abstract)

Diese Thesis umfasst vier Forschungsstudien zu bildungs- und arbeitsökonomischen Fragestellungen, die dem Gebiet der empirischen Wirtschaftsforschung zuzuordnen sind. Im Zentrum der Analyse steht bei allen Studien, wie viel Zeit Individuen für bestimmte Tätigkeiten aufbringen. Die Allokation der zur Verfügung stehenden zeitlichen Ressourcen ist eine Entscheidung, die Individuen ein Leben lang treffen. In dieser Thesis werden Individuen in verschiedenen Lebensabschnitten betrachtet - Schüler im Alter der Sekundarstufe, Studenten an Universitäten und abhängig Beschäftigte im Berufsleben.

Teil I enthält zwei bildungsökonomische Studien im Bereich der sekundären und tertiären Bildung.

In Kapitel 2 wird erforscht, ob Lernende, die innerhalb ihrer Lernkohorte relativ gesehen jünger oder älter sind, eine unterschiedliche Zeitallokation aufweisen. Aufbauend auf früheren Erkenntnissen, dass relativ jüngere Schüler schlechtere Leistungen in der Schule aufweisen, zeigt die Studie, dass relativ jüngeren Schülern ihr schlechteres Abschneiden in der Schule bewusst ist und sie eine höhere Belastung hierdurch verspüren. Dennoch lassen sich in Bezug auf die Zeit für Hausaufgaben keine eindeutigen Unterschiede finden, während relativ junge Schüler mehr Zeit mit Fernsehen und weniger Zeit für sportliche Aktivitäten verbringen. Die Ergebnisse deuten damit darauf hin, dass die niedrigeren Lernergebnisse nicht mit einer unterschiedlichen Zeitaufteilung zwischen schulischen und außerschulischen Aktivitäten einhergehen.

In Kapitel 3 wird analysiert, wie sich zum einen die individuelle Kompetenz und zum anderen Arbeitsmarktaussichten auf das Studierverhalten auswirken. Die theoretische Modellierung sagt voraus, dass beide Determinanten die Studieranstrengungen erhöhen. Die empirische Untersuchung basiert auf Querschnittsdaten aus dem Nationalen Bildungspanel (NEPS) und umfasst tausende Studierende in Deutschland. Die Analysen zeigen, dass höher begabte Studierende niedrigere subjektive Anstrengungen aufweisen und weniger Zeit in ihr Studium investieren. Hingegen führen positive Arbeitsmarktaussichten zu mehr Einsatz der Studierenden, sowohl in qualitativer als auch in quantitativer

Hinsicht. Das potentielle Endogenitätsproblem wird berücksichtigt, indem regionale Arbeitslosendaten als Instrument genutzt werden.

Teil II beinhaltet zwei arbeitsmarktökonomische Studien zu Determinanten von Überstunden. Beide Studien sind im Themengebiet der industriellen Beziehungen angesiedelt, da sie zum einen eine Gewerkschaftsmitgliedschaft und zum anderen das Zusammenspiel von Betriebsräten und Tarifbindung fokussieren.

In Kapitel 4 wird gezeigt, dass Gewerkschaftsmitglieder weniger Überstunden als Nichtmitglieder leisten. Die ökonometrischen Schätzungen berücksichtigen das Problem der unbeobachteten Heterogenität, liefern aber keine Hinweise, dass dieser Aspekt die Ergebnisse beeinflusst. Es werden verschiedene Kanäle analysiert, die zu diesem Zusammenhang führen könnten, indem entsprechende Subgruppen separat untersucht werden. Beispielsweise kann dieser Effekt der Gewerkschaftsmitgliedschaft auch in Betrieben mit Betriebsrat und für Arbeitnehmer, die sehr wahrscheinlich unter Tarifverträge fallen, beobachtet werden. Die Studie kommt zu dem Schluss, dass der beobachtete Effekt darauf zurückzuführen ist, dass eine Gewerkschaftsmitgliedschaft Arbeitnehmer vor entsprechenden erhöhten Arbeitszeitforderungen der Arbeitgeber schützen kann.

Kapitel 5 baut auf früheren Studien auf, die einen negativen Effekt von Betriebsräten auf Überstunden zeigen. Neben der Mitbestimmung durch Betriebsräte auf Unternehmensebene ist die Tarifbindung ein wichtiger Faktor im deutschen System der Arbeitsbeziehungen. Entsprechende Daten waren im SOEP lange Zeit nicht verfügbar. Daher nutzt die Studie neuere SOEP-Daten, die auch Informationen zur Tarifbindung enthalten. Es wird eine Querschnittsanalyse durchgeführt, um die Auswirkungen von Betriebsräten in Betrieben mit und ohne Tarifbindung zu untersuchen. Ähnlich wie in Studien, die andere Ergebnisvariablen analysieren, zeigen die Ergebnisse, dass der Effekt von Betriebsräten nur für Beschäftigte, die nach einem Tarifvertrag bezahlt werden, existiert.

Contents

List of Figures

List of Tables

Chapter 1

Overall Introduction

This thesis includes four papers – two papers on education economics and two on the economics of industrial relations – and is accordingly divided into two parts. All studies include an empirical investigation in order to answer the underlying research question and thus contribute to the field of empirical research in economics. Throughout this thesis, we use individual data on thousands of observations and apply several econometric techniques.

The common denominator in all of the papers is the analysis of variables which capture time spent on specific activities. For decades, time allocation is subject of economic analysis and empirical studies. The seminal paper of Becker (1965) extended the classical labour supply model and laid foundation for further research in this field. The upcoming availability of data sets on time measurements and computational resources for data analysis allowed researchers to conduct empirical studies.

Clearly, the individual division of time between work and leisure is an important question throughout one's lifespan. Children in school, students at universities and employees at the workplace have to decide on how to divide their time. Basis economic theory treats individuals as utility maximisers facing budget constraints. However, several circumstances might affect individuals and their decisions and, thus, correspond to different allocations of time. More specifically, some children might suffer from maturity disadvantages, which might positively or negatively affect time allocation. Moreover, students in the system of higher education differ in their abilities or prospects; due to these differences there might be a positive or negative effect on their study intensity. Lastly, employees exhibit differences in bargaining power, which can help them enforce a reduction or an extension of working time according to their preferences. We will consider the aforementioned factors in the corresponding chapters of this thesis.

Despite the commonality of analysing variables capturing time measurements, the interpretation of time allocation differs across the papers in this thesis. Considering students in the educational system, increased time spent on studying, may it be doing homework or self-study, is often interpreted as effort. This applies in particular to students at universities, who become increasingly independent and are not under direct control of their parents anymore. In this case, the division of time is an independent and voluntary decision. Admittedly, younger students in school are not necessarily independent utility maximisers. However, analysing students' time allocation can help to identify which factors are associated with students' behaviour. Moreover, when participating in the workforce, a different amount of working time is not necessarily the result of the well-known labour-leisure decision. It might be the case that working hours are not on a voluntary basis (and not utility-maximizing). This may be true for contractual hours, if firms only offer a specific and fixed number of hours, but also for actual working time. With respect to actual working time including overtime, it may be the case that working overtime is somewhat typical for the job or that an employer requests his workers to work overtime. Admittedly, some workers might want to do more overtime, either to increase income or to climb the career ladder. If overtime arrangements are negotiated individually, the employees' bargaining power is a fundamental facet. In contrast, the interplay of actors in the industrial relations system play an important role, if overtime arrangements are negotiated on a higher collective level.

Part I contains two studies on student's behaviour at two different stages on the educational path, i.e. secondary and tertiary education. Both studies have the common feature of using a novel data set. While the Health Behaviour in School-aged Children (HBSC) data in Chapter 2 was rarely used for economic analyses so far, the German National Educational Panel Study (NEPS) data used in Chapter 3 was recently developed for scientific use. Moreover, both studies contribute to the literature by exploring a novel research question.

Chapter 2, which is joint work with Luca Fumarco, gives insights whether students' relative age within their school year is connected with a difference in time allocation. Given the well-known phenomenon that relatively young students perform worse than their older classmates, a difference in time allocation might go along with this underperforming. The study, thereby, contributes to the literature of relative age effects (RAE), which so far has focused on multiple areas of students' life, but not on time allocation. It turns out that relatively young students are aware of their underperformance and that they are suffering from high school strain, however, they do not spend a different amount of time for doing homework. In contrast, they spend a different amount of time for specific non-school related activities – less sport but more media consumption. In sum, we do not find evidence that relative age is connected with spending a different amount of time on school-related activities and on non-school-related activities. Moreover, the study disentangles relative from absolute age effects, which is rarely achieved in the literature. In the heterogeneity analyses, we divide the sample into the three wide age groups suggested by the HBSC institute. We show that most of these effects reverse or disappear throughout the lifecycle of adolescents.

In Chapter 3, which is joint work with Adrian Chadi and Marco de Pinto, we analyse the effect of ability and labour market prospects on students' time allocation by focusing on students at German universities. While previous literature mainly focuses on study outputs such as grades, our study goes one step beyond and reveals novel and important determinants of effort decisions. By using a Cobb-Douglas Educational Production Function and standard assumptions on decision making, our theoretical modelling predicts that both ability and labour market prospects increase study effort. The empirical investigation is based on a rich data set, which is beneficial due to the following reasons. Firstly, the NEPS provides representative across-subject data, while other studies mainly make us of smaller samples from specific study subjects. Secondly, the data set provides information on study time and additionally self-assessed effort levels as two alternative outcome variables. Thirdly, due to the comprehensive competency testing in the study design, the study does not have to rely on proxies for ability, such as previous grades, but includes competency data. Despite the theoretical predictions that both ability and labour market prospects affect student effort positively, the empirical analysis shows that high-ability students tend to 'lean back', i.e. spend less time for self-studying and exert lower effort levels. In contrast, very good labour market prospects increase students' effort and thus confirm the theoretical prediction. Effect heterogeneity analysis shows, firstly, that male students tend to 'lean back'. Secondly, labour market prospects affect student effort decisions only in fields of study which do not include science, technology, engineering and mathematics. These fields of study are known for excellent job prospects such that students in these fields seem to be unaffected in their study behaviour by labour market conditions.

Apart from the results, both studies in Part I illustrate the methodological importance of considering the issue of causality. In Chapter 2, we use a two-stage least square with two instrumental variables for both potentially endogenous age variables. We instrument relative age with expected relative age, and absolute age with expected absolute age. Thus, the study provides unbiased estimates for both age variables, which is rarely done in the literature. In Chapter 3, we instrument labour market prospects with an exogenous instrumental variable based on variations in official unemployment data on regional labour markets. More concretely, we define the most common career aspiration for each field of study and then calculate the recent 'unemployment inflow' for each occupation. Using instrumental variable techniques, our findings can be interpreted as causal and rule out endogeneity concerns. This issue becomes apparent in Chapter 2, where a comparison of standard and instrumental regressions show that the former would lead to spurious results.

Part II includes two papers analysing the determinants of overtime work with a special focus on the German industrial relations' system. Both studies use data from the German Socio-Economic Panel (SOEP), which includes rich data on thousands of individuals and is collected on a yearly basis. The SOEP data is frequently used in economic analysis and also in studies focusing on the determinants of working overtime. In contrast to previous studies, our papers focus more strongly on the determinants related to facets in the system of industrial relations, which played no or only a little role in existing literature. More specifically, we focus on individual trade union membership in Chapter 4 and on the interplay of works councils and collective bargaining agreements in Chapter 5. Such collective agreements are negotiated between trade unions and employers or employer associations. For roughly 30 years, the German industrial relations system faces common challenges, i.e. a decreasing number of trade union members, a decreasing collective bargaining coverage and fewer firms with co-determination through works councils. However, collective agreements and works councils are still the main pillars in the institutional framework, while still millions of employees belong to trade unions (Oberfichtner and Schnabel, 2019). Moreover, despite the reduction in importance, these institutional actors and settings affect several labour market outcomes. In this thesis, the two studies focus on overtime work.

In Chapter 4, which is joint work with Laszlo Goerke, we show that trade union members work less overtime than non-members. We take the issue of unobserved heterogeneity into account, however, we do not find evidence that this issue affects our results. Moreover, we consider the potential correlation between union membership and works councils as well as collective bargaining. The SOEP data provides limited information about firm characteristics and utilizing this data often goes along with a reduction in sample size. Nonetheless, this union membership effect can also be observed for employees in works council establishments and for employees very likely covered by collective bargaining agreements. We examine various channels, which could give rise to the relationship, by analysing several subgroups. We hereby first focus on employees who are over-employed and then on different types of overtime compensation (paid vs unpaid or compensated with leisure). We do not find evidence that any of these channels drive our results and conclude that the observed effect appears to be due to the fact that union membership can protect employees from corresponding working time demands by employers. Similar to other gains from trade union membership, this finding constitutes a further but so far unobserved facet. Moreover, given the unclear findings from previous research, our study reveals a clear-cut effect of union membership on overtime in Germany.

Lastly, the single-authored Chapter 5 focuses on the interplay of works councils at the firm-level and collective bargaining agreements at the industry level. A negative impact of works councils on overtime is well established in the literature, especially if the contractual working time is 40 hours. However, focusing on a recent SOEP wave enables us to additionally consider collective agreements. We first replicate former results on the effect of works councils with more recent SOEP data. We then carry out a cross-sectional analysis and analyse the effects of works councils in firms with and without collective bargaining agreements. The study reveals that the negative impact of works councils is only prevalent in firms covered by such an agreement. This pattern indicates that a works council can better enforce worker-friendly arrangements, if the wages are already negotiated on a higher level. While the previous literature shows similar patterns for other outcome variables, the study firstly finds this pattern for overtime. Moreover, the study shows that previous literature uses estimation methods, which are not able to address possible uneven effects on the incidence and the amount of overtime hours. Suitable methods which consider this methodical issue are therefore applied and reveal important insights.

Part I

Education Economics

Chapter 2

Does Relative Age Make Jack a Dull Student? Evidence from Students' Schoolwork and Playtime

A large literature shows that relatively young students perform worse in class. Using data from the 'Health Behaviour in School Aged Children' international survey, we additionally find robust evidence that they are aware of performing poorly, they spend more time watching TV and less time doing sports than older peers, while tending to spend as much time as older peers on their homework. We use a two-stage least square to instrument both relative and absolute age, which turns out to be an important issue. Heterogeneity analyses show that most of these effects reverse or disappear in time.[1]

Keywords: relative age, adolescence, time allocation, strain, effort, leisure

JEL Classification: C26, D04, I24

[1]This chapter is joint work with Luca Fumarco (Fumarco and Schultze, 2020).
HBSC is an international study carried out in collaboration with WHO/EURO. For the 2001/2, 2005/06, 2009/10 and 2013/14 survey waves, the International Coordinator was Candace Currie and the Data Bank Manager was Oddrun Samdal. A complete list of the HBSC study coordinators, databank managers, researchers, and participating countries is available on `http://www.hbsc.org` (January 23, 2023). We are grateful to the HBSC advisor for this study, Alessio Vieno. Moreover, we are grateful for helpful comments by participants of a brownbag seminar in Trier and of the (virtual) Annual Conference 2020 of the German Economic Association. We thank Ingeborg Foldøy Solli for discussions on relative age, and Amber Rousse for proof reading.

2.1 Introduction

Age differences between classmates (i.e. relative age) affect educational pathways and achievement. In many countries, students are streamed into high and low streams (e.g. academic versus vocational track at the end of a given grade), with the youngest students in a class being more likely selected into low streams (Fredriksson and Öckert, 2014; Mühlenweg and Puhani, 2010; Allen and Barnsley, 1993). The reason is that the youngest students obtain on average lower grades (Ponzo and Scoppa, 2014; Sprietsma, 2010; Bedard and Dhuey, 2006). This negative effect of relative age is economically relevant because it reflects directly into a lower probability to start university and lower labor market outcomes (Peña, 2017; Black et al., 2011; Plug, 2001). While these pioneering studies mainly deal with the effect of relative age on educational achievements, recent studies have investigated indirectly related performance outcomes; for instance, mental illness (Schwandt and Wuppermann, 2016) and health (Bahrs and Schumann, 2020; Fumarco et al., 2020; Fumarco and Baert, 2019). Our paper contributes to this strand of literature on relative age effects (RAEs).

We investigate whether relative age might affect self-perceived performance, self-perceived school pressure, and decisions on time allocation between school-related and non-school-related activities. Lower success at school is associated with students' self-perception and mental strain; it is important to scrutinize whether relatively young (and worse performing) students are aware of this underperformance and whether they suffer from it. Awareness of low school performance and suffering from high school strain can be connected to anxiety, self-doubt, truancy, hopelessness, and even drug use (Shek and Li, 2016; Miller and Plant, 1999). Furthermore, a lower time allocation to school activities could lead to an inescapable vicious circle, which could reinforce and perpetuate the lower performance of relatively younger students.

To the best of our knowledge, ours is the first paper to study comprehensively the effect of relative age on these outcomes, and, while we are generally agnostic on the possible effects of relative age on these outcomes, we might derive some insights from related literature. Studies on perceived school performance provide evidence that low self-perceived performance is associated with socio-demographic and educational characteristics, along with personal circumstances.[2] So far, the impact of relative age has yet to be investigated, but it is natural to assume that relatively young students' lower actual performance goes along with their lower perceived performance. Similarly, no em-

[2]Focusing on Finland, for example, Haapasalo et al. (2010) provide evidence that being male, being in lower school grades, and being from a high socio-economic background is associated with higher perceived school performance. Similarly, they find that high educational aspirations, good student-teacher relationships, and low school strain are positively associated with higher perceived school performance.

pirical study has investigated the effect of relative age on school strain.[3] Lower marks at school are likely negatively associated with school strain; moreover, Fumarco et al. (2020) show that relatively younger adolescents are less satisfied with their life and, according to Moksnes et al. (2016), stress at school – and in particular stress concerning one's school performance – decreases the level of life satisfaction. It is therefore possible that relatively young students perform worse and have higher school strain, which goes along with lower satisfaction.

Relatively little is known about the effect of relative age on students' time allocation; however, this possible effect is relevant for different reasons. Lower grades of relatively young students could simply stem from less time dedicated to studying, but also the reverse is possible. For example, adolescent relatively young students have fewer social interactions (Fumarco and Baert, 2019); this lower time with peers could be substituted with more time spent on school-related activities. Alternatively, relatively younger students might decide to invest more time in studying to compensate maturity disadvantages.[4] While Pellizzari and Billari (2012) investigate university students at the Bocconi University, they combine their results with those from PISA data on Italian students; this investigation suggests that university students who are relatively young might perform better in tertiary education, at the cost of greater time spent doing homework in adolescence. Ambiguous effects are also possible with respect to non-school-related activities. While it might be intuitive to expect that relatively young students spend more time on virtual recreational activities (e.g. watching TV and playing videogames) – since they tend to avoid social interactions (Fumarco and Baert, 2019), the expectations on the effect of relative age on sports activities is much more ambiguous. On one hand, relatively young athletes tend to: (i) hesitate to participate in sports activities (whether purely recreational or competitive ones), owed to the tough competition from relatively older athletes in the same youth category (Smith et al., 2018; Cobley et al., 2009); and (ii) they might avoid team sports, since they shy away from face-to-face interactions (Fumarco and Baert, 2019). On the other hand, relatively young students might spend as much time as (or even more than) their older classmates in individualistic sports, where social interactions are limited and relative age plays a smaller role (e.g. exercising alone at the gym). This type of sports exercising might represent a good distraction from disadvantaged life circumstances. To sum up, we are agnostic whether relatively young students put more or fewer

[3]Fenzel (1992) analyzes the association between relative age and school strain – along with other variables. However, this study provides a purely descriptive analysis and is based on a small sample size.

[4]Focusing on students in higher education in Germany, Chadi et al. (2019) show that more talented students invest less in self-study time, while good labor market prospects increase effort. These results indicate that differences in individual capabilities affect students' time allocation and show that (de)motivational effects might play a role.

hours into school-related and recreational activities, since different effects are conceivable.[5]

To investigate relative age effects (RAEs) on students' self-perception, strain, and time allocation, we use data from the international repeated cross-section survey 'Health Behaviour in School Aged Children (HBSC)'. This survey is conducted mostly in European countries and in Northern America, and the collected data are broadly used in the social sciences;[6] recently, they have been used in two economic studies as well (Fumarco et al., 2020; Fumarco and Baert, 2019).

HBSC data include rich information on students' characteristics and on those of their households. They comprise information on both perceived performance and schoolwork strain. Moreover, HBSC data provide information on time spent in school and non-school-related activities. We measure the former activities with the amount of hours spent doing homework at the weekend and on weekdays. We measure the latter activities with five variables: (i) average weekly hours spent doing sports exercises, (ii)-(iii) average quantity of hours spent watching TV on weekdays and weekend days, (iv)-(v) average quantity of hours spent playing videogames on weekdays and weekend days.

The cross-country nature and sampling procedure of HBSC data have features that allow us to draw a refined picture of RAEs on our outcome variables for two reasons. First, while the smallest sample unit is the class, students' age range is from 10 to 17; thus, the HBSC study allows us to disentangle relative age from absolute age (e.g. we can measure relative age faced by a 12 year old student and by a 16 year old student). Second, in our sample, the calendar year does not overlap with the academic year: the calendar year is January to December in all of the countries, whereas the academic year varies by country-specific cutoff dates for starting school; thus, we can disentangle RAEs from confounders related to students' season of birth (Bound and Jaeger, 2000; Musch and Hay, 1999). These two points are discussed in greater detail in Section 2.2.

Because of the characteristic of the large variation in this dataset, we can follow the evolution of relative age effects on our outcomes. This is only the second study that does so, and the first is Fumarco et al. (2020) with the same data, but on different outcomes. Instead of using an interaction term, we divide the sample into the three wide age groups suggested by the HBSC institute. The size of these groups is large enough so that we can still use control variables for absolute age. These analyses allow us to investigate non-linear effects of absolute age.

[5]Moreover, we are agnostic with respect to which level of schoolwork pressure, or level of non-school-related activity, is either beneficial or detrimental to students' performance.

[6]See the updated list of published articles that use HBSC data: `www.hbsc.org/publications/journal` (January 23, 2023).

Similar to other studies on age at school entry and on RAEs, our study must deal with the endogeneity of students' relative and absolute age. Firstly, relative age may be determined by students' characteristics and by those of their parents (Fumarco and Baert, 2019). For example, let us consider the case of redshirted students (i.e. students whose school entry is postponed) from the US. There, students born toward the end of the academic year, and from high socio-economic status families, are overrepresented among redshirted students, because their families can afford one extra year in kindergarten (Bedard and Dhuey, 2006). When these students start school, they are older than the average age of their classmates, instead of younger, as they should be given that they were born at the end of the academic year. Secondly, absolute age is positively correlated with relative age, since students born at the beginning of an academic year are older, both in relative and in absolute terms, than their classmates are. Therefore, the same endogeneity concerns as those of relative age also apply to absolute age. We deal with this endogeneity problem with a two-stage least square (2SLS) regression method, where we instrument relative age with expected relative age, and absolute age with expected absolute age – as in Peña and Duckworth (2018). This procedure allows us to analyze separately the unbiased effects of relative and absolute age.

The remainder of the paper proceeds as follows. Section 2.2 discusses the data, the descriptive statistics, and the empirical strategy. Section 2.3 proceeds with main analyses on perceived performance, schoolwork pressure, and school as well as non-school-related activities. Section 2.4 investigates whether age effects differ for different age categories. Section 2.5 summarizes the results and its main limitations.

2.2 Data, descriptive statistics, and empirical strategy

In this section, we first discuss the most relevant HBSC survey characteristics, and then the variables used in the empirical analyses.

2.2.1 Data

The HBSC survey is organized by the World Health Organization, the series of cross-sectional surveys have been conducted every four years since 1985/6, and focus on young people's health, well-being, and health behaviors. In this study, we analyze data from 2001/2, 2005/6, 2009/10, and 2013/14 survey waves, which are the most recent publicly available ones;[7] this sample is composed of 800,000 students. However, we have to drop observations on students from Germany, Canada, the US, Russia, Armenia, Turkey, Israel,

[7]The previous four waves are not accessible to non-HBSC scholars. We choose to investigate only open-access data for the sake of transparency and replicability.

Portugal, Romania, Hungary (wave 2001) and Czech Republic (wave 2006). For a number of reasons, it is not possible to compute relative age for these students; in summary, either their birthdays are missing or the cutoff under which they receive their education cannot be correctly assigned (for greater details, see Fumarco and Baert, 2019). Additionally, we exclude students from the dataset when the class identifier was either missing or wrongly assigned.[8] This exclusion amounts to trimming the sample based on class size; we drop students from classes above the 94th percentile of the class size distribution and below the 4th percentile of this distribution.

The final sample is composed of 521,544 students from 33 countries.[9] While Table A.1 in Appendix A.1 provides the quantity of observations by country and by wave, Table A.2 lists the country-specific cutoff dates, which have been retrieved from different sources listed in Table A.3.

For the purpose of RAEs studies, the HBSC sampling is particularly attractive compared to students' surveys that are most popular among economists, namely PISA, TIMSS, and PIRLS (Cordero et al., 2018). In the PISA survey, the smallest sample unit is the student who is randomly selected across classes composed of students who turn 15 in the survey year (Sprietsma, 2010). With PISA data, absolute age and relative age cannot be disentangled. In contrast, in both TIMSS and PIRLS, the smallest sample unit is the class, which is randomly selected in participant schools; however, only classes composed of students in the fourth grade (average age 9.5) and in the eighth grade (average age 13.5) are targeted (Joncas and Foy, 2013). In TIMSS and PIRLS data, absolute age and relative age are still highly collinear. For these reasons, most studies using PISA, TIMSS, and PIRLS data focus on the related issue of the effect of age at school entry (Ponzo and Scoppa, 2014; Sprietsma, 2010; Bedard and Dhuey, 2006). In contrast, the HBSC survey allows us to focus on relative age because it can be disentangled from absolute age; in the HBSC survey, the smallest sample unit is the class, which has to include students in one of three age groups: 11, 13, and 15 years. Notice that, because of different cutoff dates and school entry rules, the sampled age range is larger: the bulk of the students falls between 10 and 17 years of age.[10]

[8]For instance, students within the same school but from different classes received one unique class identifier; so, the age range of students in this huge class is between 10 and 17. Because we focus on age differences between classmates, robustness checks that included these problematic classes would be meaningless. Note that throughout this manuscript, the term 'class' is used to identify a group of classmates; although we can identify a class of students, we cannot properly associate a grade to this class. This information is not disclosed, and cannot be precisely imputed.

[9]Moreover, the class size ranges between a minimum of 8 and a maximum of 31 students, with a mean of 18 and a standard deviation of 5.5.

[10]Roberts et al. (2009) discuss in detail the survey's methodology and the main descriptive analyses of the data.

2.2.2 Outcome variables

There are two outcomes on students' perceived performance and schoolwork strain. First, we study teachers' perception of students' relative performance, according to the student (henceforth, self-perceived performance).[11] We take the reverse scale of this categorical variable to ease the interpretation: now it ranges from 1 to 4, to indicate below average, average, good, very good performance, respectively. Second, we investigate self-perceived schoolwork pressure.[12] Past studies compare this outcome to job strain (Torsheim and Wold, 2001), so we can consider it as a proxy for schoolwork strain. It is measured on a 1-4 scale, to indicate no pressure at all, little, some, or a lot of pressure, respectively.

There are two outcomes of school-related activities. First, the average amount of time spent doing homework on weekdays and, second, the same outcome but on weekend days. These two measures cannot be summed and then averaged in a meaningful way; they range from 1 to 9: 1 denotes no time used for homework, 2 denotes half an hour for homework, 3 denotes 1 hour, and so on, up to 8 that denotes 6 hours, while 9 denotes 7 or more hours.

There are five outcomes of non-school-related activities, which include important leisure activities of young students. The first outcome of those extracurricular activities is the average amount of hours spent doing sport exercises during the whole week, which ranges from 1 to 6: 1 denotes no time used for exercising, 2 denotes half an hour, 3 denotes 1 hour, 4 denotes 2-3 hours, 5 denotes 4-5 hours, and 6 denotes 7 hours or more. The second and third recreational activities we explore are the average amount of time spent watching TV during weekdays and weekend days. The fourth and fifth non-school-related activities are the average amount of time spent playing videogames during weekdays and weekend days. The latter four outcomes are measured with the same scale and with the same categories as the two school-related activities (i.e. homework on weekdays and weekend days); therefore, for the same reasons, we cannot create average TV and average videogame time over the whole week.

The outcomes explored in this paper are the main ones from the HBSC survey that concern time use in activities that can be considered either school- or non-school-related. There exist a few additional potential outcomes that we do not study, because they are repetitive of the outcomes in either this paper or in Fumarco and Baert (2019).[13]

[11] The survey question is 'In your opinion, what does your class teacher(s) think about your school performance compared to your classmates?'

[12] The survey question is 'How pressured do you feel by the schoolwork you have to do?'

[13] There is a measure of computer utilization; however, it does not distinguish between school-related use or leisure. There are a few alternative measures of sports exercising that are equivalent to the outcome we explore. There is a measure of frequency of meeting with friends in the afterschool during the whole week, which is explored in Fumarco and Baert (2019), on the study on RAEs on social networks.

Table 2.1: Descriptive statistics

Variable	N	Mean	SD	Min	Max	Waves
Schoolwork pressure	512,759	2.256	0.895	1	4	I-IV
Perceived performance	513,170	2.816	0.831	1	4	I-IV
Homework, week-days	99,160	3.457	1.459	1	9	I
Homework, weekend-days	98,665	3.132	1.715	1	9	I
Sports exercises, whole week	396,777	3.560	1.456	1	6	II-IV
TV, week-days	495,323	4.205	2.842	1	9	I-IV
TV, weekend-days	488,094	5.083	2.025	1	9	I-IV
Videogame, week-days	385,620	3.109	2.046	1	9	II-IV
Videogame, weekend-days	378,746	3.852	2.389	1	9	II-IV
Relative age	505,900	−0.300	0.452	−5.750	5.167	I-IV
Absolute age	521,542	13.458	1.647	9.833	17	I-IV
Female	521,544	0.507		0	1	I-IV
Both parents at home	502,974	0.769		0	1	I-IV
Low SES	521,544	0.202		0	1	I-IV
Medium SES	521,544	0.400		0	1	I-IV
High SES	521,544	0.398		0	1	I-IV
Season-of-birth	521,544	5.507	3.362	0	11	I-IV

Source: HBSC data. Note: 'SD' stands for standard deviation. Wave I stands for 2001/2, wave II stands for 2005/6, wave III stands for 2009/10, and wave IV stands for 2013/14.

Table 2.1 shows that, on average, adolescents declare to face little schoolwork pressure and think that, according to their teachers, they have good performance in relation to their classmates. Moreover, they do homework for about one hour per day, both during the week and the weekend. Adolescents exercise for more than one hour per day, while they watch TV for about two hours during weekdays and three hours during the weekend. Finally, adolescents declare to play videogames for one hour during the week and for about two hours during the weekend.

It is important to remark that not all of the outcomes explored in this paper are available for each wave. The two survey items on time spent doing homework are present only in wave 2001/2, while the three survey items on time spent doing sports exercises and playing videogames are absent only in that wave.

2.2.3 Independent variables of interest

2.2.3.1 Relative age

As in Fumarco et al. (2020) and Fumarco and Baert (2019), the explanatory variable of interest to proxy relative age (RA) is the difference – measured in fractions of a year – between the age of student i in class c, AGE_{ic} and the maximum absolute age of student I in the same class, AGE_{Ic}; where student I belongs to the subset of 'regular students' in class c, R_c.[14] A regular student is a student who has not suffered retention and entered school according to her age and to her country cutoff date. In this paper, relative age is converted into years such that we can compare the impact of relative and absolute age. The construction of relative age is illustrated by Equation (2.1):

$$RA_{ic} = AGE_{ic} - \max_{I=1,\dots,n} (AGE_{Ic} | I \in R_c) \tag{2.1}$$

Equation (2.1) implies that when $i \in R_c$ (i.e. student i is a regular student), we have that $RA_{ic} \in [-1, 0]$. $RA_{ic} = 0$ implies that student i is the oldest regular student in class (i.e. $i = I$) and $RA_{ic} = -1$ implies that student i is the youngest regular student.[15]

When $i \notin R_c$ (i.e. student i is not a regular student), we have that either $RA_{ic} > 0$ (i.e. student i is older than expected; she has been retained, for instance) or that $RA_{ic} < -1$ (i.e. student i is younger than expected; she entered school earlier than expected, for example).[16] This operationalization is useful for three main reasons: (i) it clearly stresses the focus on the age-grouping system; (ii) it helps to reduce the correlation between age and relative age; and (iii) it increases variation in relative age.

Table 2.2 reports the pairwise correlations between relative age and the outcome variables. A missing correlation means that the two items referring to those outcomes are not present in the same survey wave.

[14]This measure of relative age is similar to that used in Ponzo and Scoppa (2014) and Elder and Lubotsky (2009), where the authors operationalize relative age as the difference between student's i absolute age and the average absolute age of her classmates (i.e. the similarity is in the fact that their reference age varies by class).

[15]Actually, it is almost -1, because -1 means that this student was born on the exact cutoff date of the following academic year.

[16]For details on how we identified (non-)regular students, refer to Fumarco and Baert (2019).

Table 2.2: Matrix of pairwise correlations between relative age and the outcomes

Variable	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1 Relative age	1.000									
2 Schoolwork pressure	-0.005^{***}	1.000								
3 Perceived performance	-0.049^{***}	-0.158^{***}	1.000							
4 Homework, week-days	0.021^{***}	0.122^{***}	0.103^{***}	1.000						
5 Homework, weekend-days	0.034^{***}	0.108^{***}	0.107^{***}	0.652^{***}	1.000					
6 Sports exercises, whole week	-0.017^{***}	-0.032^{***}	0.063^{***}	.	.	1.000				
7 TV, week-days	0.019^{***}	0.037^{***}	-0.087^{***}	0.024^{***}	-0.004	-0.045^{***}	1.000			
8 TV, weekend-days	0.027^{***}	0.051^{***}	-0.071^{***}	0.087^{***}	0.097^{***}	-0.042^{***}	0.672^{***}	1.000		
9 Videogame, week-days	0.017^{***}	0.007^{***}	-0.093^{***}	.	.	0.004^{***}	0.363^{***}	0.280^{***}	1.000	
10 Videogame, weekend-days	0.026^{***}	0.009^{***}	-0.083^{***}	.	.	0.007^{***}	0.279^{***}	0.350^{***}	0.801^{***}	1.000

Source: HBSC data. Note: Missing correlations indicate that the items referring to those outcomes are not present in the same survey wave; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

The correlations matrix in Table 2.2 suggests four interesting descriptive results. First and foremost, relatively young students appear to face higher schoolwork pressure and believe to fare quite well at school; the second result is contrary to what one might expect naturally. Second, the correlation between perceived performance and schoolwork pressure is negative: students who perform better face lower schoolwork pressure. Third, relative age is positively correlated with hours spent doing homework – in contrast to findings in Pellizzari and Billari (2012) for Italy, negatively correlated with hours of exercises per week – at odds with the dominant strand of literature on RAEs in sports (Smith et al., 2018; Cobley et al., 2009), and positively correlated with hours spent watching TV and playing videogames. Fourth, there exist interesting correlations between outcomes. Intuitively, students who play more videogames and watch more TV feel greater schoolwork pressure and think they are perceived as not being very good students. Counterintuitively, students who do more homework watch more TV, probably because they spend more time indoors. Students who exercises more, watch less TV, but spend slightly more time playing videogames; notice though, students in this data set spend less time in general playing videogames than watching TV. Furthermore, students who play videogames more often also watch more TV.

These results should be considered with a grain of salt, however; they do not control for possible confounders (most importantly, absolute age) and they do not account for endogeneity of relative age either.

2.2.3.2 Absolute age

The second variable of interest is absolute age, which is also measured in years. This variable is included to disentangle its effect from that of relative age. This disentanglement is possible because of both the way we construct relative age and the presence of substantial between as well as within country variation in absolute age. Different from more popular surveys, such as PISA, TIMSS, and PIRLS, the HBSC survey targets a large range of adolescents, between 10 and 17 years of age (Roberts et al., 2009). This relatively wide age range corresponds to a formative time of adolescent development. Hence, it is very likely that absolute age influences attitudes and activities of the outcome variables. Disentangling relative and absolute age is therefore crucial for our analysis.[17]

It is important to remark that absolute age is likely endogenous, for the same reasons as relative age. We therefore employ an instrumental variable approach, which we describe in Section 2.2.5 in detail.

[17] Squared age could be included as an additional independent variable in order to consider a possible non-linear effect of absolute age on the outcome variable. To ease econometric analyses where absolute age has to be instrumented, we do not include such a quadratic term.

In Section 2.4, we analyze age-based subsamples, to investigate how RAEs vary over time. Based on the HBSC survey guidelines, each participant country has to select students from different grades, so that each country surveys students in three broadly defined age categories: age category 11 that includes students who should be between 11 and 12 years of age, age category 13 that contains students who should be between 13 and 14 years old, and age category 15 that comprises students who should be between 15 and 16 years of age (Roberts et al., 2009). Although the mean ages in these categories are about 11.5, 13.5, and 15.5, respectively, each country has different rules on redshirting, grade skipping, early entry, and retention; therefore, the age range in each category is somewhat larger (Roberts et al., 2009). In the heterogeneity analyses in Section 2.4, we rename these three age categories as "Below 12", "13-14", and "Above 15."

2.2.4 Control variables

While the econometric analyses include basic demographic characteristics, such as dummies for being female, and for having both parents at home, in this section we focus on some details of family socio-economic status and season-of-birth.

Family socio-economic status (SES) is originally a categorical variable derived from multiple items and is constructed according to the HBSC guidelines (Currie et al., 2008).[18] This variable ranges from 0 (for Low SES, which is the reference group) to 2 (for High SES).

Our analyses account for the so-called 'season-of-birth effects.' These are unobservable effects of birthdate that are unrelated to age differences between classmates and that could bias the estimates of RAEs. This variable for season of birth is proxied by the month of birth within the calendar year and ranges from 0 (i.e. January, the reference month) to 11 (i.e. December). In controlling for season of birth, we follow the seminal Bedard and Dhuey (2006); the underlying assumption is that season-of-birth effects do not vary across countries – and thus we do not insert interaction terms between country and season of birth.[19] This variable is inserted into our econometric analyses as a set of dummy variables, to capture month-specific effects.

It is important to remark that there is a low correlation between relative age and season of birth; see Table A.4, in Appendix A.1. This low correlation is owed to two features of relative age: (i) the reference student I varies by class; and (ii) the country specific cutoff

[18]The HBSC survey includes four questions: whether the respondent's family owns zero, one, or more than one car; whether the respondent sleeps in her own bedroom; whether the respondent has travelled for holidays in the last twelve months never, once, or more often; and whether the respondent owns zero, one, or more than one computer. For each student the numeric answers to these questions are summed and divided into three levels of family SES: low, medium, and high.

[19]While this might look as a strong assumption, no study has documented possible differences in season-of-birth effects between European countries, to the best of our knowledge.

date determines who the reference student is. In contrast, the calendar year is the same for all of the countries.

Probably, the most well-known season-of-birth effect is discussed in Buckles and Hungerman (2013). In the US, students born at the end of the calendar year are among the oldest in their classes (on average, across states with different cutoffs), but they are more likely to have disadvantageous family backgrounds. Therefore, for these students, season-of-birth effects counterbalance their advantageous higher maturity.[20]

Finally, all of the regression analyses include both wave and country-fixed effects (the dataset includes four waves and 33 countries).

2.2.5 Instrumental variables

In this subsection, we discuss the instrument for relative age and then we conduct balance tests on the orthogonality of the instrument with respect to demographic variables.

2.2.5.1 Expected relative and absolute age

Since we expect relative age to be endogenous, the benchmark analyses are conducted with two-stage least square (2SLS) regressions. In these analyses, month of birth within the academic year is used as an instrument for relative age. This variable proxies the age difference in months between student i and the hypothetical oldest regular student in class (i.e. a regular student born in the first month of the academic year), if student i was a regular student. Therefore, we call this variable expected relative age (ERA), which ranges between 0 – for students born in the month that starts with the cutoff date – and 11 – for students born in the month immediately before the cutoff date. This version of the instrument 'expected relative age' is used as in Datar (2006) and in Peña and Duckworth (2018) as well.

Similar to relative age, expected relative age does not overlap with season of birth because of the variation in cutoff dates among the countries in the sample. Consult Table A.4, in Appendix A.1, for the correlation between expected relative age and season of birth. Additional details on the identification strategy that allows us to disentangle season of birth from expected relative age are provided in Fumarco et al. (2020) and in its online appendix.

To create an instrument for absolute age, we follow closely Peña and Duckworth (2018). This instrument is the expected absolute age of student i's classmates who participated in the same survey and are from the same country; this instrument is then rounded

[20] Season-of-birth effects are discussed in several other papers, such as Fan et al. (2017), Rietveld and Webbink (2016), Currie and Schwandt (2013), Lokshin and Radyakin (2012), Bound and Jaeger (2000), and Musch and Hay (1999).

to the closest integer age. This way of constructing the expected absolute age insures that its value is the same for all students in the same class. The latter aspect is relevant because, in turn, it insures that the instruments for relative and absolute age are orthogonal with respect to each other. The correlation between these two instruments is nil according to standard interpretation ($r = 0.013$).

2.2.5.2 Balance tests

In order to test whether the instruments are balanced through observable demographic characteristics, we conduct nonparametric tests. We run OLSs of demographic characteristics on ERA and fixed effects for country, wave, and season of birth. These tests allow us to gain a measure of the conditional correlation between observable demographic characteristics and the instruments for relative age. We do not check the conditional correlation between expected relative age and absolute age since the latter is expected to be endogenous as well. The results are reported in Appendix A.1, Table A.5, and seem to be quite reassuring. The month of birth within the academic year is not driven by family composition. Moreover, although there is a statistically significant correlation between month of birth within the academic year and SES, we should note two aspects: (i) this correlation is far from being economically relevant, and (ii) it does not appear to have a meaningful pattern, when we correlate expected relative age with dummies for SES.

It is important to remark that these results on the unbiased nature of expected relative age provide evidence of the exogeneity of birth date; thus, they extend to expected absolute age as well.

2.2.5.3 Empirical strategy

In the result Section 2.3, we report benchmark results from the 2SLS, and results from the OLS. Results from the latter regression method are reported to assess how biased the estimates would have been if we did not consider endogeneity of relative and absolute age.

Independent from the outcome investigated, the 2SLS analyses include the same variables. We conduct two first stages, one per endogenous variable (i.e. relative age and absolute age); the endogenous variable is regressed on the same battery of variables: expected relative age, ERA, expected absolute age, EAA, and on a set of demographic characteristics, X. This set includes a dummy for the student being female, a dummy for having both parents at home, two dummies for Medium and High SES. Additionally, there are three sets of fixed effects, FE, namely dummies for calendar month of birth –

as proxies for season-of-birth, dummies for wave as well as for country of the survey. See Equation (2.2):

$$Endogenous_{ic} = \gamma_0 + \gamma_1 ERA_i + \gamma_2 EAA_i + \gamma_3 X_i + \gamma_4 FE_i + \mu_i \tag{2.2}$$

In the second stage, we regress the outcome variable, Y, on the predicted values of RA and AA from the two first stages, on demographic characteristics, X, and on the three sets of fixed effects, FE. See Equation (2.3):

$$Y_i = \beta_0 + \beta_1 \widehat{RA}_{ic} + \beta_2 \widehat{AA}_{ic} + \beta_3 X_i + \beta_4 FE_i + \epsilon_i \tag{2.3}$$

The OLS specification is similar to Equation (2.3); the only difference is that there is (observed) RA and AA, instead of predicted values.

In all of the subsequent analyses, we compute robust standard errors clustered on class.

All of the outcome variables are standardized in order to be able to compare the results. Therefore, the estimated coefficients are interpreted in terms of standard deviations.

2.3 Results

This section is composed of three subsections. Firstly, we investigate the two main outcomes. Secondly, we analyze the two measures of school-related activities. Thirdly, we investigate the five measures of non-school-related activities.

In order to minimize the space occupied by the numerous analyses, the tables report only the estimates for relative and absolute age, sample size, and adjusted R-squared. Results from reduced form and the first stage are reported in separate tables in Appendix A.2, together with results from weak- and under-identification tests. Moreover, tables in this appendix report results from the over-identification tests if we had disaggregated ERA – as in Fumarco et al. (2020) and Fumarco and Baert (2019) – in dummies for academic month of birth.

2.3.1 Perceived performance and schoolwork pressure

Table 2.3 reports the results on the main outcomes: perceived performance and perceived schoolwork pressure.

Results from the top panel are in line with reasonable expectations. First, relatively old students feel that their teachers think that they are over-performing compared to their younger classmates. A one-year increase in relative age (i.e. the theoretical maximum

Table 2.3: Relative age on standardized perceived performance and schoolwork pressure

	Perceived performance (1)	Schoolwork pressure (2)
2SLS		
Relative age	0.221^{***}	-0.128^{***}
	(0.015)	(0.015)
Absolute age	-0.097^{***}	0.127^{***}
	(0.001)	(0.001)
N	480,674	480,295
Adj. R-squared	0.097	0.089
OLS		
Relative age	-0.019^{***}	-0.056^{***}
	(0.004)	(0.004)
Absolute age	-0.098^{***}	0.126^{***}
	(0.001)	(0.001)
N	480,674	480,295
Adj. R-squared	0.107	0.090

Source: HBSC data. Note: The analyses include two sets of variables: (i) demographic control variables (dummy for being female, dummy for having both parents at home, and dummies for medium and high socio-economic status), and (ii) fixed effects (wave, country, season-of-birth). Standard errors clustered on class are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

age gap between regular students) increases the perceived performance by 0.221 standard deviations; it implies that a one-month increase in relative age yields an increase in perceived performance by 0.018 standard deviations. This finding is in line with the previous literature according to which relatively young students achieve lower test scores. Studies on the reliability of the information on perceived school performance show that this item is valid and reflects differences in objective grades (Felder-Puig et al., 2012). Consequently, we can ascertain that relatively younger students are aware of their lower performance. Second, relatively old students feel less schoolwork pressure than their younger classmates do. A one-year increase in relative age decreases perceived schoolwork pressure by 0.128 standard deviations, that is, 0.011 standard deviations for an age difference of one month.

With respect to absolute age, we find two results. First, an increase by one year in absolute age lowers the perceived performance by 0.097 standard deviations. A decrease of perceived school performance in higher school grades is a well-known phenomenon (Shek and Li, 2016). Second, increases in absolute age increase pressure from schoolwork: an increase by one year in absolute age increases the schoolwork pressure by 0.127 standard deviations.

It is interesting to remark two aspects of these results. First, the sign of these results: relative and absolute age work in opposite directions. Had we not controlled for absolute age, the estimated effect of relative age on these two outcomes would have been biased toward zero. Second, the relative magnitude of these effects: (i) the effect on schoolwork pressure of one additional year in absolute age is equivalent to the effect of a one-year increase in relative age; (ii) the effect on perceived performance of one additional year in absolute age is about 50% of the effect of a one-year increase in relative age.

With the OLS, which does not account for the endogeneity of either relative or absolute age, the estimated RAEs on perceived performance and on schoolwork pressure would be biased upwards and downwards, respectively. Moreover, these estimates would be much smaller in absolute terms.

Table A.6, in Appendix A.2, reports results from both the reduced form and the first stage, together with results from weak-, under-, and over-identification tests. Most importantly, the tests reject the null hypotheses that the instruments are not correlated with the endogenous variable and that they are only weakly correlated (see critical values in Stock and Yogo, 2005), respectively. When disaggregating the instrument into 12 dummy variables, the over-identification test does not reject the null hypothesis that the instruments are not correlated with the second-stage error term. Thus, these results from the over-identification test further support the findings from the balance test in Subsubsection

2.2.5.2, on the exogeneity of expected relative age, and, more in general, on the exogeneity of birth date.

We repeated the analyses on these two outcomes while substituting absolute age with grade – instrumented with expected grade. In these additional analyses, we expected very similar results given the high correlation between absolute age and grade – they are not perfectly collinear because absolute age has non-integer numbers while grade has only integer numbers. In fact, we find that the results are virtually identical; RAEs are unchanged, and sign and magnitude of grade effects are the same as those of absolute age. These same robustness checks are repeated also for school-related and non-school-related activities, and lead to the same results. These complete analyses and results can be provided upon request.[21]

2.3.2 School-related activities

Table 2.4 reports the results on school-related activities: homework during weekdays and weekend days.

Results from the top panel do not show a clear effect of relative age on the time allocated to homework, since the estimated coefficient is not statistically significant for homework during the week and only weakly significant for homework on weekends. Regarding the latter outcome, an increase in relative age of one year corresponds to a decrease in the time allocated to homework by 0.100 standard deviations. Note that this weak evidence is in line with some findings in the literature, according to which relatively younger students spend more time on homework (Pellizzari and Billari, 2012). However, it is important to remark that these analyses are conducted only on students surveyed in the 2001 wave, which is the only wave in our sample that includes the two items on homework. Additionally, for that wave, fewer countries participate in the HBSC survey; see the details in Table A.1, Appendix A.1. Therefore, the sample size is about 25% the average sample size investigated in other analyses.

For absolute age, we find evidence that students with higher absolute age spend more time on homework both on weekdays and during the weekend. Compared to homework on weekdays, the effect of absolute age on homework at the weekend is more than double in magnitude, reflecting the increasing amount of homework in advancing school grades.

With the OLS, which does not account for the endogeneity of relative age, we would find highly significant but spurious negative RAEs on time spent doing homework, both during the week or on the weekends. This would indicate that relatively young students do

[21]Note that, as discussed in footnote 8, HBSC data do not include information on the school grades and their imputation cannot be properly conducted. Therefore, we are not completely fond of the results from this robustness check.

Table 2.4: Relative age on standardized time spent doing homework, during both week-days and weekend-days

	Homework, week-days (1)	Homework, weekend-days (2)
2SLS		
Relative age	-0.057	-0.100^{*}
	(0.055)	(0.057)
Absolute age	0.026^{***}	0.064^{***}
	(0.003)	(0.003)
N	90,393	89,934
Adj. R-squared	0.177	0.205
OLS		
Relative age	-0.035^{***}	-0.050^{***}
	(0.010)	(0.010)
Absolute age	0.026^{***}	0.062^{***}
	(0.003)	(0.003)
N	90,393	89,934
Adj. R-squared	0.177	0.206

Source: HBSC data. Note: The analyses include two sets of variables: (i) demographic control variables (dummy for being female, dummy for having both parents at home, and dummies for medium and high socio-economic status), and (ii) fixed effects (wave, country, season-of-birth). Standard errors clustered on class are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

more homework. However, the findings for non-instrumented absolute age do not appear to differ from those obtained while instrumenting it.

Table A.7 in Appendix A.2 reports results from both the reduced form and the first stage, together with results from weak-, under-, and over-identification tests. Most importantly, the tests for under- and for weak-identification reject the null hypotheses that the instruments are not correlated with the endogenous variable and that they are only weakly correlated, respectively. With regard to homework at the weekend, the over-identification test rejects the null hypothesis that the over-identifying restrictions are valid, and thus it casts doubts on the non-correlation between instruments and the second-stage error term.

2.3.3 Non-school-related activities

Table 2.5 reports the results from the 2SLS on the non-school-related activities: time spent doing sports exercises (during the whole week), watching TV, and playing videogames, during both weekdays and weekend days.

Table 2.5: Relative age on standardized time spent doing sports exercises (during the whole week), watching TV and playing videogames, during both week-days and weekend-days

Variables	Sport exercises, whole week (1)	TV, week-days (2)	TV, weekend-days (3)	Videogame, week-days (4)	Videogame, weekend-days (5)
2SLS					
Relative age	0.117^{***}	-0.026^{*}	-0.051^{***}	-0.002	-0.025^{*}
	(0.015)	(0.016)	(0.016)	(0.015)	(0.015)
Absolute age	0.004^{***}	0.028^{***}	0.044^{***}	-0.016^{***}	-0.010^{***}
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
N	377,277	465,926	461,164	366,871	362,450
Adj. R-squared	0.099	0.053	0.038	0.128	0.151
OLS					
Relative age	-0.038^{***}	0.039^{***}	0.016^{***}	0.049^{***}	0.029^{***}
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Absolute age	0.002^{*}	0.029^{***}	0.044^{***}	-0.015^{***}	-0.010^{***}
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
N	377,277	465,926	461,164	366,871	362,450
Adj. R-squared	0.095	0.053	0.039	0.129	0.152

Source: HBSC data. Note: The analyses include two sets of variables: (i) demographic control variables (dummy for being female, dummy for having both parents at home, and dummies for medium and high socio-economic status), and (ii) fixed effects (wave, country, season-of-birth). Standard errors clustered on class are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Column (1) provides evidence that relatively old students spend more time doing sports exercises during the whole week. A one-year increase in relative age increases the time spent exercising by 0.117 standard deviations, equivalent to 0.010 standard deviations for a one-month decrease in relative age. This result is in line with the findings of the literature on RAEs in sports, according to which relatively young athletes have a tough time competing with older athletes from the same youth category, and thus tend to shy away from sports activities (Smith et al., 2018; Cobley et al., 2009). Moreover, relatively young students tend to shy away from general face-to-face social interactions (Fumarco and Baert, 2019). While we cannot study whether relative age effects differ based on the type of sports (i.e. team sports where relative age matters versus individual sports where relative age is less or not important), the net effect is positive on relatively old students (and thus negative on relatively young students).

Furthermore, we find a positive but small effect of absolute age, indicating that the amount of time spent doing sports activities is slightly increasing throughout the lifecycle of adolescents.

What about relative age effects on watching TV and playing videogames? We find the following results. Column (3) provides statistically significant evidence that relatively old students watch less TV on the weekends: a one-year increase in relative age decreases the time watching TV by 0.051 standard deviations, equivalent to 0.004 standard deviations for a one-month increase in relative age. On a significance level of 10%, Column (2) confirms this finding for week-days, although the magnitude of the effect is half as strong.

We should remark that increases in absolute age increase the time spent watching TV during the entire week; however, this increase would be completely offset by a one-year increase in relative age.

Column (4) shows that relatively young students play as much videogames as their older classmates on week-days, while Column (5) provides some evidence that relative age decreases the time spent playing videogames at the weekend.

Furthermore, the estimates show that time spent on videogames decreases with absolute age. Compared to the time spent watching TV, these negative effects of absolute age on playtime are much smaller in magnitude and do not offset the positive effects on watching TV.

With the OLS, which does do not account for the endogeneity of either relative or absolute age, we would find opposite RAEs compared to those from the 2SLS for all of the outcomes.

Tables A.8 to A.10 in Appendix A.2, report results from both the reduced form and the first stage, together with results from weak-, under-, and over-identification tests. Most importantly, the tests for under- and for weak-identification reject the null hypotheses that the instruments are not correlated with the endogenous variable and that they are only weakly correlated, respectively. Except for the analysis on time spent watching TV on the weekends, the over-identification test does not reject the null hypothesis that the over-identifying restrictions are valid, and thus not correlated with the second-stage error term.

Overall, results from over-identification tests reassure us on the exogeneity of expected relative age.

2.4 RAEs by absolute age groups

In this section, we present the analyses on three broadly defined age categories: "Below 12", "13-14", and "Above 15." These analyses allow us to scrutinize whether relative and absolute age effects are persistent over the school career.

Table 2.6 provides the RAEs by age category for all nine outcome variables. First, relatively old students feel that they are perceived as being better performers, except in the oldest age category (see Column (1), Outcome (1)). This positive effect is strongest for the youngest for students below 12, becomes smaller in magnitude for the middle age group and becomes slightly negative for students in the category "Above 15." Furthermore, we can observe increasing returns to absolute age (see Column (2), Outcome (1)). While absolute age is not associated with perceived performance for "Below 12" students, the relationship becomes increasingly positive for "13-14" and "Above 15" students.

The results for Outcome (2) show that the negative effect of relative age on schoolwork pressure becomes less prevalent over time, while the effect of absolute age has an inverted-U shape.

In contrast to the lack of evidence of RAEs on homework during the week on the full sample (see Section 2.3.2), Table 2.6 indicates negative relative age effects on this homework variable for the youngest age category, but confirms no effect on the two older age categories. The results are similar when considering the time spent for homework in the weekend. There is no statistical evidence of an effect of absolute age on both homework variables, when splitting the full sample into age groups.

With regard to non-school-related activities, we observe the following. RAEs on sports exercises turn from being slightly positive, for the "Below 12" age group, to being slightly negative for the "13-14" age category. At the same time, the effect of absolute age is ambiguous: it is always positive, but its magnitude changes over time; it is the same for groups of students "Below 12" and "13-14", whereas it is smaller for the age group with the oldest students. When we investigate TV consumption during the week and during the weekend (Outcomes (6) and (7)), we find a negative effect of relative age only for younger cohorts, whereas there is a positive effect for older cohorts. The effect of absolute age appears to be positive for the two youngest age groups, while it becomes negative for students in the group "Above 15."

Lastly, similar to the previous outcome variables of TV consumption, we find that RAEs on videogame consumption are negative for students "Below 12" and positive for

Table 2.6: Relative age by age category, by standardized outcomes; two-stage least square

	Relative Age (1)	Absolute Age (2)	N (3)
Outcomes			
(1) Perceived performance			
Below 12	0.169^{***}	-0.002	159,430
13-14	0.079^{***}	0.087^{***}	165,847
Above 15	-0.040^{**}	0.159^{***}	155,397
(2) Schoolwork pressure			
Below 12	-0.117^{***}	0.051^{***}	159,305
13-14	-0.077^{***}	0.106^{***}	165,664
Above 15	-0.031^{*}	0.061^{***}	155,326
(3) Homework, week-days			
Below 12	-0.206^{***}	0.088	30,707
13-14	0.036	-0.042	30,915
Above 15	0.108	-0.075	28,771
(4) Homework, weekend-days			
Below 12	-0.240^{***}	0.126^{***}	30,450
13-14	-0.014	0.066	30,796
Above 15	-0.133	0.062	28,688
(5) Sports exercises, whole week			
Below 12	0.026	0.177^{***}	123,995
13-14	-0.045^{***}	0.177^{***}	130,473
Above 15	-0.015	0.098^{***}	122,809
(6) TV, week-days			
Below 12	-0.058^{***}	0.086^{***}	153,567
13-14	0.004	0.026	161,480
Above 15	0.061^{***}	-0.134^{***}	150,879
(7) TV, weekend-days			
Below 12	-0.072^{**}	0.101^{***}	152,558
13-14	-0.027	0.041^{***}	159,439
Above 15	0.044^{***}	-0.132^{***}	149,167
(8) Videogame, week-days			
Below 12	-0.072^{***}	0.081^{***}	119,362
13-14	0.017	-0.031	128,027
Above 15	0.089^{***}	-0.157^{***}	119,482
(9) Videogame, weekend-days			
Below 12	-0.065^{***}	0.102^{***}	118,471
13-14	-0.026	-0.020	126,117
Above 15	0.052^{***}	-0.163^{***}	117,862

Source: HBSC data. Note: All of the analyses include two sets of variables: (i) demographic control variables (dummy for being female, dummy for having both parents at home, and dummies for medium and high socio-economic status), and (ii) fixed effects (wave, country, season-of-birth). Standard errors clustered on class are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

students “Above 15”. In addition, the effect of absolute age is positive for the youngest age group and becomes negative for the oldest age group.

Overall, these results provide three interesting insights. First, with respect to schoolwork pressure, homework, and mostly to perceived performance, RAEs tend to become smaller or disappear with the increase in absolute age. Second, with respect to the time consumption for watching TV and playing videogames, we find a switch in the sign of RAEs. RAEs on these four outcome variables for media consumption are negative for younger age groups and become positive for the oldest age group. Third, if we intend the investigation of absolute age by age categories as an investigation of non-linear effects of absolute age, we observe that, except for perceived performance and schoolwork pressure, absolute age seems to have a linear relationship with the explored outcomes: initial positive effects of absolute age tend to reduce and become negative throughout adolescence.

We have conducted further additional heterogeneity analyses to investigate whether relative age might have non-linear returns. However, with the exception of RAEs on homework during weekend days, we do not find evidence of such non-linear effects. These results can be provided upon request.

2.5 Conclusions

Relative age is an important determinant of school performance and it is well established that relatively young students obtain lower grades. It is important to study other possible effects of relative age (RAEs) on and related educational disadvantages of relatively young students. Although we cannot observe how relative age (and thus lower school performance) translates into students’ experiences and how it affects their decision-making, we can observe how it affects multiple behaviors that have not been investigated so far. More concretely, we analyze data from the Health Behaviour in School Aged Children (HBSC) international survey (33 countries and 4 waves) and use a two-stage least square (2SLS) regression to study how relative age and absolute age affect students’ perceived performance and schoolwork pressure – as a proxy for strain, the time students spend on school-related activities (i.e. homework), and the time they spend on non-school-related activities (i.e. sports exercises, watching TV, and playing videogames). The HBSC data allows us to use instrumental variables for both absolute and relative age. Moreover, we compare results from 2SLSs with those from OLSs to learn insights on the importance of accounting for potential endogeneity issues, which would otherwise lead to biased estimates.

We find that relatively old students feel that they are perceived by their teachers as being over-performing compared to their younger classmates; moreover, they feel lower schoolwork pressure, and seem to spend less time watching TV. However, they do not spend more time doing homework, but there are positive RAEs on time spent in sports activities.

Based on these results, does relative age make Jack a dull student? Is the relatively young student Jack "all work and no play"? Our analyses of HBSC data lead to an ambiguous answer. It does not seem that relatively young students use their time much differently from relatively older classmates; despite their lower grades, they do not appear to spend more time on school-related activities and less time on non-school-related activities.

With respect to absolute age, we find that, firstly, higher absolute age goes along with lower perceived performance and higher schoolwork pressure. Students with higher absolute age spend more time doing homework and watching TV, but less time playing videogames. With respect to sports, the instrumental variable approach shows only small differences in absolute age.

Heterogeneity analyses by age group show that most RAEs are predominantly prevalent at younger absolute ages. Therefore, these results suggest that gaps in perceived performance and schoolwork pressure as well as differences in time allocation for homework tend to disappear as students grow older. Interestingly, RAEs on the four outcome variables for TV and videogame use reverse in older age groups.

The investigations on relative age present three limitations. First, the analyses on one of the most interesting outcomes, that is, time spent doing homework, are characterized by a small sample size (about one fourth of the average sample sizes for the various analyses in this paper), which limits the statistical power. Second, we have no information on the time use for the whole day; therefore, we cannot interpret the outcomes (and thus the findings) in terms of complements and substitutes (e.g. videogame and TV time as substitute to homework time). While it is not unfounded to consider the time spent doing homework as a complementary school-related activity or watching TV rather as a substitute activity, more time for other recreational activities, such as sport activities, is not necessarily a substitute for other school-related activities, since – in turn – it could be a substitute activity of other non-school-related activities.[22] Third, our instrumental variable approach is still exposed to the possible infringement of the monotonicity assumption of the instrument (Barua and Lang, 2016), due to the fact that students born around the cutoff date are most likely non-regular students.

[22] We thank an anonymous referee for suggesting this issue.

Results from this paper suggest two interesting venues for future studies. First, the outcomes explored in this paper are quite visible to third persons (e.g. students' parents or teachers), and relative age gaps on these outcomes tend to disappear or – especially for time spent on TV and videogames – reverse in time. Regarding the latter finding, it may be that the decrease in maturity advantages cause some anxiety–and reduces well-being gaps (Fumarco et al., 2020), so relatively old students spend more time with individual activities as a coping mechanism, in particular videogames (the role of videogames as coping mechanism is recently investigated in von der Heiden et al. (2019), Blasi et al. (2019), Plante et al. (2019)). This is just our speculation, future studies should further dig in the evolution of the relationship between relative age, time use and mental health. Additionally, future research would benefit from investigating the time allocation within the entire day, so that the outcomes being analyzed could be more clearly identified as potentially substitute or complements; data of this kind are available in countrywide time use surveys.

Appendix A.1

Table A.1: Number of observations, by country, wave and cutoff dates by country

	N by wave				
Country	2001/2 N	2005/6 N	2009/10 N	2013/14 N	N by country
Austria	4,118	4,771	4,683	3,313	16,885
Belgium, Flanders	1,608	3,051	2,956	3,135	10,750
Belgium, Wallonia	3,018	3,589	3,080	4,845	14,532
Bulgaria	-	4,811	-	4,639	9,450
Croatia	4,270	4,520	6,058	5,507	20,355
Czech Republic	4,974	-	4,284	5,041	14,299
Denmark, mainland	4,474	5,326	3,924	3,783	17,507
England	3,117	4,697	3,413	5,136	16,363
Estonia	3,215	4,124	4,131	4,001	15,471
Finland	5,143	5,143	6,494	5,810	22,590
France	7,223	5,710	5,471	5,175	23,579
Greece	3,102	-	4,801	4,078	11,981
Greenland	-	-	198	141	339
Hungary	-	3,450	4,569	3,737	11,756
Iceland	-	8,494	8,651	9,128	26,273
Ireland	1,956	3,716	3,858	3,366	12,896
Italy	4,313	3,863	4,734	3,906	16,816
Latvia	3,206	4,059	4,021	4,931	16,217
Lithuania	5,577	5,574	5,211	-	16,362
Luxembourg	-	2,886	2,968	2,214	8,068
Macedonia	3,447	4,643	3,400	4,103	15,593
Malta	1,867	-	-	2,227	4,094
Netherlands	3,769	3,138	3,174	3,844	13,925
Norway	4,943	4,536	4,017	3,109	16,605
Poland	6,239	5,434	4,185	4,060	19,918
Scotland	4,317	6,098	6,540	5,609	22,564
Slovakia	-	243	4,475	4,989	9,707
Slovenia	3,894	5,005	5,288	4,795	18,982
Spain	5,418	7,729	3,792	3,574	20,513
Sweden	3,746	4,332	6,627	7,439	22,144
Switzerland	4,083	4,204	5,694	5,500	19,481
Ukraine	3,943	4,859	5,345	3,095	17,242
Wales	3,739	4,320	5,193	5,035	18,287
Total N	108,719	132,325	141,235	139,265	521,544

Source: HBSC data. Note: Flanders and Wallonia as well as Denmark mainland and Greenland hold separate surveys within Belgium and Denmark, respectively.

Table A.2: Cutoff dates by country

Country	Cutoff date
Austria	Sep 1st
Belgium, Flanders	Jan 1st
Belgium, Wallonia	Jan 1st
Bulgaria	Jan 1st
Croatia	Apr 1st
Czech Republic	Sep 1st
Denmark	Jan 1st
England	Sep 1st
Estonia	Oct 1st
Finland	Jan 1st
France	Jan 1st
Greece	Jan 1st
Greenland	Jan 1st
Hungary	Jul 1st
Iceland	Jan 1st
Ireland	Jan 1st
Italy	Jan 1st
Latvia	Jan 1st
Lithuania	Jan 1st
Luxembourg	Sep 1st
Macedonia	Jan 1st
Malta	Jan 1st
Netherlands	Oct 1st
Norway	Jan 1st
Poland	Sep 1st
Scotland	Mar 1st
Slovakia	Sep 1st
Slovenia	Jan 1st
Spain	Jan 1st
Sweden	Jan 1st
Switzerland	Jul 1st
Ukraine	Jan 1st
Wales	Sep 1st

Table A.3: Educational settings sources

Country	Source
Croatia	Sakic, M., Burusic, J., Babarovic, T. (2013). The relation between school entrance age and school achievement during primary schooling: evidence from Croatian primary schools. British Journal of Education Psychology, 83, 651-663.
Estonia	Toomela, A., Kikas, E., Mõttus, E. (2006). Ability grouping in schools: A study of academic achievements in five schools in Estonia. Trames, 10, 32-43.
Greenland	Statistics Greenland (2015). Greenland in Figures – 2015. Nuuk: Statistics Greenland.
Greenland	Rex, K. F., Larsen, N. H., Rex, H., Niclasen, B., Pedersen, M. L. (2014). A national study on weight classes among children in Greenland at school entry. International Journal of Circumpolar Health, 73, 1-6.
Luxembourg	Ministry of Education correspondence, private correspondence
Multiple	`https://eurydice.eacea.ec.europa.eu/national-education-systems`
Multiple	`http://www.oecd.org/edu/bycountry/`
Multiple	`https://op.europa.eu/en/publication-detail/-/publication/10f14860-12a3-4f9e-b10a-32bee875420d`
Multiple	`https://www.nfer.ac.uk/media/1318/44414.pdf`
Netherlands	Plug, E. J. S. (2001). Season of birth, schooling and earnings. Journal of Economic Psychology, 22, 641-660.
Norway	Lien, L., Tambs, K., Oppedal, B., Heyerdahl, S., Bjertness, E. (2005). Is relatively young age within a school year a risk factor for mental health problems and poor school performance? A population-based cross-sectional study of adolescents in Oslo, Norway. BMC Public Health, 5, 1-8.
Norway	Solli, I. F. (2017). Left behind by birth month. Education Economics, 25, 323-346.
Scotland	Gamoran, A. (2002). Standards, inequality & ability grouping in schools. Mimeo.
Ukraine	`https://www.classbase.com/countries/Ukraine/Education-System`

Table A.4: Matrix of pairwise correlations between relative age, season of birth and expected relative age

Variables	**1**	**2**	**3**
1 Relative age	1.000		
2 Season of birth	-0.214^{***}	1.000	
3 Expected relative age	-0.299^{***}	0.594^{***}	1.000

Source: HBSC data. Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A.5: Nonparametric balance test on demographic characteristics; ordinary least square of demographic characteristics on expected relative age and fixed effects

Variables	Gender	Both parents at home	SES	Low	Medium	High
				Disaggregated SES		
	(1)	(2)	(3)	(4)	(5)	(6)
ERA	0.001	-0.001	-0.001^{***}	0.001	0.001^{*}	-0.001^{***}
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Fixed effects						
Country	✓	✓	✓	✓	✓	✓
Wave	✓	✓	✓	✓	✓	✓
Season-of-birth	✓	✓	✓	✓	✓	✓
N	521,544	502,974	521,544	521,544	521,544	521,544
Adj. R-squared	0.001	0.040	0.139	0.098	0.010	0.101

Note: 'SES' stands for socio-economic status, while 'ERA' stands for expected relative age. Standard errors clustered on class are in parenthesis. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Appendix A.2

Perceived performance and schoolwork pressure

Table A.6: Reduced form and first stage, from the 2SLS, on standardized perceived performance and schoolwork pressure

Variables	Relative age	Absolute age	Perceived performance	Relative age	Absolute age	Schoolwork pressure
	First stage	First stage	Reduced form	First stage	First stage	Reduced form
	(1)	(2)	(3)	(4)	(2)	(6)
ERA	-0.034***	-0.032***	-0.004***	-0.035***	-0.032***	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
EAA	-0.001	0.958***	-0.093***	-0.001	0.958***	0.121***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
N	480,674			480,295		
2SLS tests						
Under-identification test, Lagrange-Multiplier statistic [p-value]	4538.779 [0.000]			4542.725 [0.000]		
Weak-identification test, F statistic	3490.339			3496.271		
Over-identification test of all instruments, Hansen J statistic [p-value]	7.151 [0.711]			8.361 [0.593]		

Source: HBSC data. Note: 'ERA' stands for expected relative age. 'EAA' stands for expected absolute age. The analyses include two sets of variables: (i) demographic control variables (dummy for being female, dummy for having both parents at home, and dummies for medium and high socio-economic status), and (ii) fixed effects (wave, country, season-of-birth). Over-identification restrictions test of all instruments, had we disaggregated ERA in 11 dummy variables as in Fumarco et al. (2020) and Fumarco and Baert (2019). Standard errors clustered on class are in parenthesis. $^{***}\ p < 0.01$, $^{**}\ p < 0.05$, $^{*}\ p < 0.1$.

School-related activities

Table A.7: Reduced form and first stage, from the 2SLS, on standardized time spent doing homework, during both week-days and weekend-days

Variables	Relative age First stage (1)	Absolute age First stage (2)	Homework week-days Reduced form (3)	Relative age First stage (4)	Absolute age First stage (2)	Homework weekend-days Reduced form (6)
ERA	-0.020***	-0.017***	0.001	-0.020**	-0.017***	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
EAA	0.007***	0.974***	0.025***	0.007***	0.974***	0.061***
	(0.002)	(0.002)	(0.003)	(0.001)	(0.002)	(0.003)
N	90,393			89,934		
2SLS tests						
Under-identification test, Lagrange-Multiplier statistic [p-value]	333.591 [0.000]			334.469 [0.000]		
Weak-identification test, F statistic	198.071			198.550		
Over-identification test of all instruments, Hansen J statistic [p-value]	6.684 [0.755]			18.362 [0.049]		

Source: HBSC data. Note: 'ERA' stands for expected relative age. The analyses include two sets of variables: (i) demographic control variables (dummy for being female, dummy for having both parents at home, and dummies for medium and high socio-economic status), and (ii) fixed effects (wave, country, season-of-birth). Over-identification restrictions test of all instruments, had we disaggregated ERA in 11 dummy variables as in Fumarco et al. (2020) and Fumarco and Baert (2019). Standard errors clustered on class are in parenthesis. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A.8: Reduced form and first stage, from the 2SLS, on standardized time spent doing sports exercises during the whole week

Variables	Relative age	Absolute age	Sports exercises, whole week
	First stage	First stage	Reduced form
	(1)	(2)	(3)
ERA	-0.037***	-0.035***	-0.005***
	(0.001)	(0.001)	(0.001)
EAA	-0.003***	0.955***	0.003**
	(0.001)	(0.001)	(0.001)
N	377,277		
2SLS tests			
Under-identification test, Lagrange-Multiplier statistic [p-value]	4101.012 [0.000]		
Weak-identification test, F statistic	3383.948		
Over-identification test of all instruments, Hansen J statistic [p-value]	4.811 [0.903]		

Source: HBSC data. Note: 'ERA' stands for expected relative age. 'EAA' stands for expected absolute age. The analyses include two sets of variables: (i) demographic control variables (dummy for being female, dummy for having both parents at home, and dummies for medium and high socio-economic status), and (ii) fixed effects (wave, country, season-of-birth). Over-identification restrictions test of all instruments, had we disaggregated ERA in 11 dummy variables as in Fumarco et al. (2020) and Fumarco and Baert (2019). Standard errors clustered on class are in parenthesis. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A.9: Reduced form and first stage, from the 2SLS, on standardized time spent watching TV, during both week-days and weekend-days

Variables	Relative age	Absolute age	TV, week-days	Relative age	Absolute age	TV, weekend-days
	First stage	First stage	Reduced form	First stage	First stage	Reduced form
	(1)	(2)	(3)	(4)	(2)	(6)
ERA	-0.034***	-0.032***	-0.001	-0.034***	-0.032***	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
EAA	-0.001	0.958***	0.027***	-0.001***	0.958***	0.042***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
N	465,926			461,164		
2SLS tests						
Under-identification test, Lagrange-Multiplier statistic [p-value]	4341.885 [0.000]			4204.250 [0.000]		
Weak-identification test, F statistic	3285.544			3183.773		
Over-identification test of all instruments, Hansen J statistic [p-value]	14.685 [0.144]			18.561 [0.046]		

Source: HBSC data. Note: 'ERA' stands for expected relative age. 'EAA' stands for expected absolute age. The analyses include two sets of variables: (i) demographic control variables (dummy for being female, dummy for having both parents at home, and dummies for medium and high socio-economic status), and (ii) fixed effects (wave, country, season-of-birth). Over-identification restrictions test of all instruments, had we disaggregated ERA in 11 dummy variables as in Fumarco et al. (2020) and Fumarco and Baert (2019). Standard errors clustered on class are in parenthesis. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A.10: Reduced form and first stage, from the 2SLS on standardized time spent playing videogames, during both week-days and weekend-days

Variables	Relative age	Absolute age	Videogame, week-days	Relative age	Absolute age	Videogame, weekend-days
	First stage	First stage	Reduced form	First stage	First stage	Reduced form
	(1)	(2)	(3)	(4)	(2)	(6)
ERA	-0.037***	-0.035***	0.001	-0.037***	-0.035***	0.001**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
EAA	-0.003***	0.955***	-0.015***	-0.003***	0.955***	-0.010***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
N	366,871			362,450		
2SLS tests						
Under-identification test, Lagrange-Multiplier statistic [p-value]	4158.530 [0.000]			3999.174 [0.000]		
Weak-identification test, F statistic	3440.725			3308.249		
Over-identification test of all instruments, Hansen J statistic [p-value]	4.983 [0.892]			5.725 [0.838]		

Source: HBSC data. Note: 'ERA' stands for expected relative age. 'EAA' stands for expected absolute age. The analyses include two sets of variables: (i) demographic control variables (dummy for being female, dummy for having both parents at home, and dummies for medium and high socio-economic status), and (ii) fixed effects (wave, country, season-of-birth). Over-identification restrictions test of all instruments, had we disaggregated ERA in 11 dummy variables as in Fumarco et al. (2020) and Fumarco and Baert (2019). Standard errors clustered on class are in parenthesis. *** p <0.01, ** p <0.05, * p <0.1.

Chapter 3

Young, Gifted and Lazy? The Role of Ability and Labor Market Prospects in Student Effort Decisions

This paper examines the decision-making process of students from an economic perspective to understand the determinants of an individual's willingness to provide effort. Our theoretical model predicts that ability and job market prospects are positive determinants. Analyzing a novel dataset on thousands of German students, however, we instead find that ability has a significantly negative effect on effort. It seems that the marginal gain of increasing effort in terms of higher expected income after studying is lower for high-ability students compared to low-ability students. In regard to the second determinant, the evidence rejects a similar argument, according to which great job market prospects may impair student effort. Applying an instrumental variable approach based on official unemployment data on regional labor markets, we can confirm our prediction on the positive role of perceived employment prospects in actual student behavior.[23]

Keywords: higher education, effort, study time, leisure, ability, labor market data

JEL Classification: I23, J22, J24

[23] This chapter is joint work with Adrian Chadi and Marco de Pinto (Chadi et al., 2019).
We are grateful for helpful comments by participants of the 8th Workshop on Economics in Trier, of the 1st NEPS User Conference in Bamberg, of the Workshop on Subjective Survey Data in Labour Market Research in Trier, of the Colloquium on Personnel Economics in Zurich, of the Royal Economic Society Annual Conference in Bristol, of the Workshop on Labor Economics in Trier, of the European Economic Association Annual Congress in Lisbon, of the IAW Workshop in Tübingen and by seminar participants in Luxembourg and Trier. We also thank two anonymous referees for their useful comments as well as Laszlo Goerke and Jochen Michaelis for inspiring discussions.

3.1 Introduction

The circumstances under which individuals strive are central to scientific research on human behavior. The economic approach suggests that individuals provide high efforts whenever the expected benefits of an activity exceed the expected costs. However, we know little about the determinants of effort outside of experimental laboratories. Similarly, situations when individuals – instead of providing high efforts to maximize their economic gain – make the decision to simply lean back remain largely unexplored. In some cases, individuals with particular potential and great prospects may show high motivation to provide extraordinary performance, while in other cases, a positive outlook may actually lower effort levels, as it is possible to benefit from reduced effort costs while still obtaining a satisfactory level of achievement. By focusing on students from higher education institutions, we analyze individual effort decisions, which allows us to not only shed light on the determinants of human behavior in this particular educational context, but also beyond.[24]

The decision situation faced by students in the system of higher education has a particular facet that makes it very interesting from an economic standpoint: Both society and the individual student benefit from educational achievement. The more educational achievement can be attained, the higher the individual labor market earnings are because of increased human capital (e.g. Wolpin, 1977, Kroch and Sjoblom, 1994, Chevalier et al., 2004), which thereby fosters overall economic prosperity. To achieve this, students can choose individual effort levels as a major determinant of educational outcome (e.g. Stinebrickner and Stinebrickner, 2008). This leads to a scenario in which students have incentives for putting large amounts of effort into studying, benefiting both the economy and society. In reality, however, indications suggest a lack of effort levels among students, such as declining amounts of time spent on studying (e.g. Babcock and Marks, 2011) and increasing study durations in numerous countries (e.g. Brunello and Winter-Ebmer, 2003, Bound et al., 2012, Garibaldi et al., 2012). This raises the questions regarding the determinants of study behavior and how the drivers of students' commitment to performing well can be identified.

[24] Microeconomic models of individual effort decisions typically include assumptions on the role of an individual's potential, such as ability, without providing references to empirical evidence. This is not surprising, given a lack of studies that focus specifically on the question of how ability affects effort. While a lot of evidence on determinants of individual effort levels comes from economic laboratory experiments, researchers here oftentimes inspect mindless tasks to purposely render certain inputs like ability irrelevant. Economic researchers of field data often fall back on proxies like absenteeism (see e.g. Ichino and Riphahn, 2005, Cornelissen et al., 2011b, Block et al., 2014, Chadi and Goerke, 2018), or hours worked (see Bell and Freeman, 2001), given the importance of effort and its determinants in a variety of different research contexts, as e.g. workers' performance in firms. Another option for researchers to gather evidence from the field is professional sports, which allows testing economic predictions on effort decisions in non-laboratory data (see e.g. Lackner et al., 2015).

While there is a sizable literature on the outcomes of studying, there are not many studies dealing with economic decision-making in terms of student effort in higher education. Many of the contributions to research on educational achievement of university or college students focus on study outputs, such as grades, and analyze potential determinants, such as financial incentives or working during school (e.g. Stinebrickner and Stinebrickner, 2003, Kalenkoski and Pabilonia, 2010).[25] Few papers provide a combination of empirical analysis and theoretical modeling, in which the latter considers the crucial role of student effort as a contributor to academic success (see e.g. Löfgren and Ohlsson, 1999, Krohn and O'Connor, 2005, Bandiera et al., 2015).[26] One example in this context is a study by Oettinger (2002), who discusses how university students make strategic decisions on effort levels, for which he assumes that incentives to provide effort increase in ability. In their study on student performance, Leuven et al. (2010) also take the role of effort into account but mostly focus on passing rates and how student performance can be raised by financial incentives. Similarly, Non and Tempelaar (2016) consider both effort and academic success in their empirical study on the role of time preferences, just as Chevalier et al. (2018) do in their study on incentive schemes. There is also recent theoretical work on the role of examination rules for students' effort decisions (Michaelis and Schwanebeck, 2016).

Apart from that, we concur with the conclusion of Delaney et al. (2013) who see a clear lack of knowledge on student inputs, despite the high level of interest in explaining study outcomes. While these authors provide the first empirical investigation into the determinants of student behavior in higher education using across-subject data, they omit two determinants that we consider to be as important as they are unclear in their actual role for effort decisions: ability and labor market prospects. Intuitively, one could expect that high-ability types have strong incentives to provide extra effort, as they benefit more from educational achievement, which also seems to be in line with the evidence from the above-mentioned studies. On the other hand, they could also use their promising situation to reduce effort when they are satisfied with a certain level of achievement. Similar arguments apply for job market prospects in general, which could also affect student behavior and help to explain low effort levels. Given the unclear relationships, we provide the first comprehensive discussion, theoretically and empirically, on how these factors affect students' effort.

[25] Other indicators of study outputs in the context of higher education are graduation rates (e.g. Light and Strayer, 2000) and study durations (e.g. Gunnes et al., 2013).

[26] Note that there are also some studies discussing the importance of student effort for educational achievement among pupils before they enter higher education, such as Metcalfe et al. (2011) as well as De Fraja et al. (2010) who also point out a lack of research on the role of student effort.

As a potentially important aspect, we consider multiple dimensions of student effort in our discussion. Whereas previous educational studies often focused on study time measured via lecture attendance, the role of this factor in educational achievement appears to be unclear.[27] Given the heterogeneity of empirical findings in this context, we scrutinize whether study time is sufficient for capturing individual effort and question the underlying assumption that investing the same amount of time means investing the same amount of effort. Arguably, any given hour spent in the library or in the lecture room may consist of only focused learning, but it may just as well consist of only idle daydreaming. We therefore propose a distinction into a quantitative and a qualitative dimension of effort in order to learn more about the complex factor that effort certainly is. While in our theoretical discussion we distinguish between study time (quantitative dimension) and learning intensity (qualitative dimension), we attempt to capture the quantitative component via comprehensive time-use data and the qualitative component via subjective data on self-assessed effort levels in our empirical investigation.

In our theoretical modeling of student decision-making, we make some basic assumptions that conform to the previous literature. Students decide about both effort dimensions anticipating that higher effort is associated with a utility decline today, but improves educational achievements and hence increases expected income and utility after studying. Whether high-ability students provide less or more effort compared to low-ability students depends on two factors. First, considering each effort dimension separately, high-ability students have an incentive to increase effort, such as study time (at the expense of leisure), because this raises utility in the future, i.e. the substitution effect (SE). At the same time, however, high-ability students have an incentive to reduce effort because their high abilities per se ensure relatively good educational achievement and thus a relatively high level of expected income, i.e. the income effect (IE). Second, the way both effort dimensions are interlinked is crucial. If they were complements, high-ability students that provide high learning intensity would also choose a high study time, compared to low-ability students. If both dimensions were substitutes, however, high learning intensity would come at a price of lower study time and vice versa. These mechanisms also hold for our second determinant, i.e. job market prospects.

To gain testable predictions, we assume that both the students' utility and educational production function are of a Cobb-Douglas type. This implies that a.) the SE dominates the IE and that b.) both effort dimensions are complements. As such, we expect that

[27] See Grave (2011) for a comprehensive study on the role of students' time allocation in educational achievement. While she finds rather positive relationships between the latter and both lecture attendance and self-study time for her data on German university students, Dolton et al. (2003) find a more positive role of attendance compared to self-study using data on Spanish university students. In contrast, Bratti and Staffolani (2013) find the opposite for Italian university students and view self-study as a more important predictor of academic performance than attendance.

high-ability students provide higher effort (study time and learning intensity) compared to low-ability students. In addition, better job market prospects should increase the students' effort in both dimensions during academic studies. The mechanisms in our model translate to many other economics contexts and thereby provide us with a general framework upon which we can discuss our empirical results. This is particularly helpful as our empirical findings for individuals in higher education do indeed deviate in some respect from expectations one may have at first glance.

To test theoretical predictions, we explore data from a broadly conceived investigation of students in Germany's system of higher education, the National Educational Panel Study (NEPS). The students' cohort of the NEPS has not been utilized for similar purposes so far and allows us to inspect the role of ability, as an example, in ways not possible in most cross-subject datasets. Regarding this key student input, we can exploit data from comprehensive competence testing of university students to establish a measure that allows us to inspect this determinant of university students' behavior without having to rely on proxies such as previous grades, which are likely related to an individual's overall attitude towards providing effort. As we argue in our paper, this could be a particular problem for any attempt to find out about the actual impact of individual ability on effort. Finally, we provide evidence on the direct effects of job market prospects on student effort, as the outlook on future earnings reflects the channel through which students take their economic gains of studying into account. We thereby elaborate on the work of Brunello et al. (2004) who argue that subjectively expected returns to education are a key determinant for university students' decision-making regarding educational attainment.[28]

Our results from analyzing the NEPS data reject the prediction that ability positively affects effort levels. Instead, the evidence conforms to the notion that high-ability students use their advantage over low-ability ones to obtain additional utility by having more leisure time. We find that the higher the ability is, the lower self-assessed effort levels and weekly self-study hours are. As the latter predicts educational achievement in our data more strongly than the other activities, such as lecture attendance, this empirical result supports the notion of the 'lazy genius' who puts comparatively little effort into studying. Going back to our model, this speaks for a relatively strong IE and/or a relatively weak SE. Regarding job market prospects, we not only examine standard regression results but also apply an instrumental variable (IV) approach to address the potential reverse causality between effort and labor market prospects. To that end, we exploit official unemployment data reflecting variations in regional labor market conditions. We merge the

[28]In a similar fashion, many researchers promote the use of subjective data on students' beliefs and expectations regarding the role of the labor market in student behavior, such as Betts (1996), Wolter (2000), Botelho and Pinto (2004), Webbink and Hartog (2004), Jensen (2010), Bonnard et al. (2014), Brodaty et al. (2014), Stinebrickner and Stinebrickner (2014), Huntington-Klein (2015).

data using information on prospective jobs and university location. The results from applying this approach align with those from running standard regressions and suggest that great job prospects positively influence effort, which confirms our theoretical prediction. Vice versa, we interpret our finding in such a way that not having good prospects may frustrate students, leading to decreased motivation for putting in high efforts into studying, which may contribute to the phenomenon of prolonged study durations, as argued by other researchers (e.g. Aina et al., 2011).

3.2 Theoretical model

3.2.1 Set-up

We analyze the study behavior of an individual by using a two-period model. While being a student in the present period 1, the individual expects to enter the labor market in the future period 2 to earn income. The individual's utility function is assumed to be:

$$U = V(I^1, L^1) - C(e) + \beta V(I^2, L^2), \qquad 0 < \beta \leq 1, \tag{3.1}$$

where I^1 (I^2) denotes income in period 1 (expected income in period 2), L^1 (L^2) represents leisure in period 1 (leisure in period 2) and β is the discount rate.[29] We assume that sub-utility V increases in income and leisure at a decreasing rate. e measures the student's learning intensity. Intuitively, e indicates how diligent the student is and how hard s/he works during the time span that s/he has scheduled for studying. Learning intensity is associated with utility costs (or disutility) measured by C, which is increasing and convex, $C_e, C_{ee} > 0$, where subscripts denote partial derivatives.

Income in period 1 is assumed to be exogenously given, while leisure in period 1 reads $L^1(s) = T - s$. The endogenous variable s represents the time that the individual spends on academic studies. Given the (exogenous) time stock T, which also summarizes the time required for other activities besides studying (e.g. student employment), s determines the amount of leisure a student has. In the literature, study time is often considered as an effort indicator. In our setup, however, s constitutes only a quantitative dimension of effort, while its qualitative dimension is captured by learning intensity e.

In period 2, expected income depends on the individual's achievement during academic studies (for a similar assumption see Löfgren and Ohlsson, 1999, De Fraja and Landeras, 2006, De Fraja et al., 2010). The student's achievement is typically represented

[29]To keep our analysis as simple as possible, we consider the discount rate as exogenously given and, in particular, independent of students' characteristics such as their abilities. For the same assumption, see, for instance, De Fraja et al. (2010).

by the educational production function (EPF). Following a large strand of literature on the determinants of study success (for an excellent review see Brewer and McEwan, 2010), we assume that the student's achievement positively depends on learning intensity e, study time s and the student's time invariant and exogenously given ability level a.

The EPF can be formalized as:

$$Y = Y(e, s, a, X), \tag{3.2}$$

with $Y_e, Y_s, Y_a > 0$. The vector X captures all other factors that influence Y, for example family background or the quality of the university. Note that the sign of the cross derivative Y_{se} is undetermined. As such, both effort dimensions are interrelated and can be either complements ($Y_{se} > 0$) or substitutes ($Y_{se} < 0$).

Expected income in period 2 is thus given by:

$$I^2 = \delta Y(e, s, a, X), \qquad \delta \geq 0. \tag{3.3}$$

The parameter δ captures the student's job market prospects. If these prospects are relatively good (bad), i.e. δ is relatively high (low), a given level of educational achievement Y implies a relatively high (low) value of expected income I^2. To disentangle the effects of a student's ability and job market prospects on her/his effort during studying, we focus on prospects that are not related to a student's ability. For example, a student's job market prospects could be determined by predictions about future labor market conditions or subjective perceptions thereof. As such, we interpret δ as a measure for non-ability-related labor market expectations and consider it as exogenously given and thus independent of a (as well as of s and e). Note that job market prospects can change over time, such that δ could increase or decrease during academic studies. Future leisure L^2 is also assumed as exogenously given.[30]

With these components at hand, we can rewrite the individual's utility function as:

$$U(s, e, \delta, a) = V(I^1, L^1(s)) - C(e) + \beta V(I^2(e, s, \delta, a)), L^2), \tag{3.4}$$

where we have suppressed some of the variables to save notation.

[30]This assumption is made for simplification. It can be justified because working hours are predetermined in highly regulated labor markets such as in Germany. Moreover, an endogenous determined L^2 would not qualitatively alter our results.

3.2.2 Optimization

At the beginning of period 1, the student decides about time allocation during academic studying, i.e. the student chooses how much time (of T) will be spent on studying s. The residual time is used for leisure and other (exogenous) activities. In addition, the student sets learning intensity e. Both decisions are made to maximize total utility U.

Differentiating (3.4) with respect to s implies:

$$U_s = -\underbrace{V_{L^1}(s)}_{\equiv MC_s(s)} + \underbrace{\beta V_{I^2}(s, e, \delta, a)\delta Y_s(e, s, a)}_{\equiv MG_s(e,s,a,\delta)} = 0, \tag{3.5}$$

where MC_s and MG_s denote the marginal costs and the marginal gains of an increase in study time, respectively. This implies that study time is chosen such that the utility decrease today (due to reduced leisure) is exactly offset by the utility increase in the future (due to improved educational achievements and thus raised expected income). Note that (3.5) pins down the utility maximizing study time for any given level of learning intensity, ability and labor market prospects: $\overline{s} = s(e, a, \delta)$.

The first-order condition with respect to learning intensity reads:

$$U_e = -\underbrace{C_e(e)}_{\equiv MC_e(e)} + \underbrace{\beta V_{I^2}(s, e, \delta, a)\delta Y_e(e, s, a)}_{\equiv MG_e(e,s,a,\delta)} = 0, \tag{3.6}$$

with MC_e and MG_e representing the marginal costs and the marginal gains of an increase in e, respectively. As a result, learning intensity is chosen such that the utility decrease today (due increased costs C) is balanced by the utility increase in the future (due to improved educational achievements and thus raised expected income). Note that (3.6) pins down utility maximizing learning intensity for any given level of study time, ability and labor market prospects: $\overline{e} = e(s, a, \delta)$.

By combining $\overline{s}$ and $\overline{e}$, we can determine utility maximizing study time and learning intensity as functions of exogenous parameters only: $s^* = s(a, \delta)$ and $e^* = e(a, \delta)$.[31]

3.2.3 Comparative static analysis

How do the student's ability a and labor market prospects δ affect the student's effort choices, i.e. the utility maximizing study time and learning intensity? To provide a theo-

[31] The second-order conditions for a maximum are given by $U_{ss} < 0$, $U_{ee} < 0$ and $|H| = U_{ss}U_{ee} - U_{se}U_{se} > 0$, where $|H|$ is the determinant of the Hesse-matrix. We assume that these conditions are fulfilled.

retical answer to this question, we conduct a comparative static exercise, i.e. we consider variations in δ and a.[32]

3.2.3.1 Abilities

Let us compare a high-ability student with a low-ability student. Totally differentiating (3.5) and (3.6) and rearranging the resulting expressions yield:

$$\frac{ds}{da} = -\underbrace{\frac{1}{U_{ss}}}_{<0}\left(U_{sa} + U_{se}(Y_{se})\frac{de}{da}\right), \tag{3.7}$$

$$\frac{de}{da} = -\underbrace{\frac{1}{U_{ee}}}_{<0}\left(U_{ea} + U_{se}(Y_{se})\frac{ds}{da}\right). \tag{3.8}$$

Partial derivatives read $U_{sa} = \beta\delta(V_{I^2I^2}\delta Y_a Y_s + V_{I^2}Y_{sa})$, $U_{ea} = \beta\delta(V_{I^2I^2}\delta Y_a Y_e + V_{I^2}Y_{ea})$ and $U_{se}(Y_{se}) = \beta\delta(V_{I^2I^2}\delta Y_e Y_s + V_{I^2}Y_{se})$.

To decompose the effects on the student's behavior, we first look at the effects on $\overline{s}$ and $\overline{e}$, where $de/da = 0$ and $ds/da = 0$ hold, respectively. The sign of $d\overline{s}/da$ ($d\overline{e}/da$) depends then on the sign of U_{sa} (U_{ea}). Intuitively, there are two countervailing effects. High-ability students have higher marginal gains from effort because of an increased expected income. Therefore, they substitute leisure for study time or increase learning intensity despite the associated rise of disutility. We call this the substitution effect (SE). Given the increased expected income, however, high-ability students also have an incentive to reduce study time (learning intensity) to keep the marginal gains from s (e) constant. We call this the income effect (IE). In general, the net effect is ambiguous.

The impact on s^* and e^* can be calculated by combining (3.7) and (3.8). This yields:

$$\frac{ds^*}{da} = \underbrace{\frac{1}{|H|}}_{>0}\left(U_{se}(Y_{se}) \cdot U_{ea} - \underbrace{U_{ee}}_{<0}U_{sa}\right), \tag{3.9}$$

$$\frac{de^*}{da} = \underbrace{\frac{1}{|H|}}_{>0}\left(U_{se}(Y_{se}) \cdot U_{sa} - \underbrace{U_{ss}}_{<0}U_{ea}\right). \tag{3.10}$$

Besides the SE and IE, the effect on study time and learning intensity is driven by the interrelatedness of both which is captured by Y_{se}. Suppose that high-ability students choose to increase study time (relative to a low-ability student). If both effort dimensions were

[32] Since the student's ability is constant by definition, the results of the comparative static exercise should be interpreted as predictions of how individuals with different abilities, but otherwise identical characteristics, behave during academic studies.

complements, high-ability students would, ceteris paribus, also learn with higher intensity. If, in contrast, s and e were substitutes, increased study time would, ceteris paribus, come at a price of lower learning intensity. Given the general formulation of the EPF, the relationship between the two effort dimensions is unclear. As such, the effect on s^* and e^* is in general ambiguous.

3.2.3.2 Job market prospects

Suppose now that labor market prospects of the student improve, i.e. δ increases. We abstain from analyzing the effect on $\overline{s}$ and $\overline{e}$ and immediately compute the implications for s^* and e^*. Totally differentiating (3.5) and (3.6) implies:

$$\frac{ds^*}{d\delta} = \underbrace{\frac{1}{|H|}}_{>0} \left(U_{se}(Y_{se}) \cdot U_{e\delta} - \underbrace{U_{ee}}_{<0} U_{s\delta} \right), \tag{3.11}$$

$$\frac{de^*}{d\delta} = \underbrace{\frac{1}{|H|}}_{>0} \left(U_{se}(Y_{se}) \cdot U_{s\delta} - \underbrace{U_{ss}}_{<0} U_{e\delta} \right), \tag{3.12}$$

with $U_{s\delta} = \beta Y_s(V_{I^2 I^2} Y\delta + V_{I^2})$ and $U_{e\delta} = \beta Y_e(V_{I^2 I^2} Y\delta + V_{I^2})$.

This shows that the impact of improved labor market prospects on student's behavior depends also on a.) the interplay of SE and IE and b.) the interrelatedness of study time and learning intensity. The intuition is analog to the one described in the previous subsection. As a result, the consequences for s^* and e^* are in general ambiguous.

3.2.4 Predictions

To gain testable predictions, we have to choose an explicit formulation of the sub-utility function V and of the EPF Y. With respect to the former, we assume that the utility of income and leisure is described by a Cobb-Douglas function $V = I^\alpha L^{1-\alpha}$ with $0 < \alpha < 1$ (for a similar assumption see e.g. Mankiw, 1988). Regarding the latter, we follow the literature and assume that the EPF is given by $Y = e^{\gamma_1} s^{\gamma_2} a^{\omega}$, $0 < \gamma_1, \gamma_2 < 1$ and $\omega > 0$, which is also a Cobb-Douglas type function.[33]

These assumptions have two important implications. First, the SE will always dominate the IE, i.e. $U_{sa} > 0$, $U_{ea} > 0$, $U_{s\delta} > 0$ and $U_{e\delta} > 0$. Second, study time and learning

[33]Formalizing the EPF as a Cobb-Douglas function is a widely used assumption in the literature. See, for instance, Polachek et al. (1978), Gyimah-Brempong and Gyapong (1991) and Bishop and Wößmann (2004).

intensity are complements, i.e. $Y_{se} > 0$. It is simple to show that $U_{se}(Y_{se}) > 0$ holds, too. Using (3.9), (3.10), (3.11) and (3.12), we then find:

Prediction 1 *High-ability students choose a higher study time (lower leisure) and provide more learning intensity compared to low-ability students.*

Prediction 2 *An improvement of labor market prospects raises study time and learning intensity.*

Because of our general framework, different assumptions on the functional form of V and Y would lead to different predictions. If, for instance, IE dominated SE, high-ability students would choose less effort and an improvement of labor market prospects would decrease effort. Therefore, the effects of both on study time and learning intensity remain an empirical question.

3.3 Empirical investigation

3.3.1 Data

3.3.1.1 The NEPS

To test our theoretical predictions empirically, we exploit data from the National Educational Panel Study (NEPS). The NEPS includes more than 200 institutions of higher education in Germany, which allows for analyses based on a representative nation-wide sample. To the best of our knowledge, we are the first to use this dataset to investigate the determinants of student effort. The NEPS carries data of several cohorts covering the life span of individuals from early childhood up to further education and lifelong learning. We utilize the starting cohort 5, which provides representative data on freshman students starting in winter term 2010/2011.[34]

From the perspective of our research aims, focusing on study beginners is advantageous due to the students not being affected by their own study success. Additionally, we take the issue of potentially selective dropouts into account. Study dropouts are an ongoing problem not only in Germany but also in many other countries (see Light and

[34]Other starting cohorts focus on early childhood (cohort 1), kindergarten (cohort 2), lower secondary school (cohort 3), upper secondary school (cohort 4), and adults who are out of the education system (cohort 6). The data of the fifth one (doi:10.5157/NEPS:SC5:6.0.0) was collected by the NEPS via telephone and online. Note that from 2008 to 2013, NEPS data was collected as part of the Framework Program for the Promotion of Empirical Educational Research funded by the German Federal Ministry of Education and Research (BMBF). As of 2014, NEPS is carried out by the Leibniz Institute for Educational Trajectories (LIfBi) at the University of Bamberg in cooperation with a nationwide network. For further information see Blossfeld et al. (2011).

Strayer, 2000, Di Pietro and Cutillo, 2008). In recent years, only three quarters of students completed their studies at German universities (OECD, 2013).[35]

3.3.1.2 Variables

Our focus lies on the variables covering students' efforts. For the quantity of effort, we analyze time use data. The NEPS offers detailed information on the average time allocation of each individual's daily activities (see appendix B for more information on question wording and variable definition). Specifically, students are asked about their time spent on self-study, lecture attendance, other study-oriented activities (e.g. commuting), working, household activities and childcare. Following the findings of Bratti and Staffolani (2013), we expect self-study to be the key quantitative effort variable, but we also examine the role of lecture attendance.[36] The NEPS also provides information on study time during semester breaks, which adds to time use during the semester.

As in other empirical settings using data on university students (see e.g. Bonesrønning and Opstad, 2015), there is no exact measure of the qualitative dimension of student effort in the NEPS data. Yet, by means of a subjective assessment on individual effort levels, we are confident that we capture differences in individual effort apart from the amount of time spent on studying. Hence, as a proxy variable for the qualitative dimension of effort, we use information on how strongly each student agrees to the statement "I invest a lot of energy in being successful in my studies." Possible answers reach from "Does not apply" (1) to "Applies completely" (5) on a five-point scale. About two thirds of the students state values (3) and (4), while only a few report to invest very little or the maximum amount of energy. We refer to this variable as "self-assessed effort" in the following. Figure B.2 in the appendix visualizes the distribution of this variable as well as the effort indicator based on self-study time using histograms. For both of our key effort variables, we also show separate histograms across genders.

[35] During the time when the NEPS fieldwork commenced, some of the federal states required university students to pay tuition fees. Typically, there are no tuition fees at Germany's mainly tax-payer funded universities. This was different in the winter term of 2010/2011 in the federal states of Hamburg, Lower Saxony, North Rhine-Westphalia, Baden-Württemberg and Bavaria. Those states introduced tuition fees in the years 2006 and 2007, but the size of the tuition paid by university students was relatively small, compared to countries like the UK. Still, federal state governments decided to abolish the unpopular fees over the last years. In additional analyses, we conduct robustness checks for our main findings by including a variable for the existence of tuition fees during the period of data collection. The results are robust and are available upon request, just like all other results that we mention but do not present.

[36] Especially in the German context, there are certainly differences in the nature of these two activities, self-study and course attendance. Students may well attend classes but they are usually not required to pay attention to the lecturer or to take an active part in course lessons. It is thus quite common at German universities that some students do not attend lectures but still take the exam at the end of the semester after intense self-study and exam preparation.

We exploit data collected by the NEPS to develop a rather unique measure of ability. While many social sciences discuss the concept of ability in general and the role of ability tests in particular (see e.g. Nash, 2001), we benefit from several advantages of the NEPS data. First, in contrast to other data sources in educational research, the survey designers paid particular attention to measuring individual ability levels of each student, and conducted expensive and comprehensive competency tests. In spring of 2011, thousands of students participated for a reward of 20 Euros in voluntary tests organized at the universities. Second, this payment was unconditional on test outcomes to avoid having an economic incentive manipulate the test results. Third, the actual contents of the tests were not announced in advance, which further eliminates the possible influence of test preparation. Thus, even if students had an incentive to manipulate the outcomes of the test by putting in additional efforts, this would have hardly been possible. Arguably, in the education context, most proxies of ability are affected by individual effort levels, preventing a clean identification of actual ability.[37] The three parts of the test include reading speed, reading competency and mathematical tests, which relies on the idea that reading and mathematical skills cover most of the required core competencies of present students. Being able to conduct quantitative analyses has become increasingly important in most sciences, while an ambitious workload of reading is part of nearly every field of study. We standardize the variables of each test with a mean of zero and a standard deviation of one to generate a comprehensive variable of all three competency tests.[38]

To inspect the role of students' future employment outlooks as a potential driver of differences in effort decisions, we exploit information on job market prospects. Part of the NEPS questionnaire is the following question: "And once you do complete the degree course, what are your chances of getting a good job?" Five answers are possible, which reach from "very bad" (1) to "very good" (5). Only very few respondents (about 4%) state very bad or bad prospects (value 1 or 2), while about 46% have good prospects and about 30% of the sample have very good prospects. We build a dummy variable, which equals one if prospects are very good (5). Besides data on subjective job prospects, our dataset provides further information on students' prospective jobs. This allows us to analyze the impact of job prospects in greater depth, as we merge the NEPS data with labor

[37] A correlation analysis shows a positive relationship between high school grades and student effort. High school grades are also strongly related with our ability indicator, as better grades at high school generally go along with better competence scores. Yet, there is remaining variation in the competence test scores that cannot be explained by grades. As our ability indicator is also related to grades at university, the competence scores appear to be a reliable predictor of student performance.

[38] Due to the voluntary nature of participation in the test, data on ability is not provided for all participants of the NEPS surveys, which leads to a reduced sample size in the first part of our analysis. In turn, we can increase observation numbers in the second part of the analysis by focusing on a survey-based determinant of effort, while we make use of NEPS weights to foster representativeness of the data throughout our analyses.

market data, differentiating them by occupation and region. As job market prospects are not necessarily exogenous, we include official unemployment statistics from the Federal Employment Agency for an instrumental variable (IV) approach, which we describe in more detail later on.

The NEPS offers a large amount of data on students in Germany's system of higher education, which allows us to consider several control variables on important aspects of students' lives. As illustrated in table B.1 in the appendix, we categorize the relevant information into socio-demographic background, life circumstances, school history and university background as well as economic factors, with the latter category also including regional information. As part of the variables reflecting the factors of interest at university, we consider a subjective variable on the enjoyment of studying, which allows us to capture differences with respect to intrinsic motivation. We also consider possible measurement differences over time by using control variables for interview month throughout our analysis. We only restrict the sample by excluding outliers in regard of age (students above 40 years). Consideration of non-responses (i.e. missing values) to all the relevant survey items that we take into account leads to a sample that includes more than 4400 students who have completed all parts of the competency test. Table B.1 shows the descriptive statistics of the variables used.

To illustrate the available information on study-related time (i.e. self-study, attending classes and further study-oriented activities) and other activities, such as working, figure B.1 in the appendix shows the distribution of students' time use during the semester. The remaining share of time not reflected in one of the survey items offers us another variable of interest for our analyses. To that end, we calculate students' free-time (including weekends and sleep) by adding all hours up and deducting that number from the weekly stock of ($24 \times 7 =$) 168 hours. This leisure indicator allows us to consider the fact that some students have less time for studying than others, not because of laziness, but because of having a job, for example.

3.3.2 Student effort and ability

3.3.2.1 Main results

To investigate the role of ability for student effort decisions, we exploit our continuous measure and first inspect its basic relationship to self-assessed effort levels. We run standard regression analyses and consider relevant control variables to inspect whether the basic difference between high- and low-ability types in regards to their self-assessed effort levels is sensitive to some key characteristics. In a second step, we focus on the quantitative dimension of effort and examine various time-use variables as outcomes.

Table 3.1: Ability and (qualitative) student effort

	(1) Self-assessed effort	(2) Self-assessed effort	(3) Self-assessed effort	(4) Self-assessed effort
Ability	-0.075^{***}	-0.056^{**}	-0.053^{**}	-0.085^{***}
	(0.02)	(0.02)	(0.02)	(0.02)
Socio-demogr. controls		✓	✓	✓
School history/ life circumstances			✓	✓
Work-/university-related controls				✓
Observations	4431	4431	4431	4431
Adj. R^2	0.005	0.019	0.023	0.046

OLS estimations; NEPS weights used; robust standard errors in parentheses $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Additional dummy variables for the interview month are included in columns (2)-(4). Full results are reported in table B.4 in the appendix. Note that the appendix also contains extensive regression tables for the other analyses shown in the main part of the paper.

The main finding of table 3.1 is a significant and negative effect of ability on self-assessed effort that does not change much throughout the specifications. Column 2 adds socio-demographic controls while column 3 additionally controls for life circumstances and school history. The effect remains strong when we consider all variables together, including university- and work-related variables in column 4.

As our outcome variable here may reflect not only the qualitative dimension of effort, as in our model, but also the quantitative dimension, we inspect the consequences of controlling for the latter using time-use data. The finding of a negative ability effect holds, which is further evidence contrary to our prediction. Other sensitivity analyses confirm our main finding. In fact, while we prefer standard regression analyses in this part of the empirical investigation, we can also estimate an ordered probit model to take the ordinal scale of our dependent variable into account. Again, we reach the same conclusion, independent of whether we look at ability as a continuous variable or whether we use dummy variables for different quantiles of ability. In regard to sample selection, we estimate a Heckman model in order to check whether the results are subject to a selection bias caused by the voluntary participation in the competency test but no evidence points towards such type of selectivity.[39] In further robustness checks, we exclude students with a foreign mother tongue who might have a disadvantage in the reading tests. To fully rule out that the findings are related to interview timing, we replace the control variables for interview month with more detailed information for the exact week of the interview. We also rerun specification 4 excluding the control variable for enjoyment of studying, which

[39] To employ an instrument, we can make use of information on interviewer contact attempts. This indicator reflects the interviewer effort needed to reach targeted students. Basically, the more interviewer contact attempts are necessary, the less likely the student is a test participant. While this fact allows for having a relevant instrument on the first stage of the two-step estimator, we have to assume that interviewer contacts are unrelated to subjectively assessed student effort.

could be considered as potentially endogenous. Finally, the rich data allows adding further control variables for school history, such as the subjects chosen by students for their school-exit-examinations, which does not change our finding either.

Table 3.2: Ability and (quantitative) student effort

	(1) log(Self-study) (term)	(2) log(Self-study) (holidays)	(3) log(Attend classes)	(4) log(Free-time)
Ability	-0.034^{*}	-0.090^{***}	0.011	0.011^{+}
	(0.01)	(0.03)	(0.01)	(0.01)
Observations	4431	4431	4431	4431
Adj. R^2	0.052	0.065	0.047	0.059

OLS estimations; NEPS weights used; robust standard errors in parentheses $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Full set of control variables is used in each column, as in column (4) of table 3.1. Additional dummy variables for the interview month are included in each column. Full results are reported in table B.5 in the appendix.

Next, we focus on the effect of ability on quantitative study efforts. Table 3.2 shows our findings for several potential indicators, for which we use the log of each time-use variable. As our main outcome variable of interest, we first find a negative effect of ability on self-study time. The finding is the same, although with an even stronger effect size for the self-study time between terms. Arguably, this quantitative effort indicator could be even more telling, taking into consideration that, during the holidays, students manage their time completely autonomously, deciding whether or not to put in extra effort for university. Using lecture attendance as a dependent variable in column 3, the estimation shows a slightly positive but statistically insignificant impact of ability. This is in line with the expectation that attendance does not reflect effort levels in the way that self-study does. Given the background of the German system of higher education with its focus on intense exam preparation at the end of the semester, one could interpret the results as many students attending classes for other reasons than improving academic achievement, such as socializing.[40] Finally, we regress our generated variable of students' free-time on ability. The result in column 4 shows that ability has a contrary effect in the sense that having higher ability implies enjoying more free-time if we accept a significance level of 10%. While this illustrates the main finding of our empirical investigation really well, the free-time measure certainly is somewhat noisy, as it indirectly considers the heterogeneity in students' lives and the fact that some students are influenced by factors such as work or children in addition to studying. Overall, however, we conclude from our additional sensitivity analyses that the main finding of leisure-enjoying high-ability types holds, which stands in contrast to our theoretical prediction.

[40]In additional analyses, we examine the link between the different time-use variables and academic success, measured in grades and study progress. While self-study is significantly related to our indicators of study success, lecture attendance seems to play no role.

3.3.2.2 Discussion

Our results reject the expectation of stronger efforts among high-ability compared to low-ability types of students. Rather it seems that a 'lazy genius' phenomenon exists of students with great prospects in their lives who put in relatively little effort during their studies. Apart from these interpretations, however, other aspects could (possibly) influence our empirical analysis. One of the concerns is linked to the understanding of ability as an exogenous factor that is unaffected by effort. While we argue that our ability measure based on competence testing is less affected by past learning efforts than alternative proxies, such as high school grades, these test results are still not immune to past choices on effort. Having said that, one could also suspect that those past efforts are related to current efforts, in consequence of which the empirical link between ability and effort could be biased upwards. As illustrated above, we consider the role of study effort in our ability measure to be negligible, given a clear lack of opportunity and incentives for any preparation of the test. Most importantly, such a potential bias cannot explain the result of a negative relationship between our ability measure and effort. If anything, our finding is even more striking, as we would underestimate the negative effect of ability on effort.

Another possible concern regarding the empirical procedure relates to our idea of analyzing effort based on subjective self-assessments, which might be susceptible to measurement issues. Concretely, the research on subjective data considers the possibility of a reference bias in self-reports (Groot, 2000). One could argue that high-ability students generally report lower effort levels, as their reference point is different. During the course of their studies, they cultivate different social contacts at the university and compare themselves to other students with high abilities rather than to those with low abilities. Yet, several aspects speak against such type of measurement problem.

First, we rely on data from the outset of the individuals' studies. In this phase, the freshmen are not segregated according to their ability yet, which reduces the likelihood of peer effects in self-reporting within their subjects.[41] Second, if peer effects led to reduced self-assessed effort levels, while the effort levels of the high-ability types were actually higher than those of the low-able, one would expect a narrowing of the gap between the two groups, not a complete reversal. Third, any effect of ability on self-reports is hampered by the fact that students are not necessarily fully informed about their actual ability

[41] A relevant phenomenon in this context is the orientation week at German schools of higher education where the freshmen get into social groups through a randomization procedure (Girard et al., 2015). For the first year or even longer, students typically stay together during lectures, when learning and in their non-university life. In consequence, not only work willingness but also the ability levels of their 'random' friends are typically very heterogeneous for German students, at least in the outset of their studies.

levels.[42] Fourth, the fact that we observe a fairly similar picture in both, self-assessed effort and the amount of time students spend on self-studying, conforms to our interpretation. In the case of the latter, we again observe that high-ability students report putting less effort into their studies in regard to the key factor self-study time, which is a variable that is arguably more objective. Fifth, while reference bias is seen as a relevant issue in the analysis of life satisfaction (see Odermatt and Stutzer, 2019), as happy people may have different standards of a happy life, we are not aware of any evidence supporting a similar idea for self-reported effort. If anything, individuals could possibly have a desire to over-report efforts due to image concerns (Ewers and Zimmermann, 2015), which could be a particular issue for individuals with high standards. If one is willing to assume that such form of approval-seeking in self-reported effort is more likely among high-ability types in high-ability environments, this again would imply that we underestimate our result regarding the 'lazy genius'. Lastly, we can check the role of varying contexts regarding high- vs low-ability peers empirically by considering differences across fields of study. While additional inspections of the data do not reveal any striking differences in variation of ability or effort across subjects, we are still cautious in considering subjects directly in our analysis, since endogenous self-selection related to our variable of interest could occur. Instead, we prefer using an alternative effort indicator as dependent variable for assessing the role of subject heterogeneity in our analysis. To do so, we define an adjusted effort variable as the individual deviation from the average effort in each field of study according to the available 2 digit ISCED classification.[43] While this alternative outcome variable considers differences in effort across subjects, the analysis of this subject-adjusted effort leads to the same finding regarding the negative impact of ability on effort. The results are shown in table B.2 in the appendix and rather indicate that the effects become stronger when we consider subject differences. As a further check, we re-run our analysis for subsamples based on subjects but again do not find any evidence for positive effects of ability on effort. We conclude from these considerations that the empirical findings are valid and do not result from giving special attention to subjective data in our investigation.

[42] See Gary-Bobo and Trannoy (2008) for a discussion of students with imperfect knowledge of their ability. Also see Stinebrickner and Stinebrickner (2014) who provide empirical evidence on how students misperceive their ability to perform well.

[43] The 2 digit ISCED 1997 classification distinguishes between 21 fields of study (ranging from "Teacher Training and education science" to "Environmental protection") and was promulgated by UNESCO in 1997.

3.3.3 Student effort and job market prospects

3.3.3.1 Main results

To empirically test our second theoretical prediction, we begin with standard regressions in the vein of the previous chapter. The determinant of student effort in this case is their job market prospects, for which we analyze subjective self-assessments. We include a dummy variable that distinguishes between students with excellent job market prospects and those who do not report such a positive outlook. For a comparison, we show estimation results based on the small sample used so far, which includes ability. We then show results based on a sample with more than 10,000 individuals, which we can use by not considering ability.

The main finding of table 3.3 is that subjective job prospects are positively linked to both effort dimensions, i.e. self-study time during term (columns 1 and 2) and self-assessed effort levels (columns 3 and 4), irrespective of the sample that we use. In contrast to our analysis of the impact of students' ability levels, this is in line with our theoretical considerations. However, while there is no reason to believe that students' current effort levels vice versa affect their ability levels (negatively), the potential problem of reverse causality is certainly an issue here. One may argue that it is not the great prospects that spur effort but that great prospects result from high effort levels.

Table 3.3: Job market prospects and student effort

	(1) Self-assessed effort	(2) Self-assessed effort	(3) log(Self-study) (term)	(4) log(Self-study) (term)
Very good job prospects	0.189***	0.136***	0.120***	0.099***
	(0.04)	(0.03)	(0.03)	(0.02)
Ability	−0.088***		−0.036*	
	(0.02)		(0.01)	
Observations	4409	10233	4409	10233
Adj. R^2	0.052	0.046	0.058	0.051

OLS estimations; NEPS weights used; robust standard errors in parentheses $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Full set of control variables is used in each column. Additional dummy variables for the interview month are included in each column. Full results are reported in table B.6 in the appendix.

To check the direction of the effect of perceived job market prospects on effort levels, we employ an IV approach for which we exploit labor market data. The idea is that variation in labor market conditions in the region where the university is located is a) effectively influencing perceptions of one's own future employment outlook and b) plausibly determined exogenously and thus not dependent on effort decisions of students at the university.[44] Current variations in local labor market conditions may not influence a student's actual employment situation in the future, but we expect an impact on his or her percep-

[44]For a similar IV approach using regional labor market data from Germany, see Reichert et al. (2015).

tion of it. Specifically, we focus on labor market dynamics by using data on the amount of newly unemployed persons, i.e. changes in regional labor market conditions, which occur after the start of our investigation period. This is important, as given local unemployment rates may be related to institutional quality or other relevant regional characteristics, while, arguably, the dynamic changes in labor market conditions are not. To implement our idea, we identify the relevant labor market segment of each student by exploiting available information on the most common career aspiration for each field of study in the NEPS data. These career aspirations are measured on a standard classification of occupations (KldB1988), which allows us to merge the NEPS data with employment statistics from Germany's Federal Employment Agency at the industry level. As these industry-specific statistics are available at the regional level (German federal states), we can merge the employment statistics with our NEPS data based on the occupation identifier and the federal state where the university is located. This procedure yields a large number of cells (= industry $\times$ state), for which we can attach information on varying labor market conditions to the NEPS data. As our IV, we use actual numbers of additional unemployed persons within each regional industry sector divided by overall employment numbers per cell. This weighted inflow from employment to unemployment varies between 0 and 20 percentage points with a mean of 1.3 and a median of 0.8 percentage points. In a last step, we multiply this inflow variable with 100 and take the square root to consider outliers. As an exogenous and unanticipated influence during their studies, we expect our instrument 'unemployment inflow' to decrease students' perceived job prospects after enrollment. Since most students were interviewed at the end of the first term and our particular interest lies in job prospects during the term, we use data on unemployment entries from October to December 2010 and relate those to employment data of the same labor market segment from September 2010. Additional robustness checks show that our results do not depend on choosing those time points.

In terms of methodology, we prefer a procedure proposed by Wooldridge (2002) for potentially endogenous binary variables over the standard linear probability model. Hence, in a first step, we estimate a probit model for very good job prospects using unemployment inflow as independent variable. In a second step, we include the predicted values as an instrument in standard two-stage least-squares regressions. Employing this procedure, statistical significance tests yield asymptotically valid results (Wooldridge, 2002).

The IV results in table 3.4 substantiate the idea that job market prospects have a positive effect on effort levels. This holds for both of our effort variables, the subjective assessment and the key time-use variable of self-study. As shown in column 1, the probit estimation confirms our expectations regarding the impact of our instrument on job

Table 3.4: Job market prospects and student effort (IV estimations)

	(1) Very good job prospects	(2) Self-assessed effort	(3) log(Self-study) (term)
Unemployment inflow	-0.628^{***}		
	(0.05)		
Very good job prospects		0.606^{***}	0.351^{***}
		(0.13)	(0.09)
Observations	10233	10233	10233

IV-2SLS Wooldridge procedure estimations. Probit estimations in column (1), second stage (IV) estimations in columns (2) and (3); NEPS weights used; robust standard errors in parentheses $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Full set of control variables is used in each column. Additional dummy variables for the interview month are included in each column. Full results are reported in table B.7 in the appendix.

prospects. The instrument has the expected sign and is statistically significant. Thanks to the exogenous nature of changes in regional labor conditions, we interpret the second-stage results presented in columns 2 and 3 as evidence for a direct impact of students' employment outlook on their behavior.

3.3.3.2 Discussion

In this subsection, we briefly discuss our main findings regarding the role of job market prospects in student effort, starting with a report on a few robustness checks. Our findings are qualitatively similar when running a standard linear IV model with clustering of standard errors at the cell level (results are shown in table B.3 in the appendix). The first stage regressions again show that our instrumental variable has a strong effect on job market expectations, which ensures sufficient instrumental power (the F statistic is 34.66). Given the ordinal nature of one of our two effort variables, we conduct another sensitivity check for self-assessed effort and confirm the finding of a significantly positive impact of great employment prospects by using bivariate ordered probit (see Sajaia, 2008).

In further checks, we repeat the analysis by using different definitions of the instrument. For instance, we apply the logarithm instead of square root to determine the IV, and we create an alternative instrument by exploiting changes in unemployment rates over time. Furthermore, we conduct robustness checks by restricting the sample in different ways. For instance, we exclude students from our sample who reported having earned 60 credit points or more, which may conflict with their assumed freshman status. We can further exploit the information on interview dates by excluding observations from students who were interviewed late during the NEPS fieldwork. These modifications do not change our results in a qualitative way.

When comparing OLS results from the analysis at the beginning of the section with our IV results, we observe larger effect sizes when making use of an instrumental approach. This suggests that the former underestimate the importance of job market prospects in student effort and that the actual effect is stronger when considering endogeneity. This might be explained by heterogeneity in responses to the question on job prospects. Some highly motivated students may be less concerned about their individual future than others, and they thus may underreport their prospects despite actually having very good job market perspectives.

3.3.4 Effect heterogeneity

The main results from the previous sections indicate that students with the perception of a great job market outlook do not lean back like high-ability types do, but instead provide high effort levels. To better understand these results, it is helpful to find out which groups of students drive our findings about the role of ability and job prospects in student effort. We suspect the group of male students to explain lacking efforts, given the gender gap in educational outcomes (Goldin et al., 2006). Note that we observe gender differences in effort throughout our analyses (see figure B.2 as well as tables B.4 to B.7 in the appendix). A second group of interest in our context studies STEM (Science, technology, engineering, and mathematics) subjects.[45] In these fields, job prospects are typically believed to be excellent for German students, so that the current labor market development may not hold significance for them.

To learn more about effect heterogeneity, we expand our analysis by considering interaction terms in regression models. We inspect subgroup differences in the effects of ability and job market prospects. Regarding the latter, we use the instrument of unemployment inflow in a reduced-form fashion to allow for a consistent interaction analysis. Since we investigate effect heterogeneity in joint specifications, we use the smaller sample of students who participated in the competency tests.

Regarding ability, the results in table 3.5 show significant negative effects in student effort across both genders (columns 1 and 2). The interaction term is not significant, though the effect appears to be somewhat stronger for males. However, adding the coefficients for ability and the interaction (female $\times$ ability) yields a significant effect for females only with respect to subjectively assessed effort, not for time spent on self-study. This suggests that in particular high-ability males provide lower efforts than their low-ability counterparts by putting in less time for studying. Given that STEM subjects are selected more often by male students than by female students, it comes as no surprise that we yield similar findings for STEM students (columns 3 and 4). A weakly significant

[45] We use the 2 digit ISCED classification to make the distinction between STEM and non-STEM subjects. Accordingly, STEM subjects are life science, physical science, mathematics and statistics, computing, and engineering.

Table 3.5: Analysis of effect heterogeneity using interaction variables

	(1) Self-assessed effort	(2) log(Self-study) (term)	(3) Self-assessed effort	(4) log(Self-study) (term)
Ability	-0.111^{***}	-0.055^{*}	-0.084^{***}	-0.027
	(0.03)	(0.02)	(0.02)	(0.02)
Unemployment inflow	-0.165^{**}	-0.082^{+}	-0.186^{***}	-0.072^{**}
	(0.06)	(0.05)	(0.03)	(0.03)
Female	0.219^{**}	0.094	0.245^{***}	0.143^{***}
	(0.08)	(0.06)	(0.04)	(0.03)
Female × Ability	0.039	0.034		
	(0.04)	(0.03)		
Female × Unemployment inflow	-0.005	0.012		
	(0.07)	(0.05)		
STEM			-0.172	-0.001
			(0.11)	(0.07)
STEM × Ability			-0.033	-0.049^{+}
			(0.04)	(0.03)
STEM × Unemployment inflow			0.294^{**}	0.151^{*}
			(0.10)	(0.07)
Observations	4371	4371	4371	4371
Adj. R^2	0.052	0.056	0.056	0.065

OLS estimations; NEPS weights used; robust standard errors in parentheses $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Full set of control variables is used in each column. Additional dummy variables for the interview month are included in each column.

interaction term in column 4 (STEM × ability) indicates that the negative ability effect in effort is driven by students of STEM subjects.

Regarding job market prospects, table 3.5 reveals significant subgroup differences for STEM and non-STEM students in the effects of unemployment inflow (columns 3 and 4). In fact, adding up the effects yields the result that STEM students are not affected at all by unemployment inflows. While there are no gender differences in the effects of job prospects on students effort (columns 1 and 2), it is clear that non-STEM students drive our findings in section 3.3.3. Possibly, concerns about labor market developments are more important for those students than for STEM students. The latter may feel confident about their own future career independent of current developments and do not adjust effort levels. Given that STEM students with high ability are providing relatively low effort levels, we conclude that even worsening labor market conditions are not capable of triggering higher efforts among those talented individuals.

3.4 Conclusion

The aim of our paper is to improve the understanding of students' decision-making concerning their willingness to provide effort. We build a simple model and show that both students' abilities and job market prospects are predicted to increase effort. Our empirical findings on the impact of labor market prospects on student behavior are in line with this prediction. Accordingly, having a positive outlook on future earnings positively influences current effort levels. From a policy perspective, this appears to be a convenient finding, even if happiness researchers are correct in arguing that individuals to some extent overestimate the utility of future income (see e.g. Easterlin, 2001, Frey et al., 2007). However, this finding implies, vice versa, that bad labor market prospects could reduce effort levels at universities. Our evidence thus conforms to the conclusions of other researchers, such as Aina et al. (2011), who consider weak job prospects as an explanation for long study durations at Italian universities. In a similar fashion, van der Klaauw and van Vuuren (2010) argue that low labor market returns to academic performance explain the phenomenon of lacking ambition among Dutch students. It appears that we can provide empirical support for the important role of the labor market in student behavior for another country.[46]

The empirical evidence on the role of ability in student effort decisions is, however, in contrast with our prediction. Nevertheless, we can use the general form of our model to explain the finding of the 'lazy genius'. Since both effort dimensions decline in ability, we can conclude that study time and learning intensity are complements. The theoretical reasoning then suggests that the IE (SE) of high-ability students is relatively strong (weak) compared to that of low-ability students.[47] This could be because the future marginal utility gain of putting in high effort is relatively low for high-ability students since their abilities per se ensure a high level of educational achievement and thus a high level of expected income after studying. It is then rational for those students to reduce efforts as this is associated with a marginal utility gain during academic studies (e.g. students can increase leisure), which does not come at the expense of a high reduction of future utility gains. As a result, the assumed Cobb-Douglas specification of the EPF seems appropriate when we empirically investigate the relationship between job market prospects and effort but fails with respect to the 'lazy genius'.

Besides this interpretation, it might also be the case that some assumptions of the model fail. For example, it could be simply false that individuals benefit from raising their future income levels. One may question previous studies in favor of the human-capital theory and instead argue that higher education is in itself not necessarily relevant for one's potential to perform well in the labor market.[48] However, from the individual student's perspective, even if there were no human capital effect for income levels, educational achievement should work as a signal to potential employers. In this context, Arcidiacono et al. (2010) argue that graduation helps reveal ability to the labor market and thereby affect earnings, according to which high-ability students benefit in particular

[46] Also see Kahn (2010) who provides empirical evidence for the US on how bad labor market conditions during graduation negatively affect students' labor market outcomes later on.

[47] For related discussions, see Card (1995) and Bandiera et al. (2015).

[48] See e.g. Bedard (2001) and Frazis (2002) for more skeptical views on the human capital argument.

from successful studying. This leads to the question whether students believe in the importance of such a signal or whether they doubt that their degrees are key to labor market success in the future. Phenomena like grade inflation could be relevant in explaining why students are reluctant with their effort if they expect that the informative power of their educational success is not very effective.[49] Note, however, that schools of higher education have particular interest in avoiding the reduction of their degrees' value, as pointed out by Ehlers and Schwager (2016). Hence, the finding of little effort among high-ability types remains intriguing, especially since we find that students with better job market prospects are indeed motivated to put greater efforts into studying.

[49]For discussions on the informative value of grades, see e.g. Grant (2007) as well as Chadi and de Pinto (2018).

Appendix B

Variable definition of time use variables

English Questionnaire:
"How many hours in a typical week during term time do you spend doing the following activities?"

- Attend classes (lectures, seminars, tutorials, internships, etc.)
- Self-study (e.g. preparing, reviewing for class, preparing presentations, specialist reading, revision courses, student learning groups, homework, papers, exam preparation)
- Other study-oriented activities (e.g. library work, office hours, travel time)
- Employment
- Household (cleaning, shopping, etc.)
- Child care

Additional information for each item: "Please enter a figure for each activity, rounded to the full hour. Mark 'no time expenditure/not applicable' if you do not spend any time doing that activity or the activity does not apply to you."

Variable definition:
If the activity is not applicable, we replace the value of the respective variable by zero. This affects only a few cases (with the exception of child care, which is not applicable for most students). In order to consider outliers, we use logarithmic values of each time use variable. Observations with a zero are manually set to zero after taking the logarithm.

Tables

Table B.1: Descriptive statistics

	mean	sd	min	max
Dependent variables and further time use data				
Self-assessed effort	3.584	1.010	1.0	5.0
Time: Self-study (term)	13.261	9.225	0.0	90.0
Time: Self-study (holi.)	12.420	14.859	0.0	99.0
Time: Attending	22.884	7.045	0.0	60.0
Time: Study-oriented	5.139	4.231	0.0	90.0
Time: Job	3.807	5.650	0.0	45.0
Time: Household	4.544	3.453	0.0	35.0
Time: Childcare	0.352	3.754	0.0	99.0
Socio-demographics				
Female	0.557	0.497	0.0	1.0
Age	21.519	2.290	18.0	38.9
Migration	0.076	0.266	0.0	1.0
Foreign citizenship	0.024	0.152	0.0	1.0
Foreign mother tongue	0.055	0.228	0.0	1.0
School years (father)	14.808	2.532	9.0	18.0
School years (mother)	14.370	2.434	9.0	18.0
No partner	0.451	0.498	0.0	1.0
Partner, living apart	0.421	0.494	0.0	1.0
Partner, living together	0.129	0.335	0.0	1.0
Life circumstances and school history				
Children in household	0.010	0.101	0.0	1.0
Single person household	0.281	0.449	0.0	1.0
Living with parents	0.238	0.426	0.0	1.0
Living in dorm	0.132	0.339	0.0	1.0
Living in rented flat	0.587	0.492	0.0	1.0
Living in own flat	0.010	0.102	0.0	1.0
Living in a sublet	0.033	0.179	0.0	1.0
Repeated high school year	0.112	0.315	0.0	1.0
Gymnasium	0.771	0.420	0.0	1.0
Nontraditional A levels	0.020	0.141	0.0	1.0
University and work related controls				
U of Applied Science	0.331	0.471	0.0	1.0
Teaching track	0.104	0.305	0.0	1.0
Change of subject	0.076	0.264	0.0	1.0
Enjoyment of studying	4.387	0.726	1.0	5.0
Region: North	0.133	0.340	0.0	1.0
Region: West	0.187	0.390	0.0	1.0
Region: South	0.402	0.490	0.0	1.0
Region: East	0.277	0.448	0.0	1.0
Working	0.478	0.500	0.0	1.0
Income	898.305	677.609	0.0	10870.0
Funding: Family	0.743	0.437	0.0	1.0
Funding: BAföG*	0.332	0.471	0.0	1.0
Funding: Bank loan	0.031	0.174	0.0	1.0
Funding: Earnings	0.564	0.496	0.0	1.0
Funding: Apprentice pay	0.050	0.218	0.0	1.0
Funding: Own resources	0.250	0.433	0.0	1.0
Funding: Gov. benefits	0.312	0.463	0.0	1.0
Funding: Scholarship	0.058	0.233	0.0	1.0
Funding: Other	0.014	0.117	0.0	1.0
Funding: Third parties	0.650	0.477	0.0	1.0
Variables of interest				
Job prospects**	4.081	0.799	1.0	5.0
Ability	0.000	1.000	-4.9	4.0
Observations	4431			

NEPS weights used.
* BAföG is the abbreviation for the Federal Training Assistance Act, which regulates grants and loans for students in Germany.
** Number of observations for job prospects is 4409 and is slightly lower than for the sample shown here (due to 22 missing values).

Table B.2: Study effort and ability using deviation from the mean

	(1) Adjusted self-assessed effort	(2) Adjusted self-study time (term)	(3) Adjusted self-study time (holidays)	(4) Adjusted time for attending classes	(5) Adjusted free-time
Ability	-0.095^{***}	-0.648^{***}	-1.216^{***}	-0.058	0.619^{+}
	(0.02)	(0.19)	(0.29)	(0.14)	(0.35)
Observations	4431	4431	4431	4431	4431
Adj. R^2	0.048	0.024	0.055	0.056	0.081

OLS estimations; NEPS weights used; robust standard errors in parentheses $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The dependent variables are the individual deviations from the mean in each field of study (2-digit ISCED classification). Full set of control variables is used in each column. Additional dummy variables for the interview month are included in each column.

Table B.3: Job market prospects and student effort (IV) - Robustness check

	(1) Very good job prospects 1st stage	(2) Self-assessed effort 2nd stage	(3) log(Self-study (term)) 2nd stage
Unemployment inflow	-0.180^{***}		
	(0.03)		
Very good job prospects		0.660^{***}	0.336^{*}
		(0.20)	(0.17)
Observations	10233	10233	10233
F statistic		34.660	34.660

Standard IV-2SLS estimations. First stage estimations in column (1), second stage estimations in columns (2) and (3); NEPS weights used; robust standard errors clustered at the cell level (industry $\times$ state) in parentheses $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Full set of control variables is used in each column. Additional dummy variables for the interview month are included in each column.

Full regression tables

Table B.4: Ability and (qualitative) student effort

	(1) Self-assessed effort	(2) Self-assessed effort	(3) Self-assessed effort	(4) Self-assessed effort
Ability	-0.075^{***}	-0.056^{**}	-0.053^{**}	-0.085^{***}
	(0.02)	(0.02)	(0.02)	(0.02)
Female		0.151^{***}	0.158^{***}	0.172^{***}
		(0.04)	(0.04)	(0.04)
Age		0.017^{*}	0.017^{+}	0.028^{**}
		(0.01)	(0.01)	(0.01)
Migration		-0.060	-0.070	-0.056
		(0.12)	(0.12)	(0.12)
Foreign citizenship		0.077	0.089	0.019
		(0.19)	(0.19)	(0.19)
Foreign mother tongue		0.000	0.018	0.053
		(0.14)	(0.14)	(0.14)
School years (father)		0.010	0.012	0.008
		(0.01)	(0.01)	(0.01)
School years (mother)		-0.018^{+}	-0.017^{+}	-0.020^{*}
		(0.01)	(0.01)	(0.01)
Partner, living apart		0.170^{***}	0.174^{***}	0.162^{***}
		(0.04)	(0.04)	(0.04)
Partner, living together		0.135^{*}	0.202^{***}	0.199^{***}
		(0.06)	(0.06)	(0.06)
Children in household			-0.127	-0.116
			(0.27)	(0.28)
Single person household			0.093^{*}	0.082^{+}
			(0.04)	(0.04)
Living with parents			0.038	0.072
			(0.11)	(0.11)
Living in dorm			0.035	0.064
			(0.11)	(0.11)
Living in rented flat			-0.091	-0.049
			(0.11)	(0.10)
Living in own flat			-0.075	-0.078
			(0.22)	(0.21)
Repeated high school year			-0.104^{+}	-0.099^{+}
			(0.06)	(0.06)
Gymnasium			-0.024	-0.069
			(0.05)	(0.05)
Nontraditional A levels			0.217^{+}	0.178
			(0.13)	(0.12)
U of Applied Science				-0.179^{***}
				(0.05)
Teaching track				-0.036
				(0.04)
Change of subject				0.009
				(0.06)
Enjoyment of studying				0.129^{***}
				(0.03)
Region: West				0.038
				(0.06)
Region: South				-0.009
				(0.06)
Region: East				-0.095
				(0.06)
Working				-0.036
				(0.04)
Log-Income				-0.005
				(0.03)
Source of Funding				✓
Interview month		✓	✓	✓
Observations	4431	4431	4431	4431
Adj. R^2	0.005	0.019	0.023	0.046

OLS estimations; NEPS weights used; robust standard errors in parentheses $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table B.5: Ability and (quantitative) student effort

	(1) log(Self-study) (term)	(2) log(Self-study) (holidays)	(3) log(Attend classes)	(4) log(Free-time)
Ability	-0.034^{*}	-0.090^{***}	0.011	0.011^{+}
	(0.01)	(0.03)	(0.01)	(0.01)
Female	0.090^{***}	-0.020	-0.013	-0.020^{***}
	(0.03)	(0.05)	(0.01)	(0.01)
Age	0.015^{*}	0.034^{**}	-0.008^{*}	-0.004^{*}
	(0.01)	(0.01)	(0.00)	(0.00)
Migration	-0.176^{*}	-0.002	-0.053	0.011
	(0.07)	(0.16)	(0.04)	(0.02)
Foreign citizenship	0.019	-0.081	-0.028	0.017
	(0.11)	(0.24)	(0.07)	(0.04)
Foreign mother tongue	0.178^{*}	-0.020	0.124^{**}	-0.064^{*}
	(0.09)	(0.19)	(0.05)	(0.03)
School years (father)	0.013^{*}	0.013	0.001	-0.003
	(0.01)	(0.01)	(0.00)	(0.00)
School years (mother)	0.007	0.011	0.001	0.000
	(0.01)	(0.01)	(0.00)	(0.00)
Partner, living apart	-0.033	-0.046	0.015	0.004
	(0.03)	(0.05)	(0.01)	(0.01)
Partner, living together	0.007	-0.158^{*}	0.005	-0.017^{+}
	(0.04)	(0.08)	(0.02)	(0.01)
Children in household	-0.309	0.132	0.107^{*}	-0.287^{***}
	(0.22)	(0.26)	(0.05)	(0.08)
Single person household	0.030	0.094	0.007	-0.005
	(0.03)	(0.06)	(0.01)	(0.01)
Living with parents	-0.047	0.091	-0.020	-0.005
	(0.08)	(0.16)	(0.03)	(0.02)
Living in dorm	-0.006	-0.049	0.017	0.012
	(0.08)	(0.16)	(0.03)	(0.02)
Living in rented flat	-0.037	-0.067	-0.031	0.015
	(0.08)	(0.15)	(0.03)	(0.02)
Living in own flat	-0.041	0.566^{*}	-0.022	-0.026
	(0.17)	(0.24)	(0.05)	(0.03)
Repeated high school year	-0.012	-0.061	-0.018	0.016
	(0.04)	(0.08)	(0.02)	(0.01)
Gymnasium	-0.055	-0.080	-0.029^{+}	0.020^{+}
	(0.04)	(0.07)	(0.02)	(0.01)
Nontraditional A levels	0.016	0.066	0.005	0.015
	(0.08)	(0.18)	(0.04)	(0.02)
U of Applied Science	-0.304^{***}	-0.518^{***}	0.078^{***}	0.025^{**}
	(0.04)	(0.07)	(0.02)	(0.01)
Teaching track	-0.017	-0.114^{*}	0.045^{***}	-0.002
	(0.03)	(0.05)	(0.01)	(0.01)
Change of subject	-0.031	-0.087	-0.002	0.016
	(0.04)	(0.08)	(0.02)	(0.01)
Enjoyment of studying	0.034^{+}	0.095^{**}	0.012	-0.002
	(0.02)	(0.03)	(0.01)	(0.00)
Region: West	0.018	0.016	-0.039^{+}	-0.007
	(0.04)	(0.09)	(0.02)	(0.01)
Region: South	-0.075^{+}	-0.314^{***}	0.008	0.013
	(0.04)	(0.08)	(0.02)	(0.01)
Region: East	-0.050	-0.316^{***}	-0.007	-0.012
	(0.04)	(0.08)	(0.02)	(0.01)
Working	-0.056^{*}	-0.055	-0.009	-0.011^{*}
	(0.03)	(0.05)	(0.01)	(0.01)
Log-Income	-0.029	-0.028	-0.002	-0.003
	(0.02)	(0.04)	(0.01)	(0.00)
Source of Funding	✓	✓	✓	✓
Interview month	✓	✓	✓	✓
Observations	4431	4431	4431	4431
Adj. R^2	0.052	0.065	0.047	0.059

OLS estimations; NEPS weights used; robust standard errors in parentheses $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table B.6: Job market prospects and student effort

	(1) Self-assessed effort	(2) Self-assessed effort	(3) log(Self-study) (term)	(4) log(Self-study) (term)
Very good job prospects	0.189^{***}	0.136^{***}	0.120^{***}	0.099^{***}
	(0.04)	(0.03)	(0.03)	(0.02)
Ability	-0.088^{***}		-0.036^{*}	
	(0.02)		(0.01)	
Female	0.216^{***}	0.252^{***}	0.118^{***}	0.085^{***}
	(0.04)	(0.02)	(0.03)	(0.02)
Age	0.029^{**}	0.020^{***}	0.015^{*}	0.021^{***}
	(0.01)	(0.01)	(0.01)	(0.00)
Migration	-0.064	0.008	-0.184^{*}	0.001
	(0.12)	(0.07)	(0.07)	(0.05)
Foreign citizenship	0.039	-0.099	0.030	-0.154^{+}
	(0.19)	(0.11)	(0.11)	(0.09)
Foreign mother tongue	0.052	0.082	0.181^{*}	0.016
	(0.15)	(0.09)	(0.09)	(0.07)
School years (father)	0.007	-0.000	0.013^{*}	0.003
	(0.01)	(0.01)	(0.01)	(0.00)
School years (mother)	-0.020^{*}	-0.014^{*}	0.007	0.008^{+}
	(0.01)	(0.01)	(0.01)	(0.00)
Partner, living apart	0.162^{***}	0.131^{***}	-0.034	-0.033^{+}
	(0.04)	(0.03)	(0.03)	(0.02)
Partner, living together	0.198^{***}	0.200^{***}	0.007	-0.040
	(0.06)	(0.04)	(0.04)	(0.03)
Children in household	-0.093	-0.108	-0.297	-0.173^{*}
	(0.29)	(0.12)	(0.23)	(0.08)
Single person household	0.086^{+}	0.065^{*}	0.036	0.044^{+}
	(0.04)	(0.03)	(0.03)	(0.02)
Living with parents	0.077	-0.012	-0.047	-0.028
	(0.11)	(0.08)	(0.08)	(0.06)
Living in dorm	0.069	0.045	-0.004	0.043
	(0.11)	(0.08)	(0.08)	(0.06)
Living in rented flat	-0.042	-0.079	-0.036	-0.019
	(0.10)	(0.08)	(0.08)	(0.06)
Living in own flat	-0.107	-0.172	-0.063	0.070
	(0.21)	(0.11)	(0.17)	(0.09)
Repeated high school year	-0.084	-0.097^{**}	-0.003	-0.013
	(0.06)	(0.04)	(0.04)	(0.03)
Gymnasium	-0.066	-0.053	-0.054	-0.060^{*}
	(0.05)	(0.03)	(0.04)	(0.03)
Nontraditional A levels	0.179	0.157^{*}	0.018	0.051
	(0.12)	(0.07)	(0.09)	(0.05)
U of Applied Science	-0.183^{***}	-0.078^{*}	-0.307^{***}	-0.264^{***}
	(0.05)	(0.03)	(0.04)	(0.02)
Teaching track	-0.032	-0.002	-0.016	-0.003
	(0.04)	(0.02)	(0.03)	(0.02)
Change of subject	0.016	0.015	-0.022	0.010
	(0.06)	(0.04)	(0.04)	(0.03)
Enjoyment of studying	0.123^{***}	0.117^{***}	0.032^{+}	0.051^{***}
	(0.03)	(0.02)	(0.02)	(0.01)
Region: West	0.058	0.012	0.034	-0.011
	(0.06)	(0.04)	(0.04)	(0.03)
Region: South	0.013	-0.022	-0.063	-0.079^{**}
	(0.06)	(0.04)	(0.04)	(0.03)
Region: East	-0.078	-0.121^{**}	-0.041	-0.063^{*}
	(0.06)	(0.04)	(0.04)	(0.03)
Working	-0.033	-0.031	-0.053^{+}	-0.024
	(0.04)	(0.03)	(0.03)	(0.02)
Log-Income	-0.006	-0.009	-0.030	0.007
	(0.03)	(0.02)	(0.02)	(0.02)
Source of Funding	✓	✓	✓	✓
Interview month	✓	✓	✓	✓
Observations	4409	10233	4409	10233
Adj. R^2	0.052	0.046	0.058	0.051

OLS estimations; NEPS weights used; robust standard errors in parentheses $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table B.7: Job market prospects and student effort (IV estimations)

	(1) Very good job prospects	(2) Self-assessed effort	(3) log(Self-study (term))
Unemployment inflow	-0.628^{***} (0.05)		
Very good job prospects		0.606^{***} (0.13)	0.351^{***} (0.09)
Female	-0.526^{***} (0.03)	0.355^{***} (0.04)	0.140^{***} (0.03)
Age	-0.007 (0.01)	0.022^{***} (0.01)	0.022^{***} (0.00)
Migration	0.023 (0.10)	0.001 (0.07)	-0.003 (0.05)
Foreign citizenship	-0.218 (0.15)	-0.069 (0.12)	-0.138 (0.09)
Foreign mother tongue	-0.047 (0.12)	0.089 (0.09)	0.020 (0.07)
School years (father)	0.011 (0.01)	-0.002 (0.01)	0.003 (0.00)
School years (mother)	0.001 (0.01)	-0.014^{*} (0.01)	0.008^{+} (0.00)
Partner, living apart	0.027 (0.04)	0.124^{***} (0.03)	-0.037^{+} (0.02)
Partner, living together	0.025 (0.06)	0.191^{***} (0.04)	-0.044 (0.03)
Children in household	0.206 (0.14)	-0.138 (0.12)	-0.189^{*} (0.08)
Single person household	-0.055 (0.04)	0.070^{*} (0.03)	0.047^{*} (0.02)
Living with parents	-0.032 (0.10)	-0.017 (0.08)	-0.031 (0.06)
Living in dorm	-0.016 (0.10)	0.034 (0.08)	0.037 (0.06)
Living in rented flat	-0.059 (0.10)	-0.079 (0.08)	-0.019 (0.06)
Living in own flat	0.159 (0.15)	-0.203^{+} (0.12)	0.054 (0.09)
Repeated high school year	-0.068 (0.05)	-0.082^{*} (0.04)	-0.006 (0.03)
Gymnasium	0.027 (0.05)	-0.054 (0.03)	-0.060^{*} (0.03)
Nontraditional A levels	0.024 (0.10)	0.161^{*} (0.07)	0.053 (0.05)
U of Applied Science	0.016 (0.04)	-0.084^{**} (0.03)	-0.267^{***} (0.02)
Teaching track	-0.416^{***} (0.04)	0.012 (0.03)	0.004 (0.02)
Change of subject	-0.051 (0.06)	0.026 (0.04)	0.016 (0.03)
Enjoyment of studying	0.114^{***} (0.02)	0.101^{***} (0.02)	0.043^{***} (0.01)
Region: West	-0.228^{***} (0.06)	0.046 (0.04)	0.007 (0.03)
Region: South	-0.316^{***} (0.05)	0.009 (0.04)	-0.062^{*} (0.03)
Region: East	-0.051 (0.06)	-0.091^{*} (0.04)	-0.046 (0.03)
Working	0.012 (0.04)	-0.032 (0.03)	-0.024 (0.02)
Log-Income	0.038 (0.02)	-0.015 (0.02)	0.004 (0.02)
Source of Funding	✓	✓	✓
Interview month	✓	✓	✓
Observations	10233	10233	10233

IV-2SLS Wooldridge procedure estimations. Probit estimations in column (1), second stage (IV) estimations in columns (2) and (3); NEPS weights used; robust standard errors in parentheses $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Figures

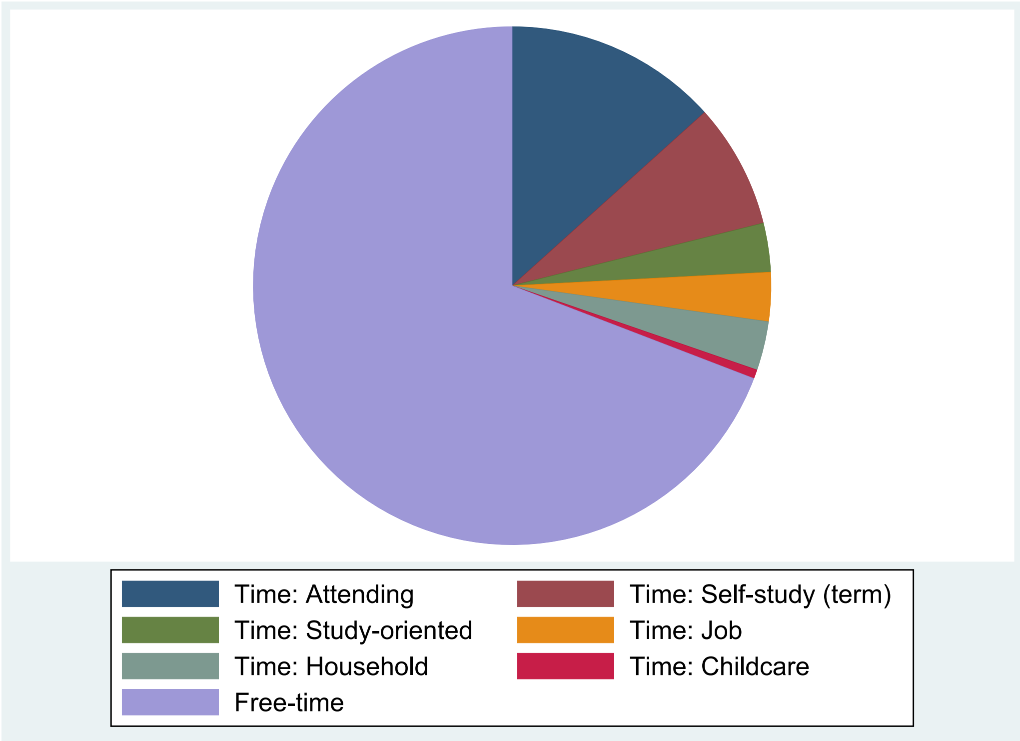

Figure B.1: Students' time allocation
Source: NEPS. Note that free-time includes sleep and weekends.

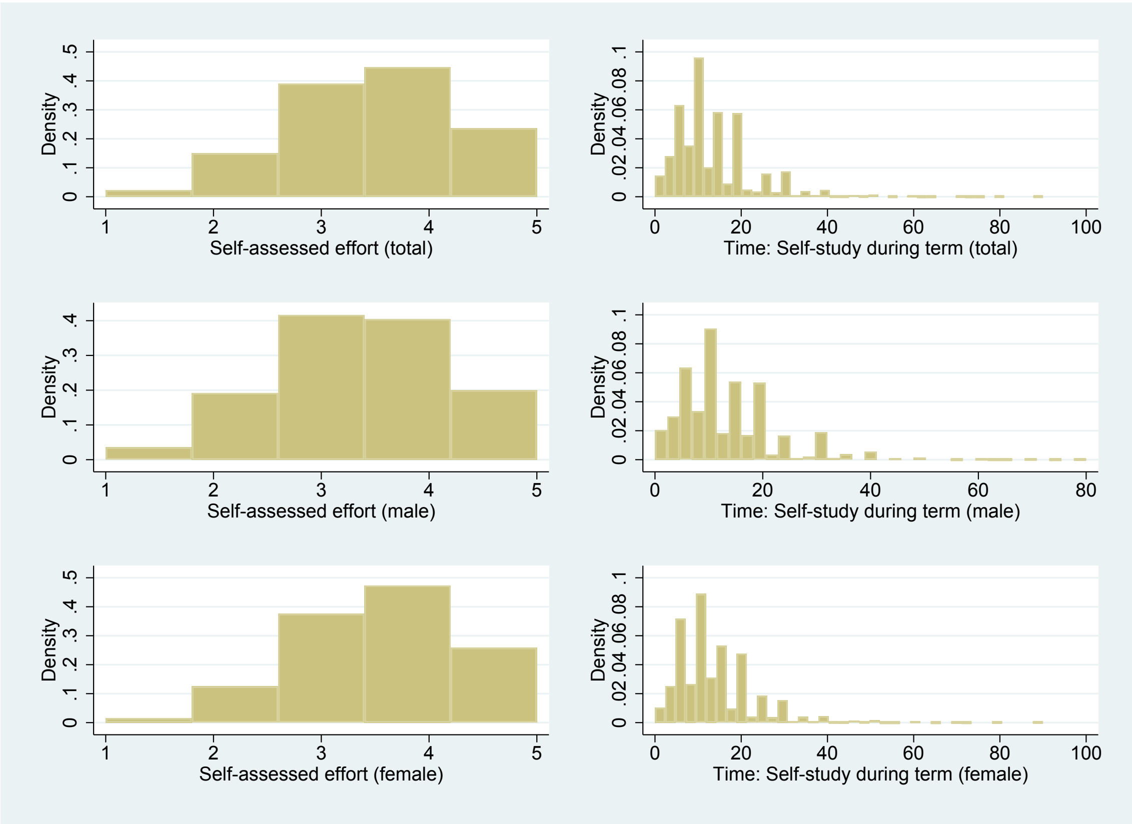

Figure B.2: Histograms of effort variables
Source: NEPS.

Part II

Industrial Relations

Chapter 4

Trade Union Membership and Overtime in Germany

Trade unions can support members in conflicts with employers and help them to reduce overtime. However, union members may also have easier access to (paid) overtime than non-members. Hence, the impact of individual trade union membership on overtime is theoretically ambiguous. We empirically investigate the relationship for Germany. Based on SOEP data and using panel data models, we observe for a multitude of specifications that membership reduces overtime in the private sector. This effect appears to arise because union membership can protect employees from according working time demands by employers.[50]

Keywords: overtime, working time, trade union membership, German Socio-Economic Panel

JEL Classification: J2, J51

[50]This chapter is joint work with Laszlo Goerke.
We are grateful for helpful comments by participants of a seminar in Trier, the ZiF Workshop 2018 in Bielefeld, the 13th International German Socio-Economic Panel User Conference in Berlin, the Autumn Workshop WK Personal in Munich, the Labour Market Workshop 2018 in Luxembourg, the 11th Interdisciplinary Ph.D. Workshop at the IAB in Nuremberg, the 22nd Colloquium on Personnel Economics in Augsburg, the 11th Workshop on Labour Economics in Trier, the RES Symposium of Junior Researchers in Warwick, the 33rd Annual Conference of the European Society for Population Economics in Bath, the 34th Conference of the Italian Association of Labour Economists in Novara, the Annual Conference 2019 of the German Economic Association in Leipzig and of a research seminar in Hamburg. We also thank John Addison, Adrian Chadi, Wolter Hassink, Boris Hirsch, Markus Pannenberg and all colleagues from the IAAEU for their useful comments.

4.1 Introduction

Long working hours and overtime are associated with numerous adverse outcomes, such as a deterioration in health (Berniell and Bietenbeck 2020, Cygan-Rehm and Wunder 2018), industrial accidents (Lee and Lee 2016), lower well-being (Lepinteur 2019) and a reduction in job quality (European Parliament 2009, OECD 2014, ch.3). Accordingly, the determinants of overtime have been looked at extensively. Nonetheless, one essential aspect for the evaluation of overtime has found relatively little attention: Can employees actually enforce their working time preferences and ward off demands or overcome corresponding restrictions by their employers?

The answer to this question depends crucially on the employees' bargaining power. It is likely to be higher if they belong to a trade union because membership can make it more costly for the firm to ignore an employee's preferences. Membership provides individuals with information, legal advice and can guarantee representation by experts in conflicts with the firm. If employers want their workforce to work more hours, being a member of a trade union can help to fend off such demands. Hence, trade union membership can be expected to reduce overtime. Alternatively, it can be hypothesised that employees like to work more hours, to enhance their income. If union members have better access to (paid) overtime because of their bargaining power, a positive relationship between individual union membership and overtime can be predicted. Hence, the direction of the effect is theoretically ambiguous.

In this paper, we address this issue by empirically analysing the relationship between individual trade union membership and overtime in Germany. To illustrate the relevance, note that weekly contractual working hours in Germany generally range from 35 to 40 hours for full-time employees. Overtime, both paid and unpaid, is widespread (Pannenberg 2005). According to the Institute for Employment Research (IAB 2020), in 2019 each employee worked on average 45 hours overtime per year, of which 50% were unpaid. More than 50% of full-time employees worked two or more hours of overtime per week according to the Working Time Report Germany (BAuA 2016). Hence, the large number of overtime hours has been a recurring topic of the public debate. The German Trade Union Federation (DGB) has even gone so far as to criticise the large share of unpaid hours as "a scandal" (DGB 2018).

While the subsequent investigation does not constitute the first analysis of the association between unionisation and overtime, previous contributions have often considered collective bargaining coverage and not an individual's membership. This may be equivalent in some countries, in particular, if there is a closed-shop system. In Germany, though, individual trade union membership is not tightly related to collective bargaining. The German economy is an interesting case to consider also for a variety of further reasons: First, about half of all employees are overemployed, in that they would like to work fewer hours per week than they currently do, taking into account resulting income adjustments (Bell et al. 2012, Knaus and Otterbach 2019). For full-time employees, the share of overemployed is even larger. Second, a sizeable minority of employees belong to a trade union. The fraction of trade union members has been highest just after German unification in 1990 and since then declined substantially. In 2018, net union density was around 16.5% (Visser 2019). Thus, individual union membership can potentially have a significant effect. Finally, studies on other countries have obtained diverse results on the relationship between individual union membership and overtime. Hence, the linkage may depend on the industrial relations system. Germany is characterised by a relatively high collective bargaining coverage and the widespread existence of works councils. These institutions may affect the impact of individual membership on overtime.

Using the nine waves of the German Socio-Economic Panel (SOEP) for the period from 1985 to 2015 that contain information on union membership, we find that about half of full-time private-sector employees worked overtime in the preceding month. On average each employee worked 2.6 hours of overtime per week. When considering the overall effect, members of a trade union worked about 13 minutes per week less overtime than non-members, which corresponds to a reduction of more than 8% (evaluated at the sample mean). Findings from pooled Tobit estimates are corroborated by Tobit models for panel data. Moreover, comparable estimates can be obtained for a variety of subgroups, for paid and unpaid overtime, for a subsample of employees who have not changed employers, for those working in firms with a works council and those likely to be covered by a collective bargaining agreement. In contrast to the robust negative correlation in the private sector, we find no such relationship in the public sector.

Our findings have important implications: First, we establish a further potential determinant of overtime. Second, we detect an individual gain from union membership, namely differential overtime, which has hitherto gone unnoticed. Third, since union membership has declined over the last decades in Germany, our findings can assist in explaining the observed development of overtime. Fourth, greater flexibility in working time arrangements, for example, due to digitisation, automation and also better opportunities to work from home, suggest a more extensive use of overtime in the future. Our findings indicate the relevance of trade union membership in such a more individualised work environment.

The remainder of the paper is structured as follows: In Section 4.2, we survey earlier contributions. Subsequently, in Section 4.3 we develop a number of channels which can rationalise how trade union membership impacts on overtime. Section 4.4 describes the data and outlines the empirical strategy. We present results in Section 4.5, where we also shed light on the potential channels by which membership can affect overtime outcomes. The last section contains some concluding remarks.

4.2 Previous contributions

In this section we, first, describe earlier contributions on the relationship between trade unions and overtime. Second, we survey findings on the determinants of overtime in Germany, based on longitudinal household data.

Studies focusing on collective bargaining or the presence of a trade union, in general, find that covered employees are less likely to work overtime at all (see Trejo (1993) for the US, Scheuer (1999) for Denmark, and Jirjahn (2008) for West Germany[51]). The results for overtime hours are more mixed. Some analyses show that covered employees work fewer overtime hours (Trejo (1993) for the US, Bell and Hart (1999) for British males and Veliziotis (2013), also for Britain, and both relating to unpaid hours, as well as Jirjahn (2008) for blue-collar workers in West German manufacturing establishments). This is consistent with the evidence by the OECD (2018, p.100f) that any form of employee representation is associated with a lower likelihood of working excessive hours. In contrast, Chiang (2012) finds a positive correlation between union coverage and overtime hours in Japan, and Miller and Mulvey (1991) indicate that the same may be true for men in Australia, focusing on what they label 'compulsory unionism'. Bell and Hart (1999) also document a positive association between union status and paid overtime hours, using the UK Labour Force Survey,[52] while Kalwij and Gregory (2005) consider union coverage and observe no correlation. Finally, Bell and Hart (2019) analyse the Annual Survey of Hours and Earnings and document a higher incidence and more overtime hours for individuals covered by collective bargaining agreements in Britain. In sum, the relationship between collective bargaining and overtime varies across countries and industrial relations systems and depends on the indicator of overtime.

Turning to individual union membership, empirical evidence is much scarcer. Miller and Mulvey (1991) use the Australian Longitudinal Survey for 1985. They show that male union members are more likely to work overtime, while the number of overtime hours is unaffected by union status. In addition, there are contributions for Britain. Pannenberg

[51]The estimated coefficient of the dummy indicating a collective bargaining agreement becomes insignificant in Jirjahn's (2008) analysis, once contractual hours are included as control variable. Using National Compensation Survey data for the United States, Barkume (2010) finds no linkage between a job being unionised and the incidence of overtime either. Based on the New Earnings Survey, Kalwij and Gregory (2005) obtain a similar result for women and observe a positive effect of union coverage for men in Britain.

[52]The question Bell and Hart (1999) use refers to individual union membership. They interpret the response as indication of union coverage. Bell and Hart (1998) also present evidence that collective bargaining coverage raises overtime hours, based on the New Earnings Survey.

and Wagner (2001) employ data for 1991-1998 from the British Household Panel Survey (BHPS). They report that the likelihood of working paid overtime is higher for a member of a trade union, while the respective probability for unpaid overtime is negatively correlated with union membership.[53] This is consistent with more recent evidence by Hart and Ma (2010) for a sample of male employees drawn from 15 waves of the BHPS. Veliziotis (2013) only considers unpaid overtime and distinguishes between the for-profit and the caring sector, relying on the BHPS, as well. While pooled estimates reveal differences between sectors, the union membership variable is no longer significantly different from zero in correlated random-effects Tobit and linear fixed-effects specifications. Turning to Germany, Zapf (2015) employs SOEP data for 2011, since this wave contains information on working conditions, which is the focus of her research. She finds the probability of working unpaid overtime to be lower for full-time employees who belong to a trade union, while there is no correlation with paid overtime. To summarise: The empirical evidence on the relationship between individual trade union membership and overtime is inconclusive.

Empirical analyses of the determinants of overtime in Germany focusing on individual choices often employ data from the SOEP (see Hübler 1989, Bauer and Zimmermann 1999, Pannenberg and Wagner 2001, and Bell et al. 2001). While there are no studies on overtime behaviour of union members, a number of analyses have looked at works councils. Using two SOEP waves, Kraft and Lang (2008) show that respondents work fewer hours of overtime if they are employed in an establishment in which a works council exists. Moreover, employing an additional wave, Gralla et al. (2017) obtain evidence of a positive correlation in the private sector, if employees have a contractual working time of 35 hours/week. The correlation is generally negative for employees with 40 contractual hours/week. Finally, Zapf (2015) shows that in 2011 the incidence of paid and of unpaid overtime is lower in works council establishments.[54] Consequently, the direction of the relationship between overtime hours and the existence of a works council is still in dispute.

[53] Pannenberg and Wagner (2001) state – but do not document – that union members in Germany are less likely to work unpaid overtime, utilising two SOEP (1989, 1998) waves.

[54] Based on a linked employer-employee data set (SOEP-LEE) for 2011, Zapf and Weber (2017) confirm some of the findings by Zapf (2015). Their specifications not directly comparable to those based on the SOEP since they contain substantially different sets of independent variables and do not include a works council dummy. Schank and Schnabel (2004) base their analysis on data from the IAB establishment panel. Overtime is sometimes related positively to the existence of a works council in pooled samples but no such relationship can be observed in fixed-effects specifications. Jirjahn (2008) observes no correlation between overtime and works councils.

4.3 Impact of individual union membership on overtime

In Germany, currently somewhat more than one half of all employees work in firms which are covered by collective bargaining agreements (i.e. unionised in the US terminology). This percentage has declined substantially in recent decades. Moreover, about 50% of those employees whose working conditions are not regulated collectively are actually paid in line with such agreements because firms voluntarily follow its regulations (Addison et al. 2016, Ellguth and Kohaut 2020). In addition to wages, standard working hours and overtime regulations are generally negotiated collectively. While only signatories are legally bound by the content of collective bargaining agreements, firms usually apply them to all employees, irrespective of whether they are members of a trade union or not. Moreover, employment in a covered establishment and the application of a collective bargaining agreement cannot be conditioned on an individual's union membership. Accordingly, there is no evidence of a wage difference between union members and non-members (Schmidt and Zimmermann 1991, Fitzenberger et al. 1999, Goerke and Pannenberg 2004).

This feature has an impact on the relationship between union membership and overtime. Although wages are an important determinant of overtime (Hart 2004), in Germany overtime hours will not vary with membership on account of differential wage returns to working (paid) overtime. Therefore, we emphasise other aspects of the employment relationship which may provide union members and non-members with differential incentives and opportunities to work overtime.

First, about half of employees are overemployed, in that they would like to work fewer hours than they currently do, taking into account income adjustments (Bell et al. 2012, Knaus and Otterbach 2019). Additionally, there is some evidence that many employees work overtime either because the employer requests them to do so or since it is necessitated by the workload (Hart 2004, p.91, BAuA 2016). This suggests that employees cannot fully enforce their working time preferences. Being a member of a trade union can help to achieve the desired working time. This is because trade unions can inform members of their legal entitlements and help to enforce them. Moreover, unions can prevent disadvantageous treatments of their members who do not want to work the number of hours requested by the employer and protect them from possibly ensuing consequences, such as dismissals or incorrect wage payments. Therefore, the informational and protective role of trade unions, resulting in greater bargaining power vis-a-vis the employer, suggests that members work less overtime. This effect can be expected to be particularly pronounced for individuals who regard themselves as overemployed.

Second, union members may not only be able to realise given working-time preferences better, but may also be characterised by different preferences or work attitudes than non-members. If, for example, individuals identify with their work and attach high importance to it, they may also be more inclined to participate in the production of public goods provided by unions, such as higher wages for members and non-members alike and more favourable working conditions. If this identification with work extends to the firm, union members may be more willing to work overtime. However, if the identification focuses more strongly on employees, while the firm is viewed as an adversary, then a negative effect on overtime can be expected. Therefore, the direction of the impact of union membership on overtime on account of such a preference- or attitude-based linkage is a priori ambiguous.

Third, overtime hours can be used to increase income. In Germany, the magnitude of the overtime premium is often determined in collective bargaining agreements and usually amounts to 25% to 50%. Therefore, paid overtime has an immediate income effect. Union members will work more paid overtime if membership helps them to obtain such extra and better-paid work more often than non-members. Consequently, union membership can represent a means to earn additional income, despite the absence of a membership wage premium. Moreover, models of collective bargaining indicate that if wages and standard hours of work are negotiated, working time is less than the amount that would be chosen by individual employees (Oswald and Walker 1994, Fuest and Huber 2000). Hence, employees would be willing to work extra hours. Therefore, an income-enhancing role of trade unions suggests that members work more paid overtime.

Fourth, overtime can also increase income in the longer run. Overtime and, in particular, unpaid overtime may be an investment by employees to signal an imperfectly observable effort to the employer (Pannenberg 2005, Anger 2008). The returns to such investments are reaped in the future through a steeper wage profile or a greater probability of promotion. It can be argued that union members work more unpaid overtime since they have greater chances of earning the returns to such investments. This effect occurs because they have longer tenure and are less likely to be dismissed individually than non-members (Goerke and Pannenberg 2011, Berglund and Furåker 2016). Therefore, union membership may constitute a kind of insurance to guarantee the returns to an investment in the form of (unpaid) overtime.

Fifth, union members may choose to work in firms which require different amounts of overtime than other employers. Hence, any observed correlation between individual trade union membership and overtime could represent a selection effect.

Sixth, there is some evidence (see above) that overtime varies with collective bargaining status. As collective bargaining is more likely to exist, the stronger trade unions are, and because unions are likely to be powerful, the more members they have, a union membership-overtime nexus may also arise because of the correlation with collective bargaining coverage. The direction of the collective bargaining effect is ambiguous though because the impact of collective negotiations on overtime is not well determined (Jirjahn 2008).

Finally, labour relations in Germany are strongly influenced by codetermination. In 2019, about one in ten of eligible private-sector establishments had a works council. Since the likelihood of codetermination increases with firm size, 40% of private sector employees are represented by a works council (Ellguth and Kohaut 2020). The Works Constitution Act (WCA) establishes consultation and codetermination entitlements for a variety of aspects. One of the most important is working time and overtime regulations (§87(1) WCA). More specifically, overtime work arrangements have to be agreed upon between the employer and the works council. Such an arrangement may also include rules relating to the selection of employees who work overtime. Since works councillors are in their majority union members (Goerke and Pannenberg 2022, Behrens 2009b) and help trade unions to expand their membership at the plant level (Behrens 2009a), union membership is likely to be higher in codetermined plants. Therefore, we can expect a relationship between individual union membership and overtime on account of the correlation with codetermination at the establishment level. The direction of this works council effect is, once again, a priori uncertain. This is because the empirical evidence on the impact of works councils on overtime is not clear-cut (see Schank and Schnabel 2004, Jirjahn 2008, Kraft and Lang 2008, Zapf 2015, Gralla et al. 2017, Zapf and Weber 2017).

To summarise: We have identified various channels via which individual trade union membership can affect overtime or give rise to an empirically observable correlation. A priori considerations do not unambiguously predict the direction of the relationship.

4.4 Data and empirical strategy

4.4.1 Data

For the empirical investigation we use data from the German Socio-Economic Panel (SOEP), a representative panel study of German households conducted annually since 1984 (see SOEP v33 and Goebel et al. 2019). Our main variables of interest are derived from the responses to the following question, which has been asked with slight modifications in each wave since 1988: *"How was your situation with regards to overtime last month? Did you work overtime? If yes, how many hours?"* Hence, information on overtime is related to the month before the interview. We divide the number of overtime hours by 4.3 to analyse weekly hours. The question on individual union membership has been asked in 1985, 1989, 1993, 1998, 2001, 2003, 2007, 2011 and 2015. For the main analysis, we utilise all these nine survey waves and, hence, consider the thirty-year time span from 1985 to 2015. We additionally employ information provided in the questionnaire in 1985 on contractual and actual working hours including overtime. We use their difference as a measure for overtime since the direct information is not available in that wave.[55]

We control for a broad set of personal characteristics (gender, nationality, education, marital status, whether a child is living in the household, labour market experience in years and experience squared), features of the job (inter alia, log of the hourly wage, contractual working time, tenure in years and tenure squared, whether individuals are concerned about job security and have changed their job in the previous year), and the occupational status (10 categories, differentiating between various types of blue- and white-collar employees). Moreover, we include firm-size variables, industry (1-digit), regional (at the level of federal states, 'Bundesländer'), year and month of the interview dummies. Table C.1 in the Appendix contains variable definitions.

When evaluating the channels by which union membership affects overtime, we also employ variables that are not included in all waves or have missing values. For instance, we use the information on the usual type of overtime compensation: The feasible answers are (1) leisure, (2) partly leisure, partly paid, (3) paid and (4) unpaid. The relevant information is missing in 2015, but otherwise available for all individuals who work non-zero overtime hours and, additionally, for a considerable fraction of those who do not.[56] Furthermore, the SOEP provides information on whether a respondent works in an establishment in which there is a works council every five years since 2001. Because these years do not always coincide with those for which information on union membership is available, we impute the information on the existence of a works council for relevant years from 1998 to 2015.[57]

We restrict our sample to paid employees aged between 18 and 65 years who state that they work full-time and have a contractual working time of at least 35 hours per week (cf. Hunt 1999, Gralla et al. 2017). We focus on this group because the reasons

[55]We might also focus on the time span from 1998 to the present, as Germany was re-united in 1990 and because of the economic transformation especially in Eastern Germany at the beginning of the 1990s. Such restriction does not qualitatively change results reported in the following tables. They are available upon request.

[56]Respondents are asked about the usual type of overtime compensation if they work overtime at least occasionally. Some individuals did not work overtime in the previous month, but nonetheless responded to this question. If individuals never worked overtime, no information on overtime compensation is available.

[57]In order to ensure that the imputed works council information applies to the same employer, we require that the respondent has sufficient tenure in the respective survey year. Since works council elections usually take place every four years between March and May, we additionally take the interview month in the SOEP into account. For a similar approach, see Goerke and Jeworrek (2021). Due to panel attrition, imputation is not feasible for all respondents, resulting in a drop in sample size. Note that the SOEP provides information on collective bargaining coverage for 1995 and since 2015 only.

for working overtime may be different for those who can already retire or do not work full-time, such that our restrictions allow for a more consistent interpretation of union membership effects.[58] We exclude self-employed individuals, apprentices, and marginal employees. In preliminary regressions, we estimated the specifications outlined below separately for the private and the public sector. They consistently revealed the absence of any significant correlation in the public sector. Therefore, we subsequently focus on the private sector.[59] This yields our main sample of nearly 33,000 observations.

Table C.2 in the Appendix shows descriptive statistics for our main sample. Slightly more than 50% of respondents worked overtime in the month before the interview. The unconditional average amount of overtime, i.e. including employees who did not work more than their contractual obligations, is 2.6 hours per week. This number is comparable to those calculated in other studies (Pannenberg and Wagner 2001, Gralla et al. 2017) and higher than the figure provided by the IAB (2020), inter alia, because it also includes overtime compensated by time off in lieu (leisure). This is also the most common form of compensation, namely for 43% of the sample. A quarter of the respondents are compensated partly by leisure and partly by a payment, while fully paid and unpaid overtime are less common. On average, 23% of the sample are union members. The data also reflects the time trend of decreasing union membership, from more than 30% in the 1980s to around 15% nowadays (see also Visser 2019).

4.4.2 Empirical strategy

For analysing the data set, consideration of the structure of overtime hours is crucial. Since these hours are censored to the left by zero, the standard way is to estimate Tobit models, as employed by Trejo (1993), Bauer and Zimmermann (1999), Kraft and Lang (2008) and Gralla et al. (2017). A corresponding Tobit model reads as the follows:

$$OT^*_{it} = X_{it}\beta + UM_{it}\beta^{UM} + \alpha_i + \epsilon_{it} \ , \tag{4.1}$$

$$OT_{it} = \begin{cases} OT^*_{it} & \text{if } OT^*_{it} > 0 \\ 0 & \text{if } OT^*_{it} \leq 0, \end{cases} \tag{4.2}$$

where the underlying latent variable is denoted by OT^*_{it}, while OT_{it} indicates the observed overtime hours within the last month by individual t in year t. UM_{it} takes the value of one if individual i is a member of a trade union, and zero otherwise. X_{it} denotes

[58]Results reported in Tables 4.2 to 4.4 below also hold for a sample which additionally includes individuals who do not work full-time.

[59]Results for specifications including public sector employees and for a public sector sample are available upon request.

the vector of covariates, while ϵ_{it} is the error term. α_i could be thought as an additional component, which captures time-invariant individual heterogeneity. In the pooled Tobit, we do not consider α_i and the underlying assumption is that our variable of interest UM_{it} and the covariates X_{it} are neither correlated with α_i nor with ϵ_{it}. Moreover, the approach assumes independent errors for all observations of each individual.

However, panel data may not fulfil the assumption of independent errors within individuals. Random effects (RE) models consider this correlation in the panel data structure. Consequently, RE specifications yield more efficient estimates than pooled specifications. But still, the underlying assumption is that the heterogeneity is uncorrelated to the observables.

A common approach introduced by Mundlak (1978) allows for correlation between the heterogeneity and the controls by adding correlated random effects (CRE) to nonlinear models. In this approach, the individual heterogeneity α_i is assumed to be correlated with the individual time averages of control variables $\overline{X}_{it}$ and our variable of interest $\overline{UM}_i$:

$$\alpha_i = \overline{X}_i\delta + \overline{UM}_i\delta^{UM} + \mu_i \tag{4.3}$$

Combining equations 4.1 and 4.3 yields the estimated regression equation for the CRE approach.[60] With respect to overtime, Kalwij and Gregory (2005) employ a CRE approach with first differences. The authors replicate former studies on overtime and conclude that the consideration of a correlation between α_i and X_{it} is of empirical importance. However, if δ and δ^{UM} in equation 4.3 are not different from zero, the CRE approach collapses to a standard RE approach (equation 4.1). Hence, a comparison of the different Tobit models allows us to constitute the proper method.

We use cluster-robust standard errors at the individual level, in order to accommodate the panel structure of our data.

4.5 Results

In a first step, we depict descriptive evidence for our key variables – overtime incidence and hours – by union member status. Subsequently, we present the results of regression analyses. Moreover, by looking at various subsamples we can analyse the channels discussed in Section 4.3, by which union membership may impact on overtime outcomes.

[60] In contrast to fixed effects models, which aim to eliminate the unobserved heterogeneity by differentiating out, a nonlinear CRE model does not suffer from the 'incidental parameter problem' which could lead to inconsistent estimates (Cameron and Trivedi 2005). Tobit models are preferred over *linear* panel methods, which are not affected by this potential inconsistency but are not able to consider the high fraction of zero overtime hours and are likely to lead to inconsistent estimates (Cameron and Trivedi 2005).

4.5.1 Descriptive evidence

In the month before the interview, 47.2% of all union members worked overtime, while 56.5% of non-members did. The unconditional average number of weekly overtime hours was 2.1 hours for union members. They worked 0.64 fewer hours than non-members (see Table 4.1).

Table 4.1: Mean comparison

	Union members		Non members		Difference	
	mean	sd	mean	sd	diff	t-value
OT dummy	0.472	(0.499)	0.565	(0.496)	-0.093^{***}	(-14.20)
OT hours	2.128	(3.525)	2.768	(3.973)	-0.640^{***}	(-13.44)
Observations	7,568		25,176		32,744	
Unpaid dummy	0.066	(0.249)	0.208	(0.406)	-0.142^{***}	(-29.56)
Paid dummy	0.162	(0.369)	0.147	(0.354)	0.016^{**}	(2.61)
Partly paid dummy	0.265	(0.442)	0.232	(0.422)	0.034^{***}	(4.71)
Leisure dummy	0.506	(0.500)	0.413	(0.492)	0.092^{***}	(11.32)
Observations	4,863		16,039		20,902	

Source: SOEP. t-tests on the equality of means; $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
Information on the usual type of compensation is not always available (see Chapter 4.4.1 for details).

Union members are much less likely to work unpaid overtime (only 6.6%) than non-members (20.8%). The reverse is true for overtime which is compensated by time off in lieu (50.6% for members versus 41.3% for non-members). Regarding fully and partly paid overtime, the means are more similar.

The descriptive results provide first evidence on the hypothesis that union membership insures employees against the loss of returns from an investment in the form of (unpaid) overtime. This mechanism could explain the observed relation if union members obtain greater expected returns and, thus, worked more such unpaid hours. The substantially lower likelihood of union members, relative to non-members, to work unpaid overtime and the considerably higher fraction of overtime compensated by leisure do not support such a view.

4.5.2 Main results

Table 4.2 reports our main regression results. It depicts the estimated coefficients for the variable of interest, that is, the union membership dummy. Column (1) in Table 4.2 contains results for a pooled Tobit specification, column (2) according estimates for a RE Tobit model, which takes the individual heterogeneity into consideration, and columns (3) and (4) for the CRE Tobit model, which includes time averages of the explanatory variables. The first row in Table 4.2 shows the coefficients of the Tobit models, while the subsequent rows delineate corresponding average marginal effects for the probability of working overtime ($PR(OT > 0)$), for the impact on the number of overtime hours of individuals working overtime ($E(OT|OT > 0)$) and for the overall effect on observed overtime hours ($E(OT)$), respectively.

Table 4.2: Main regression results (1985-2015)

	(1)	(2)	(3)	(4)
	Pooled Tobit	RE Tobit	CRE Tobit	
Union member			UM_{it}	$\overline{UM}_i$
Coeff.	-0.455^{***}	-0.423^{***}	-0.388^{*}	0.025
	(0.112)	(0.104)	(0.179)	(0.220)
$Pr(OT > 0)$	-0.029^{***}	-0.027^{***}	-0.025^{*}	0.002
	(0.007)	(0.007)	(0.012)	(0.014)
$E(OT\|OT > 0)$	-0.177^{***}	-0.163^{***}	-0.150^{*}	0.010
	(0.044)	(0.040)	(0.069)	(0.085)
$E(OT)$	-0.239^{***}	-0.221^{***}	-0.203^{*}	0.013
	(0.059)	(0.054)	(0.093)	(0.115)
N of obs.		32,744		
left-censored		14,949		
uncensored		17,795		
N of groups		16,590		

Source: SOEP. Standard errors clustered on the individual level in parentheses; $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The dependent variable is overtime hours. Coefficients and average marginal effects are displayed. Full set of controls is used in each specification. Time averages of the all covariates are included in the CRE model. Full results of the RE model are displayed in Table C.3 in the Appendix.

The first main insight from Table 4.2 is that individual trade union membership has a negative effect on overtime across all specifications. Compared to the pooled Tobit, panel data models yield only slightly smaller estimated coefficients and average marginal effects for the union membership dummy. Moreover, the point estimates for the time averages of union membership $\overline{UM}_i$ suggest no selection effects, since it is very small in magnitude and not significantly different from zero. Therefore, we find no evidence that union members work less overtime than non-members because of different preferences or attitudes relating to working time.

Since there is no indication of unobserved heterogeneity in the CRE Tobit, the RE Tobit constitutes our preferred model. Accordingly, full estimation results are displayed for this approach in Table C.3 in the Appendix. Moreover, we focus on the RE Tobit in the further course of our analysis.[61]

Focusing on the preferred RE specification, we see from column (2) that union members have a probability of working overtime, which is 2.7 percentage points lower than that of non-members. Furthermore, union membership reduces the amount of overtime

[61] Full estimations results for the pooled Tobit and the CRE approach are available upon request. These approaches yield qualitatively similar results for most subgroup analyses presented below.

by 0.16 hours, conditional on working overtime. Consequently, the overall effect amounts to 0.22 hours, that is, union members work about 13 minutes less overtime per week than non-members. The overall effect in our preferred RE specification corresponds to a reduction of 8.4% (=-0.221/2.62), relative to average overtime hours of 2.62 of the sample.

When estimating our preferred specification without any and with restricted sets of control variables (results available upon request), we obtain a coefficient which is 1.75 times greater (around -0.75), as long as we exclude the 10 dummy variables distinguishing between different qualifications for blue- and white-collar employees. This shows that the extent of overtime depends to a large extent on the occupational status, while it is also related to an individual's union membership.

Inspection of covariates (see Table C.3) reveals that women work less overtime (see Kraft and Lang 2008, Zapf 2015, Gralla et al. 2017). Having a university degree is associated with higher overtime outcomes. Furthermore, overtime rises with the employees' hierarchical level and is generally higher for white-collar than for blue-collar workers. Moreover, the hourly wage is negatively associated with overtime. Having children in the household, being married and the number of contractual working hours have no statistically significant impact,[62] while working in a larger firm raises overtime. In addition, individuals who are concerned about the security of their job work less overtime than those who are not. Finally, we see from Table C.3 that labour market experience has a hump-shaped effect, while we find a linear and negative effect of tenure.[63]

We have conducted various robustness checks, which confirm our main findings, and are described in a summary manner below. Detailed results for these and all further estimations not reported at large are available upon request. In particular, we find only small effects on the magnitude of the coefficients of interest and on their significance levels if we estimate all specifications from Table 4.2 for a sample of individuals, for whom we have information on personality traits and also include these indicators as additional control variables.[64] This suggests that the observed correlation does not arise because union members exhibit certain attitudes or characteristics, reflected by the Big Five, which induce them to work a different amount of overtime.

[62] In Tobit specifications contractual working hours are sometimes positively associated with overtime (Bauer and Zimmermann 1999, Gralla et al. 2017), and sometimes negatively (Kraft and Lang 2008). Pannenberg and Wagner (2001) present fixed-effects estimates and observe a negative effect.

[63] While some of the previous analyses of overtime incorporate linear measures of tenure (Hübler 1989, Zapf 2015) or tenure classes (Kraft and Lang 2008), Bauer and Zimmermann (1999), Pannenberg and Wagner (2001), Kraft and Lang (2008), and Gralla et al. (2017) include age and/ or tenure polynomials in their specifications. Bauer and Zimmermann (1999) and Gralla et al. (2017) observe overtime to be inversely U-shaped in age, which is commensurate with our findings with regard to experience.

[64] Information on personality traits is provided for the years 2005, 2009 and 2012/2013. Variables for 15 items on personality are condensed into five factors which correspond to the widely used Big Five traits. Since no survey year containing the personality traits coincides with the information on union membership, we have to impute the Big Five factors for each observation.

Moreover, our results do not change qualitatively if we omit variables, which could be considered as potentially endogenous or as capturing time-invariant characteristics, such as wages and perceived job insecurity. Also, the exclusion of white-collar workers with highly qualified or executive tasks does not affect our findings. Furthermore, data on time use and income might be prone to measurement errors. Excluding according outliers (upper 5th percentile, respectively) has no impact.

In addition, we have investigated whether the findings depicted in Table 4.2 differ for men and women, blue- and white-collar employees, and between West and East Germany, because overtime sometimes varies according to these dimensions (Hübler 1989, Bauer and Zimmermann 1999, Zapf 2015, Gralla et al. 2017). Our estimates show that the effects are somewhat stronger for men and in West Germany. Interestingly, a separate estimation for blue-collar workers reveals a quantitatively smaller correlation between individual union membership and overtime, which is no longer significant (t = 1.51). This is consistent with our findings (see Table C.3 in the Appendix) and results reported by Bauer and Zimmermann (1999), Gralla et al. (2017) and also Zapf (2015) for unpaid overtime that blue-collar employees work less overtime than white-collar employees. Hence, there is less scope for union-membership effect.

4.5.3 Channels

In this section, we systematically assess the channels by which union membership could affect overtime. The basic idea is that comparing the estimates for the RE specification depicted in Table 4.2 to those for various subsamples, we can evaluate the empirical relevance of a proposed channel.

First, we consider individuals who are overemployed. If we do not observe a membership effect for this subsample, our findings would not be consistent with the idea that unions protect their members against excessive overtime demands by employers. Second, we consider the impact of union membership on paid and unpaid overtime work separately. This will enable us to investigate the notion of an income-enhancing role of trade union membership and to offer further evidence on the idea that membership provides insurance for the returns from investing into unpaid overtime. Third, we look at a subsample of individuals who have not changed their employer between subsequent SOEP waves in our sample, i.e. waves including information on union membership.[65] The results of this stayer sample can help to evaluate various effects, especially if we incorporate the findings from the correlated random-effects Tobit models: (1) If employees have to establish their working time preferences when starting a job in a new firm or have to signal their work attitudes, they may work different amounts of overtime after a job change, relative to a situation in which they have already communicated this information to their superiors (Pannenberg 2005, Anger 2008). If this informational role of overtime work differs between union members and non-members, it is unlikely to generate a union membership effect in a subsample of individuals who have not changed their employer. (2) The estimated union membership effect in correlated random-effects Tobit models cannot be due to a works council or collective bargaining impact, unless firms altered that status. (3) Based on these estimates, focusing on individuals without job change allows us to also account for the selection argument. In a fourth subsample, we look at establishments for which we have information about the existence of a works council. If the effects for council plants are weaker than for non-council ones, this would be consistent with the works council argument. Finally, we look at large firms with 2000 or more employees, which are very likely covered by a collective agreement (see Schneider and Vogel 2018). Therefore, we may expect the union membership impact to be less pronounced in such establishments if collective bargaining reduces overtime.

In Table 4.3 we present results for RE Tobit specifications for subsamples of individuals who are overemployed (column 1), of those who did paid overtime (column 2), or worked overtime which was either completely or partially unpaid (column 3). Finally, column 4 is based on a subsample of individuals who have not experienced an employer change.

Column 1 shows that the union membership effect for overemployed is significantly negative and larger than for the full sample (cf. Table 4.2). This is consistent with a protective role of union membership. Moreover, corresponding regressions for non-overemployed individuals yield estimates, which are substantially smaller in absolute magnitude (and statistically insignificant).

Column 2 indicates that union membership is associated with a lower level of paid overtime, but certainly not a higher one as the discussion in Section 4.3 suggested. Thus,

[65] For each individual, we identify the longest employment spell at one employer. If there are several spells with a maximum number of included waves for an individual, we take the earliest one. This procedure leads to a maximum number of observation, compared to the following alternatives. First, we can consider only the first employment spell for generating a stayer sample or, second, only the last spell. Third, we can exclude all individuals who experienced any job change during the period in which they were interviewed by the SOEP. All definitions yield similar results. Note that for the identification of a job change we do not employ the variable provided by the SOEP, which may also refer to within-firm changes of the job. Instead, we use the information on tenure, which should not capture internal career developments.

Table 4.3: Subsamples I

Union member	(1) Over-employed	(2) Paid OT	(3) Unpaid OT or leisure comp.	(4) Stayer sample
Coeff.	-0.572^{***}	-0.570^{*}	-0.620^{***}	-0.496^{***}
	(0.116)	(0.272)	(0.143)	(0.132)
$Pr(OT > 0)$	-0.038^{***}	-0.036^{*}	-0.040^{***}	-0.032^{***}
	(0.008)	(0.017)	(0.009)	(0.009)
$E(OT \vert OT > 0)$	-0.267^{***}	-0.267^{*}	-0.294^{***}	-0.193^{***}
	(0.054)	(0.128)	(0.068)	(0.052)
$E(OT)$	-0.370^{***}	-0.371^{*}	-0.405^{***}	-0.261^{***}
	(0.075)	(0.177)	(0.093)	(0.070)
Observations	20,486	3,145	12,749	18,690
left-censored	6,520	1,011	3,872	8,473
uncensored	13,966	2,134	8,877	10,217
Number of groups	11,937	2,405	7,778	6,181

Source: SOEP. Standard errors clustered on the individual level in parentheses; $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. RE Tobit estimates and corresponding marginal effects, the dependent variable is overtime hours. Full set of controls is used in each specification.

we can reject the hypothesis that union membership has an income-generating function via greater use of paid overtime.

Column 3 indicates that union membership also has a significantly negative impact on overtime which is completely unpaid or compensated by additional leisure. Accordingly, we can reject the hypothesis that union membership insures against the loss of returns from (unpaid) overtime investments because there is no evidence that union members work more unpaid overtime. We additionally estimated the model of column 3 separately for overtime hours that are unpaid on the one hand and that result in time off in lieu on the other, because additional leisure compensation is the predominant form of compensation (see Table 4.1). The findings support the conclusion that trade union membership does not help to guarantee gains from working unpaid overtime.

The fourth column in Table 4.3 indicates that union members who have not changed their employer work significantly less often and fewer overtime hours than non-members. Accordingly, there is no indication that our findings are due, for example, to union members selecting into firms with different overtime practice. Moreover, because the results for the stayer sample are qualitatively the same, irrespective of whether we consider a RE Tobit model (as depicted in column 4) or a CRE Tobit model (not shown), the findings do not provide evidence commensurate with the works council and collective bargaining channel.

We can also incorporate information about the works council status of an establishment explicitly. As indicated above, this data is available for a substantially reduced number of observations. We initially estimate our standard specification for this reduced sample, encompassing about 14,300 observation. The estimated coefficient of the union membership dummy is significantly negative (column 1 of Table 4.4) and quite similar to the coefficient for the full sample (see Table 4.2). We then additionally split the reduced sample into one for employees working in plants without works councils and one in which such institution is present. The estimated membership coefficient becomes statistically insignificant for employees in firms without works council (not depicted) and stays significant for codetermined firms (see column 2 of Table 4.4). Moreover, these separate regressions do not show substantial differences in the estimated coefficient of the membership dummy. Thus, we can reject the hypothesis that the impact of trade union membership on overtime is due to a correlation between union membership and works councils on the one hand and works councils and overtime on the other.

Table 4.4: Subsamples II

Union member	(1) Council sample	(2) Council exists	(3) Less than 2000 employees	(4) 2000 employees or more
Coeff.	-0.419^{**}	-0.334^{*}	-0.363^{**}	-0.456^{**}
	(0.138)	(0.151)	(0.124)	(0.177)
$Pr(OT > 0)$	-0.029^{**}	-0.024^{*}	-0.023^{**}	-0.030^{**}
	(0.010)	(0.011)	(0.008)	(0.012)
$E(OT\|OT > 0)$	-0.177^{**}	-0.143^{*}	-0.138^{**}	-0.187^{**}
	(0.058)	(0.064)	(0.047)	(0.073)
$E(OT)$	-0.244^{**}	-0.196^{*}	-0.186^{**}	-0.253^{**}
	(0.080)	(0.089)	(0.064)	(0.098)
Observations	14,298	8,458	25,066	7,666
left-censored	5,659	5,125	11,673	3,272
uncensored	8,639	3,333	13,393	4,394
Number of groups	9,258	5,464	13,473	4,541

Source: SOEP. Standard errors clustered on the individual level in parentheses; $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. RE Tobit estimates and corresponding marginal effects, the dependent variable is overtime hours. Full set of controls is used in each specification. Columns (1) and (2) are restricted to firms with more than 5 employees since this is the legal threshold for works council elections. Due to a very small number of big firms in agriculture, we exclude the corresponding industry in column (4).

Furthermore, an inspection of columns 3 and 4 of Table 4.4 reveals that the union membership effect does not considerably vary with firms' size. Since collective bargaining coverage is almost universal in very large firms, we obtain no evidence in support of the conjecture that the union membership-overtime nexus is actually such a bargaining artefact. While our data does not allow us to explicitly control for coverage, other findings support this interpretation. First, we do not observe a different union membership effect for works council firms. Since the existence of a works council and coverage by collective bargaining agreements often coincide, the findings presented in columns 1 and 2 of Table 4.4 are not compatible with the view that the union membership-overtime nexus is a consequence of collective bargaining. Second, the estimates from CRE Tobit specifications for the sample of employees who do not change their job and for which we implicitly hold constant (time-invariant) firm characteristics, show a significantly negative union membership effect. This suggests that firm-fixed effects do not play a role. Since, moreover, changes in bargaining status are relatively seldom for large firms, this calls the validity of the collective bargaining channel further into question. Third, all of our specifications include contractual hours as control variable (see Table C.3 in the Appendix). Because they are determined by collective bargaining, if a firm is covered by such agreement, this variable implicitly controls for bargaining status.

4.6 Conclusion

Using data from the SOEP from 1985 to 2015, we find robust evidence that union members who work full-time in the private sector in Germany work about 0.22 hours – or 13 minutes – overtime per week less than comparable non-members. This effect can be observed across a variety of sub-groups. The difference in overtime hours is equivalent to about 0.5% of total weekly working time of about 40 hours and worth about half of the annual membership fee of 1% of gross income in Germany if overtime hours are evaluated at the average gross wage.

We have considered a number of channels, which may enable union members to work less overtime. Our results allow us to rule out most of them. In particular, we obtain no evidence that is consistent with the view that union membership helps individuals to secure the returns from an investment in the form of working unpaid overtime. Furthermore, we have argued that a linkage between overtime use and union membership may arise because works councils influence their allocation and union members are more likely to work in works council establishments than non-members. If works councils reduced overtime, the negative union membership effect may be a works council impact. However, our results for a subsample of observations, for which we know about the works council status, allows us to reject this hypothesis, as the negative overtime effect can also be observed in works council firms. We have also considered the possibility that the union membership-overtime nexus is due to collective bargaining. Unfortunately, the SOEP provides no direct information on collective bargaining status for the relevant period. Looking at large firms, which in their overwhelming majority negotiate wages and working conditions collectively, we observe the trade union membership impact, as well. Considering a sample of individuals who stay with their employer and taking into account that we observe the union membership effect also for works council firms, which are often covered by collective negotiations, provides no support for the collective bargaining hypothesis either. Finally, comparing the results for the RE Tobit and the CRE approach does not suggest that the preference- or attitude-based perspective can explain our findings. The finding that incorporating information on personality traits does not qualitatively alter estimates further supports this view.

The only remaining explanation for the negative correlation is that union members are better able to enforce their working time preferences than non-members. Given the evidence that a large fraction of employees in Germany would like to reduce working time, this impact of union membership is reflected in less overtime. Since we do not observe a union membership impact on overtime in the public sector, the above interpretation suggests that such a protective role as possibly effective in the private sector is not required for employees in the public domain.

While we have been careful not to interpret our findings as establishing a causal impact of an employee's membership in a trade union on overtime, the channels discussed above suggest such perspective. Our data does not allow us to establish a causal relationship, for example, in the absence of a suitable instrument for membership. However, we find no indication that some unobserved factor, which is correlated both with membership and overtime, is responsible for the observed correlation. The feature that we obtain comparable findings for the union membership-overtime nexus in our extensive set of robustness checks would require that the unobserved factor has a similar impact in all subgroups.

One could also be concerned about reverse causation, namely that employees who experience a reduction in overtime and, thus, work fewer hours than they prefer, decide to join a trade union to defend their income. In this case, we would expect particularly strong effects for paid overtime and for underemployed individuals. We have demonstrated above that the nexus is similar for paid and unpaid overtime (see Table 4.3) and we observe no significant effect of union membership on overtime for underemployed individuals (not documented). This makes us confident that causation is from union membership to overtime and not the other way around.

Putting our findings into an international perspective, it is noteworthy that the studies on union membership and overtime in Great Britain tend to find a positive correlation. Once again, this may be interpreted as indicating a differential role of union membership, in this instance, in two countries. Importantly, the findings suggest that research on the effects of trade unions should not ignore specificities of industrial relations systems.

We can lastly interpret the results in the context of recent Covid 19-related developments. Preliminary regressions based on SOEP-data suggest that individuals who can work from home provide more overtime than those who cannot stay away from their normal place of work. Moreover, people working at home are less likely to be union members. This suggests that the gain from becoming a union member, on account of the impact on overtime, may be higher in a work environment in which the office in the firm is increasingly replaced by the desk at home.

Appendix C

Table C.1: Variable definitions

Dependent variables	
OT dummy	Individual worked overtime in the month before the interview
OT hours	Weekly overtime hours in the month before the interview
Independent variable of interest	
Union member	Individual is a trade union member
Type of overtime compensation	
Unpaid dummy	Overtime is generally not compensated
Paid dummy	Overtime is generally compensated with additional pay
Partly paid dummy	Overtime is generally partly compensated with additional pay and partly compensated with time-off
Leisure dummy	Overtime is generally compensated with time-off
Socio-demographic controls	
Woman	Individual is female
Foreign citizen	Individual has a non-German citizenship
Married	Individual is married and living together
Child in household	At least one child is living in the household
Apprenticeship	Individual has completed an apprenticeship
University degree	Individual holds a university degree
Job-related controls	
Log. hourly wage	Logarithm of current gross income from employment
Experience	Labour market experience in years (fulltime)
Tenure	Job at the same employer in years
Contractual work time	Contractual working hours per week
Side job	Individual has a side-job
Concerns about job security	
Job Sec.: No concerns	Individual states no concerns about job security
Job Sec.: Some concerns	Individual states some concerns about job security
Job Sec.: Big concerns	Individual states big concerns about job security
Job change	
Job Change: No	Individual states no recent job change or entered employment for the first time
Job Change: No info	No information available if recent job change occurred
Job Change: Yes	Individual states recent job change
Further categorised controls	
Firmsize	4 dummies indicating the number of employees in individual's firm (1-19, 20-199, 200-1999, 2000+)
Occupation	10 dummies indicating occupational status, distinguishing between different qualifications for blue- and white collar employees
Industry	8 industry dummies on a one digit level (energy and mining combined)
Federal State	15 regional dummies at the level of the federal states with Rhineland-Palatinate and Saarland combined
Interview month	8 month dummies (August-December combined)
Survey Year	9 year dummies
Variables used for robustness checks	
Works council	Existence of work council in firm
Overemployed	Difference between actual and desired working hours greater than zero

Table C.2: Descriptive statistics

	count	mean	sd	min	max
OT dummy	32744	0.54	0.50	0	1
OT hours	32744	2.62	3.88	0	40
Unpaid dummy	20902	0.17	0.38	0	1
Paid dummy	20902	0.15	0.36	0	1
Partly paid dummy	20902	0.24	0.43	0	1
Leisure dummy	20902	0.43	0.50	0	1
Union member	32744	0.23	0.42	0	1
Woman	32744	0.29	0.45	0	1
Foreign citizen	32744	0.18	0.38	0	1
Married	32744	0.63	0.48	0	1
Child in household	32744	0.42	0.49	0	1
Apprenticeship	32744	0.68	0.47	0	1
University degree	32744	0.17	0.37	0	1
Log. hourly wage	32744	2.64	0.46	0	6.473891
Experience	32744	17.71	11.22	0	49.2
Experience squared	32744	439.52	454.71	0	2420.64
Tenure	32744	10.41	9.53	0	50.9
Tenure squared	32744	199.33	318.35	0	2590.81
Contractual work time	32744	39.19	2.60	35	78
Side job	32744	0.06	0.23	0	1
Job Sec: No concerns	32744	0.43	0.50	0	1
Job Sec: Some concerns	32744	0.42	0.49	0	1
Job Sec: Big concerns	32744	0.15	0.36	0	1
Job Change: No	32744	0.85	0.35	0	1
Job Change: No info	32744	0.02	0.14	0	1
Job Change: Yes	32744	0.13	0.33	0	1
BC: Unskilled	32744	0.04	0.19	0	1
BC: Semiskilled	32744	0.17	0.38	0	1
BC: Skilled	32744	0.23	0.42	0	1
BC: Foreman	32744	0.03	0.18	0	1
BC: Master	32744	0.01	0.12	0	1
WC: Foreman	32744	0.01	0.11	0	1
WC: Simple tasks	32744	0.08	0.28	0	1
WC: Qualified	32744	0.24	0.43	0	1
WC: Highly qual.	32744	0.17	0.37	0	1
WC: Executive	32744	0.02	0.13	0	1
Firmsize 1-19	32744	0.22	0.41	0	1
Firmsize 20-199	32744	0.31	0.46	0	1
Firmsize 200-1999	32744	0.24	0.43	0	1
Firmsize 2000+	32744	0.23	0.42	0	1
Works council	15027	0.56	0.50	0	1
Overemployed	32161	0.64	0.48	0	1

BC stands for Blue-Collar, whereas WC stands for White-Collar.

Table C.3: Main results (1985-2015)

	(1) Coeff.	(2) $Pr(OT > 0)$	(3) $E(OT \mid OT > 0)$	(4) $E(OT)$
Union member	-0.423^{***}	-0.027^{***}	-0.163^{***}	-0.221^{***}
	(0.104)	(0.007)	(0.040)	(0.054)
Woman	-2.002^{***}	-0.129^{***}	-0.773^{***}	-1.046^{***}
	(0.105)	(0.007)	(0.041)	(0.055)
Foreign citizen	-1.365^{***}	-0.088^{***}	-0.527^{***}	-0.713^{***}
	(0.142)	(0.009)	(0.055)	(0.074)
Married	-0.084	-0.005	-0.032	-0.044
	(0.095)	(0.006)	(0.037)	(0.050)
Child in household	-0.126	-0.008	-0.048	-0.066
	(0.089)	(0.006)	(0.034)	(0.047)
Apprenticeship	0.599^{***}	0.039^{***}	0.231^{***}	0.313^{***}
	(0.147)	(0.009)	(0.057)	(0.077)
University degree	1.262^{***}	0.081^{***}	0.487^{***}	0.659^{***}
	(0.191)	(0.012)	(0.074)	(0.100)
Log. hourly wage	-1.626^{***}	-0.105^{***}	-0.628^{***}	-0.849^{***}
	(0.152)	(0.010)	(0.059)	(0.079)
Experience	0.144^{***}	0.009^{***}	0.055^{***}	0.075^{***}
	(0.015)	(0.001)	(0.006)	(0.008)
Experience squared	-0.004^{***}	-0.000^{***}	-0.001^{***}	-0.002^{***}
	(0.000)	(0.000)	(0.000)	(0.000)
Tenure	-0.048^{**}	-0.003^{**}	-0.019^{**}	$-0.025**$
	(0.015)	(0.001)	(0.006)	(0.008)
Tenure squared	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Contractual work time	-0.022	-0.001	-0.008	-0.011
	(0.018)	(0.001)	(0.007)	(0.009)
Side job	0.377^{*}	0.024^{*}	0.146^{*}	$0.197*$
	(0.154)	(0.010)	(0.059)	(0.080)
Job Sec: No concerns	*Reference category*			
Job Sec.: Some concerns	-0.418^{***}	-0.027^{***}	-0.162^{***}	-0.218^{***}
	(0.077)	(0.005)	(0.030)	(0.040)
Job Sec.: Big concern	-0.514^{***}	-0.033^{***}	-0.199^{***}	-0.269^{***}
	(0.116)	(0.007)	(0.045)	(0.061)
Job Change: No	*Reference category*			
Job Change: No info	-0.790^{**}	-0.051^{**}	-0.305^{**}	$-0.413**$
	(0.300)	(0.019)	(0.116)	(0.157)
Job Change: Yes	-0.052	-0.003	-0.020	-0.027
	(0.121)	(0.008)	(0.047)	(0.063)
BC: Unskilled	*Reference category*			
BC: Semiskilled	1.099^{***}	0.071^{***}	0.424^{***}	0.574^{***}
	(0.250)	(0.016)	(0.097)	(0.131)
BC: Skilled	1.751^{***}	0.113^{***}	0.676^{***}	0.914^{***}
	(0.255)	(0.016)	(0.099)	(0.133)
BC: Foreman	3.048^{***}	0.196^{***}	1.177^{***}	1.592^{***}
	(0.317)	(0.020)	(0.122)	(0.165)
BC: Master	3.630^{***}	0.234^{***}	1.402^{***}	1.896^{***}
	(0.404)	(0.026)	(0.156)	(0.211)
WC: Foreman	4.077^{***}	0.263^{***}	1.574^{***}	2.129^{***}
	(0.418)	(0.027)	(0.161)	(0.218)
WC: Simple tasks	1.928^{***}	0.124^{***}	0.745^{***}	1.007^{***}

	(0.272)	(0.017)	(0.105)	(0.142)
WC: Qualified	3.195^{***}	0.206^{***}	1.234^{***}	1.668^{***}
	(0.262)	(0.017)	(0.101)	(0.136)
WC: Highly qual.	5.522^{***}	0.356^{***}	2.132^{***}	2.883^{***}
	(0.284)	(0.018)	(0.110)	(0.148)
WC: Executive	8.429^{***}	0.543^{***}	3.255^{***}	4.401^{***}
	(0.421)	(0.027)	(0.163)	(0.219)
Firmsize 1-19	*Reference category*			
Firmsize 20-199	0.672^{***}	0.043^{***}	0.260^{***}	0.351^{***}
	(0.112)	(0.007)	(0.043)	(0.058)
Firmsize 200-1999	0.715^{***}	0.046^{***}	0.276^{***}	0.374^{***}
	(0.126)	(0.008)	(0.049)	(0.066)
Firmsize 2000+	0.827^{***}	0.053^{***}	0.319^{***}	0.432^{***}
	(0.134)	(0.009)	(0.052)	(0.070)
N of obs.	32,744			
left-censored	14,949			
uncensored	17,795			
N of groups	16,590			

Source: SOEP. Standard errors clustered on the individual level in parentheses; $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. RE Tobit estimations, the dependent variable is overtime hours. Average marginal effects are displayed in columns (2)-(4). Full set of controls is used. Dummies for industry, region, survey year and interview month are not displayed.

Chapter 5

Overtime, Works Councils and Collective Bargaining Agreements

Previous studies using SOEP data, show a negative effect of works councils on overtime hours, especially if the contractual working time is 40 hours. In addition to co-determination through works councils at the firm level, collective bargaining coverage is an important factor in the German industrial relations system. Corresponding data was not available in the SOEP before 2015. Therefore, we focus on the 2016 SOEP wave, which also contains information on collective bargaining coverage. We then carry out a cross-sectional analysis, in order to analyze the effects of works councils in firms with and without collective bargaining agreements. Similar to studies analyzing other outcome variables, the results show that the effect of works councils is only prevalent for employees, who are paid according to a collective bargaining agreement.[66]

Keywords: Overtime, works councils, collective bargaining agreements

JEL Classification: J2, J51, J53

[66]This chapter is single-authored.
We are grateful for helpful comments by participants of the brownbag seminar at the IAAEU and of the (virtual) TriECON Workshop on Co-Determination 2020.

5.1 Introduction

Working overtime is a prevalent phenomenon in many employees' working life. Overtime arrangements are not always subject to individual decisions between employers and employees. Instead, the institutional setting provides a framework for such arrangements and may include regulations for overtime work. Hence, we expect that overtime work differs by the institutional setting in which an employee works. The specific institutional setting, in turn, depends on the country being analysed.

In the German system of industrial relations, trade unions and employer associations bargain on working conditions for a large share of employees, whereas works councils represent employees on the firm-level. However, a considerable share of employees are neither paid according to collective bargaining agreements nor represented by a works council. Both institutions can affect overtime since regulations regarding the amount of overtime or remuneration are included in collective agreements or negotiated by works councils at the firm-level. Moreover, interdependencies between these two institutions may play a role. Moderating effects of collective agreements on works councils effects are well established in the literature for various outcome variables (e.g. Hübler and Jirjahn, 2003), but so far not for overtime.

Hence, in this paper, we study how collective bargaining agreements influence the impact of works councils on overtime. Our empirical investigation is based on data from the German Socio-Economic Panel (SOEP). The SOEP data provides increasing information on the institutional background of the respondents' work place. In particular, the 2016 wave provides refined data on collective bargaining agreements and information on works councils. Thereby, it grants the analysis of the interdependencies between those actors.

While there are previous studies analysing the effect of works councils on overtime and using the SOEP (Kraft and Lang, 2008, Gralla et al., 2017), this study additionally considers collective agreements. We first replicate former results on the effect of works councils with more recent SOEP data. We then carry out a cross-sectional analysis and analyze the effects of works councils in firms with and without collective bargaining agreements. Moreover, we investigate the sole effect of collective agreements, since this determinant of overtime was rarely considered in previous literature.

In our study, it turns out that the works council effect prevails only if respondents are covered by a collective agreement. Although information on both institutional actors are often available in firm-level data, no study has found possible interdependencies in this context so far. Hence, to the best of our knowledge, this study provides important and novel insights into the determinants of overtime on the one hand and into the relationship of works councils as well as collective agreements on the other.

This paper is organized as follows. Section 5.2 sketches the institutional background of the German industrial relations system, summarizes theoretical considerations and previous findings from the literature. Section 5.3 introduces the data set and the empirical strategy. Section 5.4 presents results from the regression analysis, while Section 5.5 includes limitations and a summary of the findings.

5.2 Institutional background and previous findings

5.2.1 German system of industrial relations

The German system of industrial relations is characterized by the parallel existence of predominantly industry-wide regulated collective agreements and co-determination through works councils at the firm level. While collective agreements are negotiated between trade unions and employer representatives, works councils can be implemented in firms which employ at least five employees. Although the degree of formal institutionalisation decreased significantly in recent decades, collective agreements and co-determination are still the major pillars in the industrial relations system (Oberfichtner and Schnabel, 2019). In 2020, about one quarter of firms applied collective agreements. These firms employ more than one half of all employees, whose wages and other working conditions are regulated by collective agreements. Although works councils are implemented only in 8% of all firms, 38% of all employees are employed in co-determined firms. This is due to the positive correlation between firm size and works council existence. Moreover, there is a prevalent overlap of having both bargaining coverage and co-determination through works councils, especially in bigger firms with more than 500 employees (Ellguth and Kohaut, 2021).

Collective agreements are usually negotiated between trade unions and employer associations at the industry level or, alternatively, between a trade union and a single firm. These regulations take precedence over agreements between works councils and employers. In general, collective agreements regulate the employee's payment (wage agreement), while most collective agreements also include further regulations (*Mantelbestimmungen*) concerning working time and further job characteristics, e.g. vacation entitlements. Collective agreements are binding for the employer and their employees covered by them (Müller and Schulten, 2019). Alternatively, employers who are not part of a bargaining employer association can orient themselves towards collective agreements and apply these regulations. In recent years, this type of orientation has become more important and partly compensates the decline of employees covered by a binding agreement (Bossler, 2019). The Collective Agreements Act (*Tarifvertragsgesetz*) contains relatively few requirements of collective agreements. In contrast, the Works Constitution Act (WCA, *Betriebsverfassungsgesetz*) includes several components which may be subject to co-determination through works councils at the firm level. So-called opening clauses (*Öffnungsklauseln*) in collective agreements can provide a framework, in which deviations from the agreement are allowed.[67] Predominantly, opening clauses include the call for the need of agreements on the firm level, so called *Betriebsvereinbarungen*, which involves the works councils.

5.2.2 Collective agreements and overtime

If collective agreements also include regulations concerning specific aspects of excess working time or overtime arrangements, we can expect a direct effect on overtime use. Bispinck (2016) analyses various collective agreements with respect to the flexibility of working hours through collectively agreed regulations. Around one half of the analyzed collective agreements either includes a maximum number of overtime hours or a maximum daily working time including overtime (Bispinck, 2016). In most of the analyzed agreements, regulations include an employer's obligation to pay an extra remuneration for overtime work or the opportunity to offer time off in lieu. As Bispinck (2016) points out, extra-paid overtime plays a bigger role in practice than leisure time as compensation. Usually, extra remuneration equals 25% for few overtime hours and goes up to 50% for higher levels of overtime hours. If collective agreements contain a maximum daily working time including overtime, it typically limits the working day to 10 hours (Bispinck, 2016). Both a maximum amount of hours and more costly personnel costs for overtime are likely to decrease the firm's demand for overtime.

Instead of a direct effect of collective agreements on overtime, alternative and rather indirect mechanisms are possible. From the eighties onwards, the unions accepted firms' demand for a greater flexibility of working time – in exchange for an enforced reduction in working time (Hunt, 2013). Indeed, it is documented that covered firms have lower standard working hours than non-covered ones (Kölling and Lehmann, 2001, Kramarz et al., 2008). Based on the fall in standard hours, Jirjahn (2008) argues that collective agreements might have an impact on overtime use due to firms' adjustments as a response to the reduction of contractual working time. From a theoretical point of view, the response to a reduction of contractual hours might either be an increase of overtime work of a constant workforce, in order to compensate the reduction of standard hours, or a decrease of overtime due to new hires of workers or increased capital intensity. Theoretical models depend on certain conditions – including the exogeneity of wages, constant overtime pre-

[67] Opening clauses became increasingly important in recent decades. A general overview of the prevalence and importance of opening clauses is outlined in Kohaut and Schnabel (2007).

mia and the differentiation between standard-hours-only and overtime firms – and do not provide unambiguous predictions for the effect of overtime as a response to a change in standard hours (Jirjahn, 2008, Schank, 2015). In line with the unambiguous predictions of the theory, Jirjahn (2008) summarizes that existing empirical work dealing with the effect of a reduction in standard hours on overtime yields mixed results. As mentioned above, the fall in standard hours did go along with more flexibility. The increased flexibility itself might be another mechanism that affects overtime. Accordingly, Zapf (2016) expects that covered firms have a higher probability for overtime work. She argues that the prevalence of overtime regulations in collective agreements indicates an increased use of overtime in covered firms. It is worth noting that the development of falling standard hours stopped in the 2000s (Kramarz et al., 2008). However, it is likely that standard hours still differ between covered and uncovered firms. Hence, an empirical investigation of actual working hours should not neglect the impact of contractual hours.

Despite the obvious link between collective agreements and overtime due to the widespread inclusion of overtime regulations in such agreements, there are only few studies from Germany, which consider collective agreements as determinants of overtime. To the best of our knowledge, only the following two studies deal with this nexus. Firstly, Jirjahn (2008) observes a negative impact of collective agreements on overtime of blue-collar workers using the Hanover Panel.[68] However, the variable for coverage becomes insignificant when controlling for standard hours, indicating that lower standard hours in covered firms are the underlying mechanism. Secondly, Zapf (2016) finds no effects of collective agreements on the incidence of overtime in firms, based on four waves of the IAB establishment panel (2001, 2007, 2009, 2011). It is worth noting that both studies do not explicitly focus on collective agreements as a covariate, but rather have the focus to reveal general determinants of overtime. Turning to other countries, the literature provides very mixed results on the effects of collective agreements or unionisation (the equivalence in the Anglo-Saxon system of industrial relations) on overtime. These mixed findings were recently summarized in Goerke and Schultze (2021) and can be interpreted in that way that country-specific characteristics of the industrial relations' systems play an important role.

5.2.3 Co-determination and overtime

Apart from collective agreements, co-determination at the firm level can affect overtime. According to §87(1) WCA, works council participate in decision making on temporary reduction or extension of normal working hours, which includes overtime. If the company

[68] The Hanover Panel is a sample of firms from the manufacturing sector in the federal state of Lower Saxony collected in the nineties.

applies a collective agreement, the council can decide on aspects which are not covered by the agreement (§77(3) WCA). Moreover, opening clauses can determine the degree of possible agreements on the firm level. In fact, most collective agreements imply that the specific working time arrangement should be implemented by company-level agreements (Bispinck, 2016). In line with this, Walwei et al. (2021) show that working time accounts and overtime are one of the most important topics of such agreements. Although works councils are not elected by employees with executive tasks (§5(3 & 4) WCA) and obviously do not represent these employees, works councils generally also represent so-called non-tariff employees. Since it is feasible that overtime arrangements on the firm-level are regulating the overtime work of tariff (but not non-tariff) employees, a differentiated analysis of these groups seems essential.[69] Beside agreements on the firm-level, works council also provide information services for employees in a firm. Hence, there is the possibility that better informed employees work a different amount of overtime, compared to employees in not-co-determined firms.

As a representative of the employees, the question whether works councils support or resist overtime work depend on the employees' interests. If overtime hours are remunerated with extra-pay, employees are incentivized to work more overtime hours in order to enhance their income. In case of a wage premium in firms covered by a collective agreement (Gürtzgen, 2016) or even wage cushion (Jung and Schnabel, 2011), there might be higher supply for overtime in covered firms. This call for more overtime could be accentuated by councils. However, given the prevalence of 'overemployed' workers who want to reduce actual working time (Knaus and Otterbach, 2019), support for increased overtimes hours is quite unlikely. Similarly, Schank and Schnabel (2004) argue that works councils may try to increase wages but to reduce working time. Beside collective voice effects, works councils could also lead to higher adjustments costs due to their co-determination rights regarding hires and fires. This would go in line with more overtime in co-determined firms (Schank and Schnabel, 2004).

Empirical studies on the effect of works councils yield mixed results pending on the kind of data. Classical studies using firm-level data from the IAB establishment data find a positive impact (Kölling, 1997, Schank and Schnabel, 2004), whereas more recent studies using individual data from the SOEP find a negative impact (Kraft and Lang, 2008, Gralla et al., 2017). Moreover, Kraft and Lang (2008) separate introduction effects from selection effects and find that the works council introduction itself is not responsible for fewer hours of overtime. Using a special linked-employer-employee data set from the SOEP, Zapf (2015) finds a negative effect on the probability of paid and unpaid overtime.

[69] In contrast to other data sources, the SOEP data allows to identify non-tariff employees (cf. Section 5.3), such that our analysis considers this issue.

Also focussing on the incidence of overtime work and not hours, Zapf (2016) finds no works council effect in the IAB establishment data. Similarly, Jirjahn (2008) does not find any effects of works councils on the incidence of overtime and overtime hours, based on regional firm-level data.

5.2.4 The interplay of collective agreements and co-determination

The effectiveness of works councils might depend on the collective agreement status. In their significant study published in 1995, Freeman and Lazear argue that works councils can focus on further working issues at the firm-level, if bargaining on wages and other elementary working conditions are externalized to a higher level in form of industry-wide collective agreements. This would lead to a higher effectiveness of the works council's work and thereby stronger results. The literature speaks of "moderating effects" of collective agreements on works council behaviour (Brändle, 2017, Jirjahn, 2017). In line with this argumentation, the literature provides empirical evidence on heterogeneous effects for various outcome variables. Hübler and Jirjahn (2003) show that works councils in covered firms have a stronger impact on productivity but a smaller effect on wages, compared to uncovered firms.[70] Moreover, works councils may have a stronger effect on family-friendly work practices and equal opportunity practices for women in covered firms (Heywood and Jirjahn, 2009, Jirjahn and Mohrenweiser, 2021). Furthermore, councils have a stronger effect on reducing quits in covered firms (Pfeifer, 2011, Adam, 2019) or increasing apprenticeship training in covered firms (Kriechel et al., 2014). Such heterogeneous effects might also depend on the type of collective agreements. For example, Brändle (2017) finds wage and productivity effects of works councils only in firms with an industry-level agreements but not for firm-level agreements. On the one hand, it may be the case that firm-level agreements include more regulations and are, thereby, more specific. On the other hand, the externalization of bargaining processes is not applicable in case of firm-level contracts.

Previous studies also provide further explanations as to why the works council's impact can differ between firms with and without collective agreements (see Pfeifer, 2011, Kriechel et al., 2014, Adam, 2019). These studies highlight the following possible explanations. First, if a collective agreement applies to a firm, works councils are able to better ensure the compliance since they can refer to these regulations. Second, trade unions might provide legal advice which helps a works council to achieve the bargaining objectives. The fact that most works councillors are trade union members (Behrens, 2009b)

[70] While the empirical effect on productivity was confirmed by a series of studies, the effect on wages is not consistently confirmed (see Jirjahn (2017) for an overview). According to this overview, studies using the Hanover Firm Panel yield stronger effects of works councils in uncovered firms, while a series of studies using IAB data yield clearer effects in covered firms or very similar effects as in Brändle (2017).

should ease cooperation between councils and unions. Third, the bargaining atmosphere on the firm-level might be more cooperative if wages are regulated by collective agreements. All these explanations can help works councils to enforce their interests regarding regulations on overtime and, hence, lead to a more pronounced works council effect in firms covered by a collective agreement.

Empirical studies on overtime, which consider both collective agreements and works councils, are extremely rare. Exceptions are Jirjahn (2008) and Zapf (2016) who both – as already mentioned – find no effect of works councils. While Jirjahn (2008) finds a negative effect of collective agreements through a reduction in contractual working hours, Zapf (2016) yields insignificant estimates for the collective agreement variable. Instead of running separate regressions, the authors employ both variables as covariates. In a robustness check, Jirjahn (2008) uses interacting terms between works council and collective agreement status, but does not find significant effects. Similarly, Zapf (2016) does not reveal significant estimates of the corresponding interaction on the incidence of overtime work. Both studies use data on the establishment level. While the study of Jirjahn (2008) is limited to one German federal state, the IAB data only allows to consider the incidence, but not the amount of overtime.

Therefore, to the best of our knowledge, the following analysis constitutes the first investigation that extensively analyses the interplay of works councils and collective agreements and their effect on overtime incidence as well as hours, and the first to use comprehensive individual data for Germany.

5.3 Data, descriptives and empirical strategy

5.3.1 Data

For the empirical analysis, we utilize data from the SOEP. This dataset is organized by the German Institute for Economic Research (DIW) and covers responses from thousands individuals and households annually since 1984 (see Goebel et al. 2019).

Our outcome variable contains information on the incidence and on the amount of overtime work in the month preceding the interview. The corresponding set of questions reads as *"How was your situation with regards to overtime last month? Did you work overtime? If yes, how many hours?"* For our data analysis, we calculate average weekly hours by dividing monthly values by the number of weeks per month (4.3). Since the late eighties, information on overtime work is annually included in the SOEP data.

The first explanatory variable of interest is the works council status. The SOEP questionnaire included the question "*Does an employees' council exist at your place of work?*"

in a five year interval from 2001 onwards, i.e. 2001, 2006, 2011, 2016, and additionally in 2019. The answer options are binary (yes vs. no) such that we build a dummy variable for the existence of works council.

Beside works councils, our focus lies on the information whether the respondents' employer applies collective agreements. From 2015 to 2017, the SOEP provides refined information on this issue.[71] The corresponding question reads as "*Are you paid according to a collectively agreed wage agreement?*". The possible answers are (1) "*Yes, a legally binding company wage agreement.*", (2) "*Yes, paid according to a collective wage agreement that is not legally binding for this sector/company.*", (3) "*Yes, a legally binding collective wage agreement.*", (4) "*No, my job is exempt from the collective wage agreement in place where I work.*", (5) "*No, there is no collective wage agreement.*" and lastly "*Don't know.*". This highly differentiated question was surveyed by the SOEP in the years 2015, 2016 and 2017 and allows to define *coverage* by collective agreements. In 2018 and 2019, the SOEP substituted the differentiation with 5 items with a binary answer option (yes vs. no). Since the former category (2) is no longer surveyed, the binary version rather refers to collectively determined wages than coverage (see Goerke and Huang (2022) where the same distinction is applied.) If a respondent is paid according to a collective wage agreement without being covered by a collective agreement, it is unclear whether the employer also orients towards collectively determined working time regulations.[72] Hence, we abstain from considering information on collectively determined wages, but focus on collective bargaining coverage.

In the regression analysis in Section 5.4, we first investigate the impact of works councils using the above-mentioned five waves containing the corresponding information. We then investigate whether the works council effect differs by the collective bargaining status. Since only the 2016 wave provides detailed information on collective agreement and coincides with information on works councils, we use this wave for the joint analysis. Hence, we carry out a cross-sectional analysis.

For the subsequent analysis, we restrict our estimation sample to dependent employees who are older than 18 and younger than 65 years. We exclude self-employed individuals, apprentices, and marginal employees. Moreover, we restrict the sample to individuals

[71] In 1995, the SOEP questionnaire included less detailed information on the collective bargaining status (3 categories). Since the works council information was not surveyed in this time, we do not utilize this information. Moreover, a special questionnaire for migrants included information on collective agreement in 2014.

[72] This view is supported by the findings of Bossler (2019). According to his study, firms which voluntarily apply collective wage agreements pay fewer wages than firms bound by such agreements. Similarly, Ellguth and Kohaut (2021) show that only one fifth of firms that orient themselves towards collective agreements apply tariff-like regulations. While most firms apply only collectively determined wages to their majority of employees, the remaining share of 17% loosely orients towards collective agreements.

working in firms with at least 5 employees, since this threshold constitutes the possibility to implement works councils.

In order to ease interpretation and to draw general conclusions, we additionally restrict the sample to employees stating that they work full-time and at least 35 hours per week. Our sample of the survey year 2016 does not allow a proper analysis of the subgroup of part-time workers.

5.3.2 Descriptives

Since the disaggregated information on collective agreements was included in the SOEP only from 2015 to 2017, studies using this information are quite rare.[73] Information on the distribution of collective agreements in Germany is usually derived from the IAB establishment panel (WSI, 2022).[74] The establishment panel distinguishes between company-level, industry-level and no collective agreement. If the firm is does not apply an agreement, the questionnaire additionally asks whether the firm orients itself towards a current collective agreement.

Figure 5.1 shows the distribution of collective agreements according to the IAB establishment panel (left panel) and the information derived from the SOEP (right panel). Both panels show the distribution of employees, since the data from the IAB establishment panel is weighted with the number of employees. The share of firm-level agreements is relatively similar across both data sources, whereas the share of industry-wide agreements differs at a first glance. In contrast to the IAB data, the SOEP data allows for the differentiation between employees who work in a covered firm but are not paid according to a collective agreement since they take, for example, a leading role in the workforce (non-tariff-employees). Hence, the non-tariff-employees in the SOEP work in firms with industry-wide agreements, as counted by the IAB data. Furthermore, the pie charts shows that not all employees are aware of the fact that their employer orients themselves towards collective agreements or that their payment is affiliated to an agreement, since establishment data provides significantly higher shares.[75] Overall, we do not find substantial differences between these two data sources reflecting collective bargaining coverage in Germany. In line with this finding, Schulten (2019) reports similar shares of

[73] Exemptions are a descriptive study by Schneider and Vogel (2018) and a multivariate analysis of the application of minimum wages by Pusch (2018). More recent studies rather use the binary information on collectively agreed wages (Goerke and Huang, 2022, Bonaccolto-Töpfer and Schnabel, 2023).

[74] A short overview and comparison of German data on collective bargaining coverage is also provided by Schulten (2019).

[75] A share of around 8% of respondents in the SOEP stated that they do not know whether they are paid according to a collective agreement. While this can be seen as an indicator for measurement errors in the SOEP data (Fitzenberger and Sommerfeld, 2016), the option to admit not to know the collective agreement status prevents misstatements in the other categories. We drop this category from the subsequent analysis.

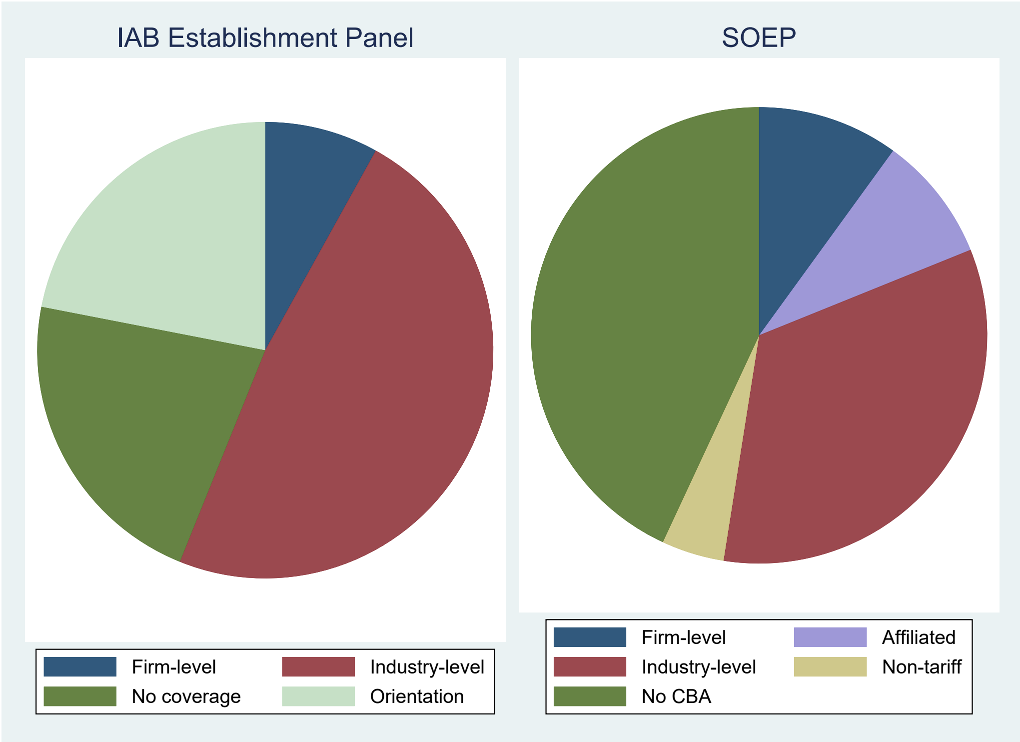

Figure 5.1: Distribution of collective agreements by different data sources
Source: WSI (2022) and SOEP.

collective bargaining coverage when comparing IAB and SOEP data, while the German Structure of Earning Survey provides a significantly lower share.

Table 5.1 shows the distribution of collective agreements and mean of works council existence for the working sample.[76] While around 70% of all firms which are bound to or affiliated with collective agreements have a works council; the average share is less than 30% for firms without applying a collective agreement. Within the former group, firms applying industry-wide collective agreements have the highest share of works council, whereas affiliated firms are co-determined by works councils less often. Around 55% of respondents in the whole sample are employed in firms with a work council.[77]

Descriptive statistics of all variables used in the subsequent analysis are shown in Table D.1. In the full sample, employees work on average 2.6 overtime hours per week. Since two out of five do not work overtime at all, non-zero overtime hours are on average considerably higher (around 4.3).

[76]The working sample constitutes full-time working employees in 2016 (for details on the sample restriction see Section 5.3.1), such that the numbers in Table 5.1 are not comparable to the shares in Figure 5.1.

[77]This number is somewhat higher than reported by Ellguth and Kohaut (2021) but comparable to other studies using the SOEP data (e.g. Gralla et al., 2017).

Table 5.1: Distribution of collective agreements and mean of works council existence

	Frequency	Percent	Mean(WC)
Firm-level	543	12.34	0.73
Affiliated	453	10.29	0.61
Industry-level	1300	29.55	0.77
Non-tariff	307	6.97	0.72
No CBA	1798	40.85	0.29
Total	4401	100.00	0.55

Source: SOEP.

5.3.3 Empirical strategy

The dependent variable in our regression analysis is overtime hours. As summarized in the last subsection, overtime hours are zero for a considerable number of observations. We assume that there is a latent variable OT^*_{it}, which determines the observed overtime hours OT_{it} for individual i in year t:

$$OT_{it} = \begin{cases} OT^*_{it} & \text{if } OT^*_{it} > 0 \\ 0 & \text{if } OT^*_{it} \leq 0 \end{cases} \tag{5.1}$$

Following the strand of literature, we first estimate Tobit models. At a first glance, Tobit models seems to be an appropriate model for our outcome variable which is censored at zero hours. A corresponding Tobit model reads as follows:

$$OT^*_{it} = X_{it}\beta + \epsilon_{it} \tag{5.2}$$

X_{it} denotes the vector which includes both the variables of interest, i.e. for works councils and collective agreement status, and further covariates, while β is the vector of regression coefficients. Lastly, ϵ_{it} denotes the usual error term.

Although Tobit models are widely used, it is possible that they are too restrictive. In the likelihood function of the Tobit model, the probability of working non-zero hours and the effect on overtime given positive overtime hours are determined by the same set of parameters. Hence, a different magnitude or signs in the two effects cannot be disentangled (Burke, 2009).

We, therefore, estimate a Cragg hurdle model and compare the estimates with the Tobit model. In a hurdle model, the observed overtime hours OT_{it} are modelled by a participation variable w_{it} and the latent variable OT^*_{it}:

$$OT_{it} = w_{it}OT^*_{it} \tag{5.3}$$

The participation w_{it} is observable and binary, either zero if $OT_{it} = 0$ or one if $OT_{it} > 0$, whereas OT_{it}^* is only observable if $w_{it} = 1$. The hurdle model can be estimated by a probit model for the participation decision and by a (truncated) linear regression for the continuous outcome (Burke, 2009). In the context of this study, this corresponds to a probit model for the incidence of overtime work and a truncated regression model for the intensity of overtime work given non-zero overtime. It is worth mentioning that Cragg hurdle models are seldom used in the literature.[78] This is due to the fact that the truncated regression is, by construction, not estimated using the whole sample, which could cause less efficient estimations (Jirjahn, 2008).

5.4 Results

5.4.1 The impact of works councils

In the subsequent analysis of this subsection, we replicate the findings of Gralla et al. (2017) with more recent SOEP data. For comparison purposes, we employ a set of control variables, which corresponds to the study of Gralla et al. (2017). In their study, the authors use three SOEP waves (2001, 2006, 2011) and regress overtime hours on the works council status without controlling for the collective agreement status, since this information was not available in those waves. We follow what the authors have done and for the moment neglect the information on the collective agreement status. Compared to their study, we utilize two further SOEP waves (2016 and 2019) which were not available for use at the time that of their study. Moreover, we first follow Gralla et al. (2017) and estimate Tobit models and, then, compare the results from the Tobit with the Cragg hurdle model. Thereafter, we run the same estimates only for the survey year 2016, since this year constitutes as the baseline of the analysis for the proceeding subsections.

Table 5.2 provides the corresponding regression results. Columns (1) to (4) show the results for all available SOEP waves including the information on works councils. In contrast, the last four columns include the estimates for the survey year 2016. For each sample, we compare the Tobit estimates with the Cragg hurdle estimates. Throughout the analysis, we report average marginal effects, in order to ease economical interpreting. We calculate the probability of non-zero overtime hours ($Pr(OT > 0)$) and the expected value given that overtime is greater than zero ($E(OT|OT > 0)$).

[78] One exemption is the study of Schank and Schnabel (2004). The authors show that the Tobit model is too restrictive to disentangle the effect of the hourly wage on overtime. Due to the importance of controlling for the wage as a determinant of overtime hours, it is somewhat surprising that Cragg hurdle models are not frequently used in this context.

Both the Tobit model and the Cragg hurdle model show that a works council is negatively connected to the probability of working overtime at all, and also to the numbers of overtime hours greater than zero. In Columns (1) and (2), we expand the sample used by Gralla et al. (2017) by including two further SOEP waves and employ the same model, i.e. the Tobit model. Compared to their results, our estimates show a somewhat stronger effect on the incidence of overtime, while the effect on non-zero overtime hours is slightly weaker. However, we do not find substantial differences compared to their study. When comparing the estimates from the two econometric models, we find that the effect on the probability is stronger in the Tobit model, whereas the effect on non-zero overtime hours is stronger in the Cragg hurdle model. This is true for both samples, i.e. all SOEP waves with information on the works council status and the 2016 wave. Regarding economic significance, the estimates reveal important effects. When considering the 2016 wave, column (7) shows that the probability of working overtime is 3.5 percentage points lower, if a works council exists. In turn, the effect of works councils on non-zero overtime hours amounts to nearly half an hour (see column (8)). This equals to a reduction of nearly 20% at the sample mean of 2.6 overtime hours.

Table 5.2: Comparison of Tobit and Cragg hurdle estimates

	All SOEP waves incl. WC				SOEP 2016 wave			
	Tobit		Cragg hurdle		Tobit		Cragg hurdle	
	OT>0	OT|OT>0	OT>0	OT|OT>0	OT>0	OT|OT>0	OT>0	OT|OT>0
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
WC	-0.043^{***}	-0.251^{***}	-0.025^{**}	-0.506^{***}	-0.051^{**}	-0.288^{**}	-0.035^{+}	-0.496^{**}
	(0.008)	(0.044)	(0.009)	(0.088)	(0.016)	(0.091)	(0.019)	(0.188)
Obs.	20984	20984	20984	20984	4401	4401	4401	4401

Source: SOEP. Standard errors in parentheses; $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Average marginal effects are displayed. Full set of controls is used in each specification. Extended results are displayed in Table D.3 in the Appendix.

In Table 5.2, we do not find direct evidence that the Tobit model is too restrictive when considering the effect of works councils. However, with respect to some control variables, we find a statistically significant effect only in one of the two tiers of the hurdle model (cf. Table D.3).[79] For example, this is the case for foreign citizenship which only affects the incidence of overtime, while a place of residence in East Germany only reduces the amount of overtime hours. Restrictions of the Tobit model apparently become clear, if the coefficients in the two tiers have the opposite sign and are statistically significant. In our regressions, this is true for contractual working time, which is negatively associated with the incidence of overtime work, but positively associated with the amount of non-zero overtime hours. If a shorter standard week reflects coverage by collective agreements, such employees would more likely work overtime, but then would work less overtime

[79] A likelihood ratio test clearly confirms that the hurdle model is preferred over the Tobit model (χ^2=348.02).

hours. Hence, we find evidence that the Cragg hurdle model is preferable over the Tobit. This is especially relevant in our context, where contractual working time is an important covariate, through its collinearity with collective agreements (cf. Subsection 5.2.2).

5.4.2 The impact of collective bargaining agreements

After having confirmed the effects of works council, we will now assess the effect of collective agreements on overtime. We do a step-wise regression analysis: Columns (1) and (2) of Table 5.3 corresponds to columns (7) and (8) of Table 5.2 and is for comparative purposes only. In Columns (3) and (4) of Table 5.3 we regress the information on collective agreement coverage on overtime without controlling for the works council status. This yields the sole effect of collective agreements on overtime. Having no collective agreement is the reference category.

Table 5.3: Cragg hurdle estimates for the full sample

	(1) OT>0	(2) OT|OT>0	(3) OT>0	(4) OT|OT>0	(5) OT>0	(6) OT|OT>0
WC	-0.035^{+}	-0.496^{**}			-0.015	-0.332^{+}
	(0.019)	(0.188)			(0.020)	(0.193)
CBA (Ref.: No)						
Firm-level			-0.064^{**}	-0.409^{+}	-0.060^{*}	-0.309
			(0.024)	(0.245)	(0.024)	(0.253)
Affiliated			-0.055^{*}	-0.512^{+}	-0.051^{+}	-0.428
			(0.028)	(0.275)	(0.028)	(0.280)
Industry-level			-0.061^{***}	-0.726^{***}	-0.057^{**}	-0.642^{***}
			(0.019)	(0.187)	(0.019)	(0.191)
Non-tariff			0.008	0.240	0.011	0.303
			(0.031)	(0.306)	(0.031)	(0.308)
Observations	4401	4401	4401	4401	4401	4401

Source: SOEP. Standard errors in parentheses; $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Average marginal effects are displayed. Full set of controls is used in each specification. Extended results are displayed in Table D.4 in the Appendix.

Table 5.3 shows that especially a firm-level, an industry-level agreement as well as an affiliation go along with a decline in overtime probability and hours. The evidence is not clear-cut for categories with a smaller sample size, such as being a non-tariff employee. In columns (5) and (6) we control for both works council and collective agreement status. The works council effects vanishes to a large extend when we control for the collective agreement status. While we still find a weak effect of works councils on non-zero overtime hours, the effect in the Probit tier becomes insignificant. In contrast, most variables reflecting collective agreement status keep statistically significance in the

Probit tier, while a collective agreement at the industry-level also reduces the amount of non-zero overtime hours significantly.[80]

We now restrict the estimation samples to observations who are neither an employee with executive tasks nor non-tariff employees. Thereby, the following analysis does not include employees who are not represented by a works council and employees whose overtime regulation are generally not regulated by collective agreements (cf. Subsection 5.2.3). This restriction enables us to focus on the interdependencies of works councils and collective agreements.

Table 5.4: Cragg hurdle estimates for the main sample

	(1) OT>0	(2) OT|OT>0	(3) OT>0	(4) OT|OT>0	(5) OT>0	(6) OT|OT>0
WC	-0.027	-0.396^{+}			-0.004	-0.317
	(0.023)	(0.215)			(0.024)	(0.223)
CBA (Ref.: No)						
Firm-level			-0.084^{**}	-0.286	-0.083^{**}	-0.198
			(0.028)	(0.268)	(0.028)	(0.275)
Affiliated			-0.079^{*}	-0.129	-0.079^{*}	-0.048
			(0.033)	(0.323)	(0.033)	(0.329)
Industry-level			-0.057^{**}	-0.367^{+}	-0.056^{*}	-0.281
			(0.022)	(0.211)	(0.023)	(0.217)
Observations	3114	3114	3114	3114	3114	3114

Source: SOEP. Standard errors in parentheses; $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Average marginal effects are displayed. Full set of controls is used in each specification. Extended results are displayed in Table D.5 in the Appendix.

In Table 5.4 we repeat the regression analysis for the new sample. Compared to Table 5.3, the new sample is reduced by 30% of the observations and now comprises around 3,100 individuals.[81] The regression results from columns (1) and (2) of Table 5.4 show that the pure effect of works councils, i.e. without controlling for the collective agreement status, on overtime is statistically less robust than before. Of course, this might be due to a shrinkage in the sample size of more than one quarter. However, the magnitude of the coefficients also mitigates to some extend. In contrast to Table 5.3, the collective agreement variables stay significant regarding the probability of working overtime, but most of them become statistically insignificant when estimating the truncated regression

[80]We have also used different specifications, which do not always include the contractual working time as a covariate. As shown in Subsection 5.2.2, collective bargaining rounds lead to a reduction of standard working hours in the last decades. Hence, we observe a negative correlation between collective agreements and contractual working time in the data. Controlling for contractual working time can lead to insignificant effects of the collective agreement variable (Jirjahn, 2008). Here, the estimations are virtually unaffected if we exclude contractual working hours. These results can be provided upon request. Moreover, Subsection 5.4.4 provides an analysis using interaction between contractual working time and works councils differentiated by collective agreement status.

[81]Descriptive statistics for this sample are shown in Table D.2 in the appendix.

in the second tier of the Cragg hurdle model (cf. columns (3) and (4)).[82] Only the dummy variable for the industry-level remains weakly significant on a 10%-level. As in the preceding regression table, variables including information on collective agreements reduce the impact and statistical significance of the works council variable (cf. columns (5) and (6)). Moreover, the variables representing the collective agreement status yield significant coefficients in the Probit tier, but none reaches statistical significance on a common level, when analysing the intensity of non-zero overtime hours.

In sum and in contrast to the studies of Jirjahn (2008) and Zapf (2016), our results show that the collective bargaining status significantly affects overtime, especially the incidence of overtime.

5.4.3 Works council effect by collective agreement

Since the potential multicollinearity between works councils and collective agreement status can explain the fact that one of the variables becomes insignificant, this subsection presents subsample regression results. More specifically, we run regressions for the works council effect separately by collective agreement status. Since collective agreements on the firm-level and affiliation to collective agreements do not have a sufficient number of observations, we focus on a comparison of collective agreements on the industry-level and no collective agreement coverage. Both subsamples include more than 1,000 observations and, thus, enable us to perform a sensible regression analysis.

Table 5.5 shows Cragg hurdle estimates of the works council effect by a collective bargaining regime. According to column (1) there is no statistically significant effect of works councils on the incidence of overtime in the sample of employees covered by an industry-wide collective agreement, although the magnitude of the estimated coefficient is larger than before. Column (2) shows a strong and statistically highly significant effect on non-zero overtime hours indicating that works council predominantly affect the intensity of overtime. In contrast, columns (3) and (4) show the effects for uncovered employees, which are virtually inexistent for both tiers. In sum, we only find a works council effect on the intensity of overtime hours for employees covered by an industry-wide collective agreement, but not for the other categories. However, it seems that works councils only decrease the amount of overtime hours, but do not have sufficient power to prevent overtime arrangements at all. The effect in column (2) is highly statistically as well as economically significant. In fact, the marginal effect corresponds to a reduction of more than one hour of overtime per week.

[82] The standard Tobit model is not able to disentangle these two effects and would provide one significant coefficient for each category. This provides further evidence that the Cragg hurdle model is preferable over the Tobit model.

Table 5.5: Cragg hurdle estimates by collective agreement regime

	Industry-level CBA		No CBA	
	OT>0	OT|OT>0	OT>0	OT|OT>0
	(1)	(2)	(3)	(4)
WC	-0.064	-1.054^{**}	0.002	-0.015
	(0.044)	(0.331)	(0.036)	(0.373)
Observations	1131	1131	1245	1245

Source: SOEP. Standard errors in parentheses; $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Average marginal effects are displayed. Full set of controls is used in each specification.

The estimates in Table 5.5 have to be interpreted as average effects of works councils across all values of contractual working time. Possible interdependencies are considered in the following subsection.

5.4.4 Interactions

The works council effects may differ along the distribution of contractual hours. Gralla et al. (2017) show that a negative effect of works council is not prevalent for shorter standard hours but for longer ones. Firms covered by collective agreements have on average shorter standard working hours than uncovered firms, as sketched in Section 5.2. Hence, a differentiated analysis seems essential.

In the subsequent analysis we include an interaction term for works councils and contractual working time in our regressions. We run these regressions separately for employees with collective agreements on the industry level and without a collective agreement, as in Table 5.5. Additionally, we run this regressions for the main sample, which corresponds to the 3114 observations from Table 5.4. We then estimate three kinds of marginal effects, which cover the two tiers of the Cragg hurdle model and, additionally, the unconditional overall effect. These marginal effects are calculated for different hours of contractual working time.[83] The results are shown graphically in Figure 5.2.

We first summarize the findings for employees with a collectively determined wage at the industry-level, i.e. column 1 of Figure 5.2. With respect to the incidence of overtime ($Pr(OT > 0)$), the works council estimate is positive (though insignificant) for shorter standard hours and becomes negative and significant with increasing numbers of hours. Regarding the intensity of non-zero hours ($E(OT|OT > 0)$), we find a negative, but constant effect of works councils along the distribution of contractual hours. The effect is weakly statistically significant for lower hours and becomes most statistically significant for a working time of around 40 hours per week. For values above 45 hours the significance vanishes due to too few observations, as shown by the confidence inter-

[83]Our calculation of marginal effects takes the interdependencies of interaction variables and thus the critique of Ai and Norton (2003) into account (Williams, 2012).

vals exceeding the zero baseline. Combining these two results reveals the overall effect ($E(OT)$), which is negative throughout the distribution of working hours and becomes significant from 38 hours onwards.

In contrast, when looking at employees without a collective agreement (column 2), we find neither an effect on the probability nor on the intensity along the whole distribution of contractual working hours. Consequently, there is also no effect, when calculating unconditional margins.

As the main sample (column 3) predominantly includes these two subsamples from the preceding discussion, we find marginal effects, which are smaller in magnitude and less robust compared with employees covered by industry-wide collective agreements. Interestingly, margins for the intensity of non-zero hours become more negative for higher standard hours. This finding indicates that one of the smaller subgroups (now shown in the figure) exhibits an increasingly negative effect of works councils along the distribution of hours.[84]

Overall, the results suggest that the previous findings of a negative works council effect are only applicable for individuals with an industry-wide collective agreement, but not for individuals without any collective agreement. Regarding the former, the effect of works councils on overtime incidence is increasingly negative with larger values of contractual hours. In contrast, a negative effect on the amount of overtime is observable for all values of contractual hours. Again, these findings show the importance of a separate view on both dimensions of overtime, as done with Cragg hurdle instead of Tobit models. For employees not covered by an collective agreement, the effect of works councils is virtually inexistent for all dimensions of overtime.

[84]A detailed analysis would go beyond the scope of this study. Moreover, the subgroups have too few observations in order to draw reliable conclusions.

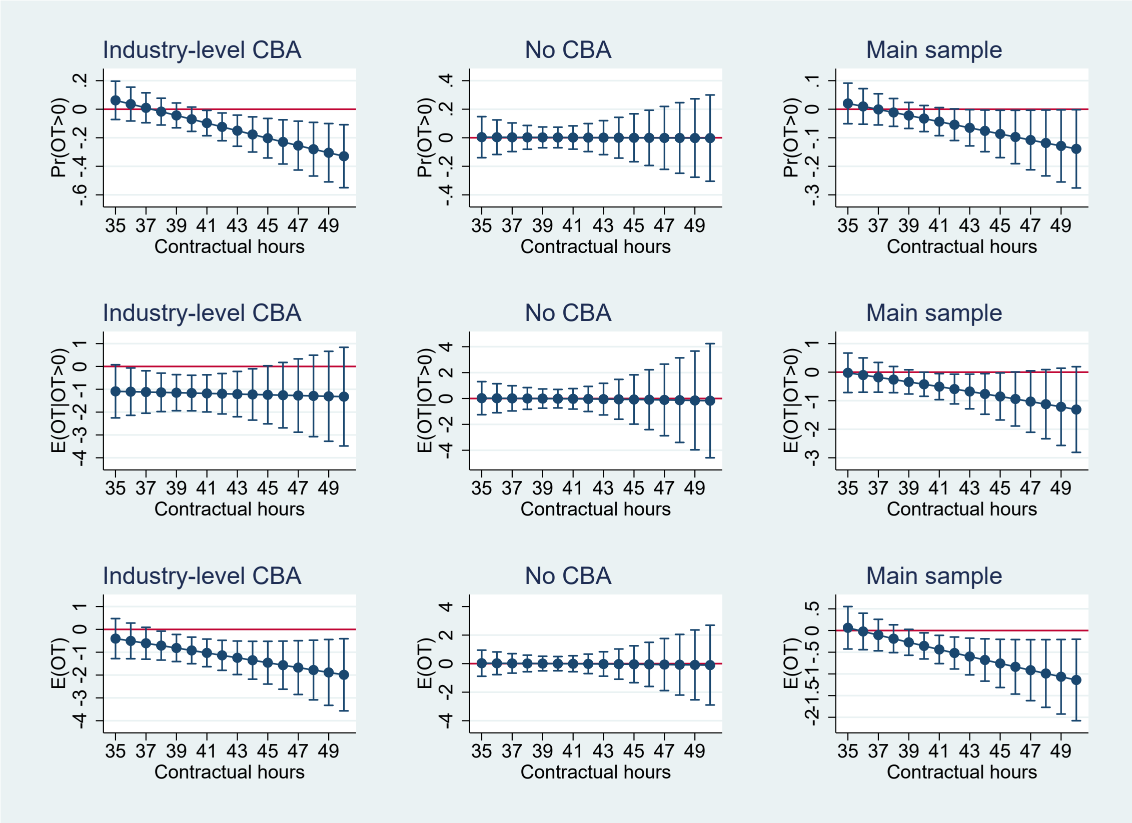

Figure 5.2: Margins plots of the works council effect
Source: SOEP. Average marginal effects and 95% confidence interval are displayed.

5.5 Limitations and conclusions

Using data from the 2016 wave of the SOEP data, we find that the works council effects differ by collective agreement coverage. The availability of only one wave, which includes information on both variables of interest, leads to a relatively small numbers of observation. This goes along with different limitations, which we discuss in the following section. In this section, not all regression results are depicted.

In many studies on the effect of works councils, the question of causality remains unsolved. Using only one SOEP wave in our main analysis, panel data models cannot be employed here. However, the works council status seldom changes, such that fixed effects estimates would be identified by very few changes in the presence of a works council.[85] An alternative way is using instrumental variable regressions. There are several studies which employ an "explorative" instrumental variables strategy (cf. Cornelissen et

[85]This aspect is supported by correlated random effects estimates for all SOEP waves including information on works council existence (i.e. the sample of ca. 21,000 observations in the left half of Table 5.2). The fixed effects estimate is smaller in magnitude and statistically less significant than the estimate for the individual time average of works council existence. This indicates that the found works council effect is not due to variation over time.

al., 2011a), meaning that averages of the potential endogenous variable within specific groups, regions or years are used as an instrument (e.g. Jirjahn and Mohrenweiser, 2021). In preliminary regressions, we calculated the works council density as averages by federal states, industry and survey year. However, meaningful results were only obtained for the five-wave-sample, indicating that the number of observations in our 2016-sample is not sufficiently large for reliable IV estimates (cf. Table D.6).[86]

Potential heterogeneous effects of works councils between collective agreements on the firm-level and on the industry-level could shed light on the underlying channel through which works council can enforce their interests in different institutional settings. For other outcome variables, the literature shows that works councils are only effective in the case of industry-wide but not firm-level agreements (Brändle, 2017, Jirjahn and Mohrenweiser, 2021). In our study, the number of individuals paid according firm-level agreements is not sufficiently high in order to draw general conclusions from those estimates. Nonetheless, corresponding estimates for firm-level agreements indicate no effect of works councils at all. This would be in line with the argument that works councils can better enforce their cause in industry-wide covered firms because bargaining takes place completely externally.

Moreover, Goerke and Schultze (2021) show that employees who are trade union members work fewer overtime hours, based on SOEP data from 1984 to 2015. Due to a potential correlation between trade union membership and collective agreements, further robustness checks seem essential. In recent years, information on trade union membership was asked in the SOEP in 2015 and 2019. Since the 2015 wave is temporally close to our main analysis, we use this wave for a robustness check. In a first step, we carry the information on trade union membership forward from 2015 to 2016, if individuals are included in the data set in both waves and have a tenure larger than one year in 2016 (for a similar approach, see Goerke and Jeworrek (2021)). This data restriction yields a somewhat selective sample but which includes reliable information on individual membership and coverage of collective agreements. We then include this imputed information as an additional control variable to our analysis. According to the regression results in Tables D.7 and D.8, the union membership dummy itself is only statistically significant in some specifications, while the coefficients of the works council dummy and variables for collective agreement are quite similar to the former results and virtually unaffected by the additional covariate. Hence, we do not find evidence that our results are affected by this potential omitted variable.

[86]In principal, it is sensible to use the density for the collective bargaining status as an additional instrumental variable. However, the problem of the limited number of observations would become even more apparent.

Despite the mentioned limitations, this study yields the following important insights into impacts of works councils on overtime. Firstly, the previously observed negative impact of works councils on overtime can be confirmed when using more recent survey data of the SOEP. The magnitudes of our estimates are comparable to the effect size estimated by Gralla et al. (2017) for older waves. Thus, we do not find evidence that the effect changed substantially over time. Secondly, the estimates show that the frequently used Tobit models do not sufficiently disentangle the impact on the incidence and the amount of overtime hours. This applies in particular to the important covariate of contractual hours, where the estimates reveal statistically significant coefficients with a reverse sign. Moreover, throughout the analysis, the estimates show significant coefficients in one but not in the other tier of the hurdle model. This constitutes further evidence for using models separating the incidence and the amount of overtime work. Thirdly, collective agreements themself have a negative impact on overtime, which is mostly in line with theoretical expectations. However, this finding stands in contrast to the few previous findings, which do not provide statistically significant effects. Fourthly, it turns out to be important for the analysis that the 2016 wave of the SOEP includes information on both works councils and collective agreements. Without information on the latter, estimates for the works council dummy would spuriously include variation of collective agreements. Fifthly and most importantly, the works councils effect on overtime depends on the institutional regime, i.e. collective agreement coverage. More specifically, works councils only have an effect on overtime if collective agreements are already negotiated at a higher level and applied to employees in a firm. Due to a limited number of observations, we can claim this only for widespread industry-wide agreements but not for other types of collective agreement. However, these results indicate the importance of the interdependence between both institutional actors, i.e. collective agreement and works councils. Regarding overtime work, these channels were so far unnoticed in the literature. Having said that, the analysis also demonstrates some data issues. A consistent definition of collective agreements over time would allow to exploit the panel structure of the SOEP in our context.

Appendix D

Table D.1: Descriptive statistics - Full sample

	count	mean	sd	min	max
OT dummy	4401	0.64	0.48	0	1
OT hours	4401	2.56	3.46	0	46.7
OT hours>0	2692	4.02	3.60	.2	46.7
WC	4401	0.55	0.50	0	1
CBA					
Firm-level	4401	0.12	0.33	0	1
Affiliated	4401	0.10	0.30	0	1
Industry-level	4401	0.30	0.46	0	1
Non-tariff	4401	0.07	0.25	0	1
No CBA	4401	0.41	0.49	0	1
Contractual hours	4401	39.31	2.53	35	77
Log. hourly wage	4401	2.94	0.44	0	5.121978
Blue collar	4401	0.30	0.46	0	1
Woman	4401	0.29	0.45	0	1
Foreign citizen	4401	0.11	0.31	0	1
Apprenticeship	4401	0.67	0.47	0	1
University degree	4401	0.24	0.43	0	1
Age	4401	43.85	11.55	18	65
Age squared	4401	2056.04	995.30	324	4225
Tenure	4401	11.39	10.66	0	48.583
East Germany	4401	0.20	0.40	0	1
Survey year 2016	4401	1.00	0.00	1	1
Industry					
Agriculture	4401	0.01	0.11	0	1
Energy/Mining	4401	0.02	0.13	0	1
Manufacturing	4401	0.41	0.49	0	1
Construction	4401	0.07	0.26	0	1
Trade	4401	0.14	0.34	0	1
Transport	4401	0.06	0.25	0	1
Bank, Insurance	4401	0.05	0.21	0	1
Services	4401	0.24	0.43	0	1
Firmsize					
Firmsize 1-19	4401	0.17	0.38	0	1
Firmsize 20-199	4401	0.30	0.46	0	1
Firmsize 200-1999	4401	0.25	0.43	0	1
Firmsize 2000+	4401	0.28	0.45	0	1

Source: SOEP.

Table D.2: Descriptive statistics - Main sample

	count	mean	sd	min	max
OT dummy	3114	0.58	0.49	0	1
OT hours	3114	2.09	3.14	0	46.7
OT hours>0	1732	3.61	3.40	.2	46.7
WC	3114	0.53	0.50	0	1
CBA					
Firm-level	3114	0.15	0.35	0	1
Affiliated	3114	0.12	0.32	0	1
Industry-level	3114	0.34	0.48	0	1
Non-tariff	3114	0.00	0.00	0	0
No CBA	3114	0.39	0.49	0	1
Contractual hours	3114	39.24	2.73	35	77
Log. hourly wage	3114	2.80	0.38	0	4.017202
Blue collar	3114	0.41	0.49	0	1
Woman	3114	0.30	0.46	0	1
Foreign citizen	3114	0.12	0.33	0	1
Apprenticeship	3114	0.80	0.40	0	1
University degree	3114	0.09	0.28	0	1
Age	3114	43.69	11.94	18	65
Age squared	3114	2051.78	1021.13	324	4225
Tenure	3114	11.33	10.77	0	48.083
East Germany	3114	0.21	0.41	0	1
Survey year 2016	3114	1.00	0.00	1	1
Industry					
Agriculture	3114	0.01	0.11	0	1
Energy/Mining	3114	0.01	0.12	0	1
Manufacturing	3114	0.42	0.49	0	1
Construction	3114	0.08	0.27	0	1
Trade	3114	0.15	0.36	0	1
Transport	3114	0.07	0.26	0	1
Bank, Insurance	3114	0.04	0.20	0	1
Services	3114	0.21	0.41	0	1
Firmsize					
Firmsize 1-19	3114	0.19	0.40	0	1
Firmsize 20-199	3114	0.32	0.47	0	1
Firmsize 200-1999	3114	0.25	0.43	0	1
Firmsize 2000+	3114	0.23	0.42	0	1

Source: SOEP.

Table D.3: Full regression results of Table 5.2

	All SOEP waves incl. WC				SOEP 2016 wave			
	Tobit		Cragg hurdle		Tobit		Cragg hurdle	
	OT>0	OT|OT>0	OT>0	OT|OT>0	OT>0	OT|OT>0	OT>0	OT|OT>0
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
WC	-0.043^{***}	-0.251^{***}	-0.025^{**}	-0.506^{***}	-0.051^{**}	-0.288^{**}	-0.035^{+}	-0.496^{**}
	(0.008)	(0.044)	(0.009)	(0.088)	(0.016)	(0.091)	(0.019)	(0.188)
Contractual hours	0.002^{*}	0.014^{*}	-0.003^{*}	0.087^{***}	0.000	0.002	-0.005^{+}	0.084^{**}
	(0.001)	(0.007)	(0.001)	(0.013)	(0.002)	(0.014)	(0.003)	(0.029)
Log. hourly wage	0.045^{***}	0.266^{***}	0.058^{***}	0.072	0.061^{**}	0.341^{**}	0.071^{**}	0.204
	(0.009)	(0.053)	(0.011)	(0.106)	(0.019)	(0.110)	(0.023)	(0.228)
Blue collar	-0.096^{***}	-0.565^{***}	-0.114^{***}	-0.237^{**}	-0.064^{***}	-0.360^{***}	-0.079^{***}	-0.143
	(0.007)	(0.043)	(0.008)	(0.089)	(0.016)	(0.091)	(0.018)	(0.194)
Woman	-0.090^{***}	-0.529^{***}	-0.060^{***}	-1.014^{***}	-0.057^{***}	-0.320^{***}	-0.024	-0.866^{***}
	(0.007)	(0.040)	(0.008)	(0.089)	(0.015)	(0.084)	(0.017)	(0.197)
Foreign citizen	-0.091^{***}	-0.534^{***}	-0.115^{***}	-0.046	-0.102^{***}	-0.572^{***}	-0.136^{***}	0.020
	(0.010)	(0.059)	(0.011)	(0.128)	(0.019)	(0.107)	(0.021)	(0.225)
Apprenticeship	0.054^{***}	0.316^{***}	0.085^{***}	-0.378^{**}	-0.023	-0.127	0.017	-0.723^{**}
	(0.011)	(0.063)	(0.012)	(0.139)	(0.022)	(0.124)	(0.025)	(0.254)
University degree	0.138^{***}	0.812^{***}	0.160^{***}	0.346^{*}	0.043^{+}	0.243^{+}	0.072^{*}	-0.116
	(0.013)	(0.074)	(0.014)	(0.156)	(0.026)	(0.146)	(0.030)	(0.288)
Age	0.013^{***}	0.075^{***}	0.009^{***}	0.137^{***}	0.012^{**}	0.070^{**}	0.011^{*}	0.100^{+}
	(0.002)	(0.012)	(0.002)	(0.026)	(0.004)	(0.025)	(0.005)	(0.056)
Age squared	-0.000^{***}	-0.001^{***}	-0.000^{***}	-0.001^{***}	-0.000^{**}	-0.001^{**}	-0.000^{*}	-0.001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
Tenure	-0.001^{***}	-0.007^{***}	-0.001^{+}	-0.014^{**}	-0.001	-0.006	-0.001	-0.013
	(0.000)	(0.002)	(0.000)	(0.004)	(0.001)	(0.004)	(0.001)	(0.009)
East Germany	-0.018^{*}	-0.105^{*}	0.004	-0.464^{***}	-0.046^{**}	-0.258^{**}	-0.020	-0.692^{***}
	(0.007)	(0.042)	(0.008)	(0.088)	(0.016)	(0.089)	(0.018)	(0.207)
Industry (Ref.: Agriculture)								
Energy	-0.098^{**}	-0.578^{**}	-0.076^{+}	-1.029^{*}	-0.214^{***}	-1.330^{**}	-0.199^{**}	-1.530^{+}
	(0.033)	(0.201)	(0.039)	(0.438)	(0.064)	(0.425)	(0.077)	(0.898)
Manufacturing	-0.055^{*}	-0.337^{*}	-0.009	-1.083^{**}	-0.145^{**}	-0.970^{**}	-0.104^{+}	-1.429^{+}
	(0.024)	(0.160)	(0.029)	(0.356)	(0.045)	(0.359)	(0.055)	(0.740)
Construction	-0.041	-0.258	-0.001	-0.885^{*}	-0.101^{*}	-0.714^{+}	-0.072	-1.032
	(0.026)	(0.167)	(0.030)	(0.368)	(0.048)	(0.377)	(0.059)	(0.771)
Trade	-0.033	-0.208	-0.008	-0.681^{+}	-0.129^{**}	-0.878^{*}	-0.110^{+}	-1.053
	(0.025)	(0.164)	(0.030)	(0.364)	(0.047)	(0.368)	(0.057)	(0.762)
Transport	-0.018	-0.116	0.012	-0.757^{*}	-0.124^{*}	-0.847^{*}	-0.096	-1.195
	(0.027)	(0.175)	(0.032)	(0.378)	(0.049)	(0.380)	(0.060)	(0.778)
Bank, Insurance	-0.083^{**}	-0.498^{**}	-0.018	-1.470^{***}	-0.179^{**}	-1.154^{**}	-0.137^{*}	-1.684^{*}
	(0.028)	(0.177)	(0.033)	(0.381)	(0.055)	(0.397)	(0.067)	(0.810)
Services	-0.056^{*}	-0.348^{*}	-0.013	-1.086^{**}	-0.151^{***}	-1.004^{**}	-0.104^{+}	-1.563^{*}
	(0.025)	(0.161)	(0.029)	(0.358)	(0.046)	(0.363)	(0.056)	(0.747)
Firmsize (Ref.: 1-19)								
Firmsize 20-199	0.063^{***}	0.351^{***}	0.060^{***}	0.382^{***}	0.108^{***}	0.551^{***}	0.119^{***}	0.315
	(0.009)	(0.049)	(0.010)	(0.098)	(0.020)	(0.099)	(0.023)	(0.216)
Firmsize 200-1999	0.074^{***}	0.415^{***}	0.081^{***}	0.351^{**}	0.125^{***}	0.647^{***}	0.141^{***}	0.386
	(0.011)	(0.059)	(0.012)	(0.118)	(0.023)	(0.120)	(0.027)	(0.252)
Firmsize 2000+	0.082^{***}	0.466^{***}	0.073^{***}	0.649^{***}	0.141^{***}	0.744^{***}	0.145^{***}	0.702^{*}
	(0.011)	(0.063)	(0.013)	(0.129)	(0.024)	(0.126)	(0.028)	(0.276)
Survey Year (Ref.: 2001)								
Survey Year=2006	0.031^{***}	0.194^{***}	0.047^{***}	-0.036				
	(0.009)	(0.056)	(0.011)	(0.118)				
Survey Year=2011	0.011	0.066	0.027^{*}	-0.193^{+}				
	(0.009)	(0.055)	(0.011)	(0.117)				
Survey Year=2016	-0.031^{***}	-0.182^{***}	0.016	-0.870^{***}				
	(0.009)	(0.053)	(0.011)	(0.111)				
Survey Year=2019	-0.068^{***}	-0.384^{***}	-0.019^{+}	-1.128^{***}				
	(0.009)	(0.053)	(0.011)	(0.112)				
Observations	20984	20984	20984	20984	4401	4401	4401	4401

Source: SOEP. Standard errors in parentheses; $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Average marginal effects are displayed. Full set of controls is used in each specification.

Table D.4: Full regression results of Table 5.3

	(1) OT>0	(2) OT|OT>0	(3) OT>0	(4) OT|OT>0	(5) OT>0	(6) OT|OT>0
Contractual hours	-0.005^{+}	0.084^{**}	-0.006^{*}	0.079^{**}	-0.006^{*}	0.075^{*}
	(0.003)	(0.029)	(0.003)	(0.029)	(0.003)	(0.029)
Log. hourly wage	0.071^{**}	0.204	0.063^{**}	0.061	0.064^{**}	0.097
	(0.023)	(0.228)	(0.023)	(0.228)	(0.023)	(0.228)
Blue collar	-0.079^{***}	−0.143	-0.074^{***}	−0.077	-0.074^{***}	−0.067
	(0.018)	(0.194)	(0.018)	(0.194)	(0.018)	(0.194)
Woman	−0.024	-0.866^{***}	−0.023	-0.861^{***}	−0.023	-0.854^{***}
	(0.017)	(0.197)	(0.017)	(0.195)	(0.017)	(0.195)
Foreign citizen	-0.136^{***}	0.020	-0.132^{***}	0.088	-0.132^{***}	0.086
	(0.021)	(0.225)	(0.021)	(0.226)	(0.021)	(0.226)
Apprenticeship	0.017	-0.723^{**}	0.018	-0.698^{**}	0.019	-0.698^{**}
	(0.025)	(0.254)	(0.025)	(0.252)	(0.025)	(0.252)
University degree	0.072^{*}	−0.116	0.063^{*}	−0.201	0.064^{*}	−0.183
	(0.030)	(0.288)	(0.030)	(0.287)	(0.030)	(0.287)
Age	0.011^{*}	0.100^{+}	0.010^{*}	0.099^{+}	0.010^{*}	0.097^{+}
	(0.005)	(0.056)	(0.005)	(0.056)	(0.005)	(0.056)
Age squared	-0.000^{*}	−0.001	-0.000^{*}	−0.001	-0.000^{*}	−0.001
	(0.000)	(0.001)	(0.000)	(0.001)	(0.000)	(0.001)
Tenure	−0.001	−0.013	−0.000	−0.013	−0.000	−0.012
	(0.001)	(0.009)	(0.001)	(0.009)	(0.001)	(0.009)
East Germany	−0.020	-0.692^{***}	−0.023	-0.747^{***}	−0.023	-0.735^{***}
	(0.018)	(0.207)	(0.018)	(0.207)	(0.018)	(0.206)
Industry (Ref.: Agriculture)						
Energy	-0.199^{**}	-1.530^{+}	-0.187^{*}	-1.538^{+}	-0.186^{*}	-1.502^{+}
	(0.077)	(0.898)	(0.077)	(0.889)	(0.077)	(0.889)
Manufacturing	-0.104^{+}	-1.429^{+}	-0.098^{+}	-1.415^{+}	-0.098^{+}	-1.393^{+}
	(0.055)	(0.740)	(0.055)	(0.733)	(0.055)	(0.730)
Construction	−0.072	−1.032	−0.061	−0.980	−0.062	−1.015
	(0.059)	(0.771)	(0.059)	(0.766)	(0.059)	(0.762)
Trade	-0.110^{+}	−1.053	-0.106^{+}	−1.067	-0.108^{+}	−1.085
	(0.057)	(0.762)	(0.057)	(0.755)	(0.057)	(0.751)
Transport	−0.096	−1.195	−0.091	−1.238	−0.091	−1.203
	(0.060)	(0.778)	(0.060)	(0.770)	(0.060)	(0.768)
Bank, Insurance	-0.137^{*}	-1.684^{*}	-0.133^{*}	-1.720^{*}	-0.132^{*}	-1.681^{*}
	(0.067)	(0.810)	(0.067)	(0.800)	(0.067)	(0.799)
Services	-0.104^{+}	-1.563^{*}	-0.096^{+}	-1.527^{*}	-0.097^{+}	-1.524^{*}
	(0.056)	(0.747)	(0.056)	(0.741)	(0.056)	(0.737)
Firmsize (Ref.: 1-19)						
Firmsize 20-199	0.119^{***}	0.315	0.116^{***}	0.256	0.120^{***}	0.305
	(0.023)	(0.216)	(0.022)	(0.218)	(0.023)	(0.213)
Firmsize 200-1999	0.141^{***}	0.386	0.136^{***}	0.279	0.145^{***}	0.425^{+}
	(0.027)	(0.252)	(0.025)	(0.239)	(0.027)	(0.250)
Firmsize 2000+	0.145^{***}	0.702^{*}	0.140^{***}	0.566^{*}	0.150^{***}	0.759^{**}
	(0.028)	(0.276)	(0.025)	(0.254)	(0.028)	(0.277)
WC	-0.035^{+}	-0.496^{**}			−0.015	-0.332^{+}
	(0.019)	(0.188)			(0.020)	(0.193)
CBA (Ref.: No)						
Firm-level			-0.064^{**}	-0.409^{+}	-0.060^{*}	−0.309
			(0.024)	(0.245)	(0.024)	(0.253)
Affiliated			-0.055^{*}	-0.512^{+}	-0.051^{+}	−0.428
			(0.028)	(0.275)	(0.028)	(0.280)
Industry-level			-0.061^{***}	-0.726^{***}	-0.057^{**}	-0.642^{***}
			(0.019)	(0.187)	(0.019)	(0.191)
Non-tariff			0.008	0.240	0.011	0.303
			(0.031)	(0.306)	(0.031)	(0.308)
Observations	4401	4401	4401	4401	4401	4401

Source: SOEP. Standard errors in parentheses; $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Average marginal effects are displayed. Full set of controls is used in each specification.

Table D.5: Full regression results of Table 5.4

	(1) OT>0	(2) OT|OT>0	(3) OT>0	(4) OT|OT>0	(5) OT>0	(6) OT|OT>0
Contractual hours	-0.009^{**}	0.016	-0.009^{**}	0.019	-0.009^{**}	0.014
	(0.003)	(0.034)	(0.003)	(0.035)	(0.003)	(0.035)
Log. hourly wage	0.032	-1.089^{***}	0.037	-1.105^{***}	0.037	-1.073^{***}
	(0.031)	(0.323)	(0.031)	(0.322)	(0.031)	(0.321)
Blue collar	-0.034	0.148	-0.029	0.154	-0.029	0.161
	(0.021)	(0.192)	(0.021)	(0.193)	(0.021)	(0.193)
Woman	0.003	-0.898^{***}	0.005	-0.903^{***}	0.004	-0.899^{***}
	(0.021)	(0.238)	(0.021)	(0.238)	(0.021)	(0.237)
Foreign citizen	-0.113^{***}	-0.016	-0.111^{***}	0.008	-0.111^{***}	0.005
	(0.025)	(0.233)	(0.025)	(0.234)	(0.025)	(0.234)
Apprenticeship	0.030	-0.201	0.029	-0.199	0.029	-0.194
	(0.028)	(0.247)	(0.027)	(0.247)	(0.028)	(0.247)
University degree	0.046	0.331	0.041	0.304	0.041	0.325
	(0.038)	(0.342)	(0.038)	(0.342)	(0.038)	(0.342)
Age	0.011^{+}	0.143^{*}	0.011^{+}	0.144^{*}	0.011^{+}	0.141^{*}
	(0.006)	(0.062)	(0.006)	(0.062)	(0.006)	(0.062)
Age squared	-0.000^{*}	-0.002^{*}	-0.000^{*}	-0.002^{*}	-0.000^{*}	-0.002^{*}
	(0.000)	(0.001)	(0.000)	(0.001)	(0.000)	(0.001)
Tenure	-0.000	-0.013	0.000	-0.013	0.000	-0.012
	(0.001)	(0.011)	(0.001)	(0.011)	(0.001)	(0.011)
East Germany	-0.030	-1.011^{***}	-0.032	-1.063^{***}	-0.032	-1.045^{***}
	(0.023)	(0.261)	(0.023)	(0.264)	(0.023)	(0.263)
Industry (Ref.: Agriculture)						
Energy	-0.244^{*}	-1.249	-0.234^{*}	-1.250	-0.234^{*}	-1.235
	(0.096)	(0.985)	(0.097)	(0.979)	(0.097)	(0.984)
Manufacturing	-0.132^{*}	-0.745	-0.124^{+}	-0.735	-0.124^{+}	-0.733
	(0.064)	(0.666)	(0.064)	(0.664)	(0.064)	(0.665)
Construction	-0.108	-0.364	-0.093	-0.272	-0.093	-0.333
	(0.069)	(0.707)	(0.069)	(0.708)	(0.070)	(0.707)
Trade	-0.132^{*}	-0.755	-0.126^{+}	-0.737	-0.126^{+}	-0.768
	(0.067)	(0.689)	(0.067)	(0.687)	(0.067)	(0.687)
Transport	-0.097	-0.642	-0.089	-0.645	-0.089	-0.635
	(0.070)	(0.712)	(0.070)	(0.709)	(0.070)	(0.711)
Bank, Insurance	-0.136	-1.474^{+}	-0.129	-1.514^{+}	-0.129	-1.478^{+}
	(0.083)	(0.806)	(0.083)	(0.798)	(0.083)	(0.803)
Services	-0.103	-1.227^{+}	-0.093	-1.218^{+}	-0.093	-1.226^{+}
	(0.066)	(0.687)	(0.066)	(0.685)	(0.066)	(0.686)
Firmsize (Ref.: 1-19)						
Firmsize 20-199	0.110^{***}	0.196	0.109^{***}	0.145	0.110^{***}	0.195
	(0.027)	(0.237)	(0.026)	(0.240)	(0.027)	(0.235)
Firmsize 200-1999	0.122^{***}	0.229	0.126^{***}	0.125	0.128^{***}	0.260
	(0.032)	(0.285)	(0.029)	(0.272)	(0.032)	(0.285)
Firmsize 2000+	0.131^{***}	0.549^{+}	0.134^{***}	0.423	0.137^{***}	0.597^{+}
	(0.034)	(0.322)	(0.030)	(0.299)	(0.034)	(0.325)
WC	-0.027	-0.396^{+}			-0.004	-0.317
	(0.023)	(0.215)			(0.024)	(0.223)
CBA (Ref.: No)						
Firm-level			-0.084^{**}	-0.286	-0.083^{**}	-0.198
			(0.028)	(0.268)	(0.028)	(0.275)
Affiliated			-0.079^{*}	-0.129	-0.079^{*}	-0.048
			(0.033)	(0.323)	(0.033)	(0.329)
Industry-level			-0.057^{**}	-0.367^{+}	-0.056^{*}	-0.281
			(0.022)	(0.211)	(0.023)	(0.217)
Observations	3114	3114	3114	3114	3114	3114

Source: SOEP. Standard errors in parentheses; $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Average marginal effects are displayed. Full set of controls is used in each specification.

Table D.6: Instrumental variable results

	(1)	(2)	(3)	(4)
	All SOEP waves incl. WC		SOEP 2016 wave	
	OT>0	OT|OT>0	OT>0	OT|OT>0
OT hours				
Works council	-0.077	-0.450	0.070	0.394
	(0.062)	(0.361)	(0.146)	(0.827)
Works council				
Works council density	0.60***		0.57***	
	(0.031)		(0.077)	
Observations	20,984		4,401	

Source: SOEP. Standard errors in parentheses; $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. IV-Tobit estimates. Average marginal effects are displayed. Full set of controls is used in each specification.

Table D.7: Cragg hurdle estimates for the full sample including (imputed) trade union membership

	(1)	(2)	(3)	(4)	(5)	(6)
	OT>0	OT|OT>0	OT>0	OT|OT>0	OT>0	OT|OT>0
WC	-0.021	-0.499^{*}			-0.003	-0.337
	(0.021)	(0.203)			(0.022)	(0.207)
Imputed TUM	-0.060^{**}	-0.064	-0.053^{*}	-0.013	-0.052^{*}	0.028
	(0.022)	(0.223)	(0.022)	(0.223)	(0.022)	(0.224)
CBA (Ref.: No)						
Firm-level			-0.060^{*}	-0.415	-0.059^{*}	-0.322
			(0.026)	(0.260)	(0.026)	(0.267)
Affiliated			-0.054^{+}	-0.548^{+}	-0.053^{+}	-0.467
			(0.030)	(0.288)	(0.030)	(0.293)
Industry-level			-0.060^{**}	-0.758^{***}	-0.060^{**}	-0.680^{***}
			(0.020)	(0.201)	(0.021)	(0.205)
Non-tariff			0.003	0.144	0.003	0.205
			(0.034)	(0.319)	(0.034)	(0.321)
Observations	3801	3801	3801	3801	3801	3801

Source: SOEP. Standard errors in parentheses; $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Average marginal effects are displayed. Full set of controls is used in each specification. TUM stands for trade union membership.

Table D.8: Cragg hurdle estimates for the main sample including (imputed) trade union membership

	(1)	(2)	(3)	(4)	(5)	(6)
	OT>0	OT|OT>0	OT>0	OT|OT>0	OT>0	OT|OT>0
WC	-0.013	-0.421^{+}			0.013	-0.361
	(0.026)	(0.236)			(0.027)	(0.244)
Imputed TUM	-0.042	0.085	-0.034	0.055	-0.036	0.105
	(0.026)	(0.244)	(0.026)	(0.243)	(0.026)	(0.246)
CBA (Ref.: No)						
Firm-level			-0.091^{**}	-0.166	-0.094^{**}	-0.071
			(0.030)	(0.287)	(0.030)	(0.294)
Affiliated			-0.089^{*}	-0.176	-0.092^{**}	-0.087
			(0.035)	(0.341)	(0.036)	(0.347)
Industry-level			-0.063^{**}	-0.337	-0.066^{**}	-0.244
			(0.024)	(0.228)	(0.025)	(0.232)
Observations	2685	2685	2685	2685	2685	2685

Source: SOEP. Standard errors in parentheses; $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Average marginal effects are displayed. Full set of controls is used in each specification. TUM stands for trade union membership.

Bibliography

Adam, Julian B., "Voluntary quits: Do works councils matter? An analysis of the reform of the German Works Constitution Act 2001," *Jahrbücher für Nationalökonomie und Statistik*, 2019, *239* (1), 67–109.

Addison, John T., Paulino Teixeira, Katalin Evers, and Lutz Bellmann, "Is the erosion thesis overblown? Alignment from without in Germany," *Industrial Relations: A Journal of Economy and Society*, 2016, *55* (3), 415–443.

Ai, Chunrong and Edward C. Norton, "Interaction terms in logit and probit models," *Economics Letters*, 2003, *80* (1), 123–129.

Aina, Carmen, Eliana Baici, and Giorgia Casalone, "Time to degree: Students' abilities, university characteristics or something else? Evidence from Italy," *Education Economics*, 2011, *19* (3), 311–325.

Allen, Jeremiah and Roger Barnsley, "Streams and tiers: The interaction of ability, maturity, and training in systems with age-dependent recursive selection," *Journal of Human Resources*, 1993, *28* (3), 649–659.

Anger, Silke, "Overtime work as a signaling device," *Scottish Journal of Political Economy*, 2008, *55* (2), 167–189.

Arcidiacono, Peter, Patrick Bayer, and Aurel Hizmo, "Beyond signaling and human capital: Education and the revelation of ability," *American Economic Journal: Applied Economics*, 2010, *2* (4), 76–104.

Babcock, Philip and Mindy Marks, "The falling time cost of college: Evidence from half a century of time use data," *Review of Economics and Statistics*, 2011, *93* (2), 468–478.

Bahrs, Michael and Mathias Schumann, "Unlucky to be young? The long-term effects of school starting age on smoking behavior and health," *Journal of Population Economics*, 2020, *33* (2), 555–600.

Bandiera, Oriana, Valentino Larcinese, and Imran Rasul, "Blissful ignorance? A natural experiment on the effect of feedback on students' performance," *Labour Economics*, 2015, *34*, 13–25.

Barkume, Anthony, "The structure of labor costs with overtime work in US jobs," *Industrial and Labor Relations Review*, 2010, *64* (1), 128–142.

Barua, Rashmi and Kevin Lang, "School entry, educational attainment, and quarter of birth: A cautionary tale of a local average treatment effect," *Journal of Human Capital*, 2016, *10* (3), 347–376.

BAuA, "Arbeitszeitreport Deutschland 2016," Bundesanstalt für Arbeitsschutz und Arbeitsmedizin, `https://www.baua.de/DE/Angebote/Publikationen/Berichte/F2398.pdf`, 2016.

Bauer, Thomas and Klaus F. Zimmermann, "Overtime work and overtime compensation in Germany," *Scottish Journal of Political Economy*, 1999, *46* (4), 419–436.

Becker, Gary S., "A theory of the allocation of time," *The Economic Journal*, 1965, *75* (299), 493–517.

Bedard, Kelly, "Human capital versus signaling models: University access and high school dropouts," *Journal of Political Economy*, 2001, *109* (4), 749–775.

_ and Elizabeth Dhuey, "The persistence of early childhood maturity: International evidence of long-run age effects," *Quarterly Journal of Economics*, 2006, *121* (4), 1437–1472.

Behrens, Martin, "Still married after all these years? Union organizing and the role of works councils in German industrial relations," *Industrial and Labor Relations Review*, 2009, *62* (3), 275–293.

_ , "Unterscheiden sich Mitglieder von Betriebs- und Personalräten vom Rest der Belegschaften?," *Industrielle Beziehungen/The German Journal of Industrial Relations*, 2009, *16* (4), 303–326.

Bell, David N.F., Alassandro Gaj, Robert A. Hart, Olaf Hübler, and Wolfgang Schwerdt, "Unpaid overtime working in Germany and the UK," Working Paper, `https://ftp.zew.de/pub/zew-docs/div/wwheworkshop/hart.pdf`, 2001.

_ and Robert A. Hart, "Working time in Great Britain, 1975–1994: Evidence from the New Earnings Survey panel data," *Journal of the Royal Statistical Society*, 1998, *161* (3), 327–348.

_ and _ , "Unpaid work," *Economica*, 1999, *66* (262), 271–290.

_ and _ , "The decline of overtime working in Britain," *IZA Discussion Paper No. 12651*, 2019.

_ , Steffen Otterbach, and Alfonso Sousa-Poza, "Work hours constraints and health," *Annals of Economics and Statistics*, 2012, *105/106*, 35–54.

Bell, Linda A. and Richard B. Freeman, "The incentive for working hard: Explaining hours worked differences in the US and Germany," *Labour Economics*, 2001, *8* (2), 181–202.

Berglund, Tomas and Bengt Furåker, "Employment protection regulation, trade unions and tenure of employment: An analysis in 23 European countries," *Industrial Relations Journal*, 2016, *47* (5-6), 492–512.

Berniell, Inés and Jan Bietenbeck, "The effect of working hours on health," *Economics & Human Biology*, 2020, *39*, 100901.

Betts, Julian R., "What do students know about wages? Evidence from a survey of undergraduates," *Journal of Human Resources*, 1996, *31* (1), 27–56.

Bishop, John H. and Ludger Wößmann, "Institutional effects in a simple model of educational production," *Education Economics*, 2004, *12* (1), 17–38.

Bispinck, Reinhard, "Arbeitszeit – Was bietet der tarifvertragliche Instrumentenkoffer," *Elemente qualitativer Tarifpolitik Nr. 82*, 2016.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes, "Too young to leave the nest? The effects of school starting age," *Review of Economics and Statistics*, 2011, *93* (2), 455–467.

Blasi, Maria Di, Alessandro Giardina, Cecilia Giordano, Gianluca Lo Coco, Crispino Tosto, Joel Billieux, and Adriano Schimmenti, "Problematic video game use as an emotional coping strategy: Evidence from a sample of MMORPG gamers," *Journal of Behavioral Addictions*, 2019, *8* (1), 25–34.

Block, Jörn, Laszlo Goerke, José María Millán, and Concepción Román, "Family employees and absenteeism," *Economics Letters*, 2014, *123* (1), 94–99.

Blossfeld, Hans-Peter, Hans-Günther Roßbach, and Jutta von Maurice (Editors), "Education as a lifelong process – The German National Educational Panel Study (NEPS)," *Zeitschrift für Erziehungswissenschaft*, 2011, *14.*

Bonaccolto-Töpfer, Marina and Claus Schnabel, "Is there a union wage premium in Germany and which workers benefit most?," *IZA Discussion Paper No. 15844*, 2023.

Bonesrønning, Hans and Leiv Opstad, "Can student effort be manipulated? Does it matter?," *Applied Economics*, 2015, *47* (15), 1511–1524.

Bonnard, Claire, Jean-François Giret, and Marielle Lambert-Le Mener, "Educational intentions, cognitive skills and earnings expectations of French undergraduates," *Applied Economics Letters*, 2014, *21* (18), 1293–1296.

Bossler, Mario, "The rise in orientation at collective bargaining without a formal contract," *Industrial Relations: A Journal of Economy and Society*, 2019, *58* (1), 17–45.

Botelho, Anabela and Ligia Costa Pinto, "Students' expectations of the economic returns to college education: Results of a controlled experiment," *Economics of Education Review*, 2004, *23* (6), 645–653.

Bound, John and David A. Jaeger, "Do compulsory school attendance laws alone explain the association between quarter of birth and earnings?," *Research in Labor Economics*, 2000, *19*, 83–108.

__, Michael F. Lovenheim, and Sarah Turner, "Increasing time to baccalaureate degree in the United States," *Education Finance and Policy*, 2012, *7* (4), 375–424.

Brändle, Tobias, "Flexible collective bargaining agreements: Still a moderating effect on works council behaviour?," *Managerial and Decision Economics*, 2017, *38* (8), 1137–1153.

Bratti, Massimiliano and Stefano Staffolani, "Student time allocation and educational production functions," *Annals of Economics and Statistics*, 2013, *111/112*, 103–140.

Brewer, Dominic J. and Patrick J. McEwan, *Economics of education*, Elsevier, 2010.

Brodaty, Thomas, Robert J. Gary-Bobo, and Ana Prieto, "Do risk aversion and wages explain educational choices?," *Journal of Public Economics*, 2014, *117*, 125–148.

Brunello, Giorgio and Rudolf Winter-Ebmer, "Why do students expect to stay longer in college? Evidence from Europe," *Economics Letters*, 2003, *80* (2), 247–253.

_ , Claudio Lucifora, and Rudolf Winter-Ebmer, "The wage expectations of European business and economics students," *Journal of Human Resources*, 2004, *39* (4), 1116–1142.

Buckles, Kasey S. and Daniel M. Hungerman, "Season of birth and later outcomes: Old questions, new answers," *Review of Economics and Statistics*, 2013, *95* (3), 711–724.

Burke, William J., "Fitting and interpreting Cragg's tobit alternative using Stata," *The Stata Journal*, 2009, *9* (4), 584–592.

Cameron, A. Colin and Pravin K. Trivedi, *Microeconometrics: Methods and applications*, Cambridge University Press, 2005.

Card, David, "Earnings, schooling, and ability revisited," in Solomon Polachek, ed., *Research in Labor Economics Vol. 14*, 1995, pp. 23–48.

Chadi, Adrian and Laszlo Goerke, "Missing at work – Sickness-related absence and subsequent job mobility markets," *Journal of Economic Behavior & Organization*, 2018, *153*, 153–176.

_ and Marco de Pinto, "Selecting successful students? Undergraduate grades as an admission criterion," *Applied Economics*, 2018, *50* (28), 3089–3105.

_ , _ , and Gabriel Schultze, "Young, gifted and lazy? The role of ability and labor market prospects in student effort decisions," *Economics of Education Review*, 2019, *72*, 66–79.

Chevalier, Arnaud, Colm Harmon, Ian Walker, and Yu Zhu, "Does education raise productivity, or just reflect it?," *Economic Journal*, 2004, *114* (499), F499–F517.

_ , Peter Dolton, and Melanie Lührmann, "'Making it count': Incentives, student effort and performance," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 2018, *181* (2), 323–349.

Chiang, Hui-Yu, "The effect of fringe benefits on the paid overtime hours in Japan," *Journal of Applied Business and Economics*, 2012, *13* (3), 123–136.

Cobley, Stephen, Joseph Baker, Nick Wattie, and Jim McKenna, "Annual age-grouping and athlete development: A meta-analytical review of relative age effects in sport," *Sports Medicine*, 2009, *39* (3), 235–256.

Cordero, José M., Víctor Cristóbal, and Daniel Santín, "Causal inference on education policies: A survey of empirical studies using PISA, TIMSS and PIRLS," *Journal of Economic Surveys*, 2018, *32* (3), 878–915.

Cornelissen, Thomas, John S. Heywood, and Uwe Jirjahn, "Performance pay, risk attitudes and job satisfaction," *Labour Economics*, 2011, *18* (2), 229–239.

_ , Oliver Himmler, and Tobias Koenig, "Perceived unfairness in CEO compensation and work morale," *Economics Letters*, 2011, *110* (1), 45–48.

Currie, Candace, Saoirse Nic Gabhainn, Emmanuelle Godeau, Chris Roberts, Rebecca Smith, Dorothy Currie, Will Picket, Matthias Richter, Antony Morgan, and Vivian Barnekow, *Inequalities in young people's health: HBSC international report from the 2005/2006 survey*, World Health Organization, 2008.

Currie, Janet and Hannes Schwandt, "Within-mother analysis of seasonal patterns in health at birth," *Proceedings of the National Academy of Sciences*, 2013, *110* (30), 12265–12270.

Cygan-Rehm, Kamila and Christoph Wunder, "Do working hours affect health? Evidence from statutory workweek regulations in Germany," *Labour Economics*, 2018, *53*, 162–171.

Datar, Ashlesha, "Does delaying kindergarten entrance give children a head start?," *Economics of Education Review*, 2006, *25* (1), 43–62.

De Fraja, Gianni and Pedro Landeras, "Could do better: The effectiveness of incentives and competition in schools," *Journal of Public Economics*, 2006, *90* (1), 189–213.

_ , Tania Oliveira, and Luisa Zanchi, "Must try harder: Evaluating the role of effort in educational attainment," *Review of Economics and Statistics*, 2010, *92* (3), 577–597.

Delaney, Liam, Colm Harmon, and Martin Ryan, "The role of noncognitive traits in undergraduate study behaviours," *Economics of Education Review*, 2013, *32*, 181–195.

DGB, "So viele Überstunden machen die Deutschen," DGB, `http://www.dgb.de/-/E0s`, 2018.

Di Pietro, Giorgio and Andrea Cutillo, "Degree flexibility and university drop-out: The Italian experience," *Economics of Education Review*, 2008, *27* (5), 546–555.

Dolton, Peter, Oscar D. Marcenaro, and Lucia Navarro, "The effective use of student time: A stochastic frontier production function case study," *Economics of Education Review*, 2003, *22* (6), 547–560.

Easterlin, Richard A., "Income and happiness: Towards a unified theory," *Economic Journal*, 2001, *111* (473), 465–484.

Ehlers, Tim and Robert Schwager, "Honest grading, grade inflation, and reputation," *CESifo Economic Studies*, 2016, *62* (3), 506–521.

Elder, Todd E. and Darren H. Lubotsky, "Kindergarten entrance age and children's achievement: Impacts of state policies, family background, and peers," *Journal of Human Resources*, 2009, *44* (3), 641–683.

Ellguth, Peter and Susanne Kohaut, "Tarifbindung und betriebliche Interessenvertretung: Aktuelle Ergebnisse aus dem IAB-Betriebspanel 2019," *WSI-Mitteilungen*, 2020, *73* (4), 278–285.

_ and _ , "Tarifbindung und betriebliche Interessenvertretung: Aktuelle Ergebnisse aus dem IAB-Betriebspanel 2020," *WSI-Mitteilungen*, 2021, *74* (4), 306–314.

European Parliament, "Indicators of job quality in the European Union," `https://www.europarl.europa.eu/document/activities/cont/201107/20110718ATT24284/20110718ATT24284EN.pdf`, 2009.

Ewers, Mara and Florian Zimmermann, "Image and misreporting," *Journal of the European Economic Association*, 2015, *13* (2), 363–380.

Fan, Elliott, Jin-Tan Liu, and Yen-Chien Chen, "Is the quarter of birth endogenous? New evidence from Taiwan, the US, and Indonesia," *Oxford Bulletin of Economics and Statistics*, 2017, *79* (6), 1087–1124.

Felder-Puig, Rosemarie, Robert Griebler, Oddrun Samdal, Matthew A. King, John Freeman, and Wolfgang Duer, "Does the school performance variable used in the international Health Behavior in School-Aged Children (HBSC) study reflect students' school grades?," *Journal of School Health*, 2012, *82* (9), 404–409.

Fenzel, L. Mickey, "The effect of relative age on self-esteem, role strain, GPA, and anxiety," *Journal of Early Adolescence*, 1992, *12* (3), 253–266.

Fitzenberger, Bernd and Katrin Sommerfeld, "A sequential decomposition of the drop in collective bargaining coverage," *Jahrbücher für Nationalökonomie und Statistik*, 2016, *236* (1), 37–69.

_ , Isabelle Haggeney, and Michaela Ernst, "Wer ist noch Mitglied in Gewerkschaften? Eine Panelanalyse für Westdeutschland," *Journal of Applied Social Science Studies (Schmollers Jahrbuch)*, 1999, *119* (2), 223–263.

Frazis, Harley, "Human capital, signaling, and the pattern of returns to education," *Oxford Economic Papers*, 2002, *54* (2), 298–320.

Fredriksson, Peter and Björn Öckert, "Life-cycle effects of age at school start," *Economic Journal*, 2014, *124*, 977–1004.

Freeman, Richard B. and Edward P. Lazear, "An economic analysis of works councils," in Joel Rogers and Wolfgang Streeck, eds., *Works councils: Consultation, representation, and cooperation in industrial relations*, University of Chicago Press, 1995, pp. 27–52.

Frey, Bruno S., Christine Benesch, and Alois Stutzer, "Does watching TV make us happy?," *Journal of Economic Psychology*, 2007, *28* (3), 283–313.

Fuest, Clemens and Bernd Huber, "Is tax progression really good for employment? A model with endogenous hours of work," *Labour Economics*, 2000, *7* (1), 79–93.

Fumarco, Luca and Gabriel Schultze, "Does relative age make Jack a dull student? Evidence from students' schoolwork and playtime," *Education Economics*, 2020, *28* (6), 647–670.

_ and Stijn Baert, "Relative age effect on European adolescents' social network," *Journal of Economic Behavior & Organization*, 2019, *168*, 318–337.

_ , _ , and Francesco Sarracino, "Younger, dissatisfied, and unhealthy – Relative age in adolescence," *Economics & Human Biology*, 2020, *37*, 100858.

Garibaldi, Pietro, Francesco Giavazzi, Andrea Ichino, and Enrico Rettore, "College cost and time to complete a degree: Evidence from tuition discontinuities," *Review of Economics and Statistics*, 2012, *94* (3), 699–711.

Gary-Bobo, Robert J. and Alain Trannoy, "Efficient tuition fees and examinations," *Journal of the European Economic Association*, 2008, *6* (6), 1211–1243.

Girard, Yann, Florian Hett, and Daniel Schunk, "How individual characteristics shape the structure of social networks," *Journal of Economic Behavior & Organization*, 2015, *115*, 197–216.

Goebel, Jan, Markus M. Grabka, Stefan Liebig, Martin Kroh, David Richter, Carsten Schröder, and Jürgen Schupp, "The German Socio-economic Panel (SOEP)," *Jahrbücher für Nationalökonomie und Statistik*, 2019, *239* (2), 345–360.

Goerke, Laszlo and Gabriel Schultze, "Trade union membership and overtime in Germany," 2021, *Mimeo.*

_ and Markus Pannenberg, "Norm-based trade union membership: Evidence for Germany," *German Economic Review*, 2004, *5* (4), 481–504.

_ and _ , "Trade union membership and dismissals," *Labour Economics*, 2011, *18* (6), 810–821.

_ and _ , "Wage determination in the shadow of the law: The case of works councilors in Germany," *Economic and Industrial Democracy*, 2022.

_ and Sabrina Jeworrek, "Paid vacation use: The role of works councils," *Economic and Industrial Democracy*, 2021, *42* (3), 473–503.

_ and Yue Huang, "Job satisfaction and trade union membership in Germany," *Labour Economics*, 2022, *78*, 102238.

Goldin, Claudia, Lawrence F. Katz, and Ilyana Kuziemko, "The homecoming of American college women: The reversal of the college gender gap," *Journal of Economic Perspectives*, 2006, *20* (4), 133–156.

Gralla, Rafael, Kornelius Kraft, and Stanislav Volgushev, "The effects of works councils on overtime hours," *Scottish Journal of Political Economy*, 2017, *64* (2), 143–168.

Grant, Darren, "Grades as information," *Economics of Education Review*, 2007, *26* (2), 201–214.

Grave, Barbara S., "The effect of student time allocation on academic achievement," *Education Economics*, 2011, *19* (3), 291–310.

Groot, Wim, "Adaptation and scale of reference bias in self-assessments of quality of life," *Journal of Health Economics*, 2000, *19* (3), 403–420.

Gunnes, Trude, Lars J. Kirkebøen, and Marte Rønning, "Financial incentives and study duration in higher education," *Labour Economics*, 2013, *25*, 1–11.

Gürtzgen, Nicole, "Estimating the wage premium of collective wage contracts: Evidence from longitudinal linked employer–employee data," *Industrial Relations: A Journal of Economy and Society*, 2016, *55* (2), 294–322.

Gyimah-Brempong, Kwabena and Anthony O. Gyapong, "Characteristics of education production functions: An application of canonical regression analysis," *Economics of Education Review*, 1991, *10* (1), 7–17.

Haapasalo, Ilona, Raili Välimaa, and Lasse Kannas, "How comprehensive school students perceive their psychosocial school environment," *Scandinavian Journal of Educational Research*, 2010, *54* (2), 133–150.

Hart, Robert A., *The economics of overtime working*, Cambridge University Press, 2004.

_ **and Yue Ma**, "Wage–hours contracts, overtime working and premium pay," *Labour Economics*, 2010, *17* (1), 170–179.

Heywood, John S. and Uwe Jirjahn, "Family-friendly practices and worker representation in Germany," *Industrial Relations: A Journal of Economy and Society*, 2009, *48* (1), 121–145.

Hübler, Olaf, "Individual overtime functions with double correction for selectivity bias," *Economics Letters*, 1989, *29* (1), 87–90.

_ **and Uwe Jirjahn**, "Works councils and collective bargaining in Germany: The impact on productivity and wages," *Scottish Journal of Political Economy*, 2003, *50* (4), 471–491.

Hunt, Jennifer, "Has work-sharing worked in Germany?," *Quarterly Journal of Economics*, 1999, *114* (1), 117–148.

_ , "Flexible work time in Germany: Do workers like it and how have employers exploited it over the cycle?," *Perspektiven der Wirtschaftspolitik*, 2013, *14* (1-2), 67–98.

Huntington-Klein, Nick, "Subjective and projected returns to education," *Journal of Economic Behavior & Organization*, 2015, *117*, 10–25.

IAB, "Durchschnittliche Arbeitszeit und ihre Komponenten in Deutschland," IAB, http://doku.iab.de/arbeitsmarktdaten/AZ_Komponenten.xlsx, 2020.

Ichino, Andrea and Regina T. Riphahn, "The effect of employment protection on worker effort: Absenteeism during and after probation," *Journal of the European Economic Association*, 2005, *3* (1), 120–143.

Jensen, Robert, "The (perceived) returns to education and the demand for schooling," *Quarterly Journal of Economics*, 2010, *125* (2), 515–548.

Jirjahn, Uwe, "On the determinants of shift work and overtime work: Evidence from German establishment data," *British Journal of Industrial Relations*, 2008, *46* (1), 133–168.

_ , "Works councils and collective bargaining in Germany: A simple theoretical extension to reconcile conflicting empirical findings," *Journal of Institutional and Theoretical Economics*, 2017, *173* (2), 322–346.

_ **and Jens Mohrenweiser**, "Works councils and organizational gender policies in Germany," *British Journal of Industrial Relations*, 2021, *59* (4), 1020–1048.

Joncas, Marc and Pierre Foy, "Sample design in TIMSS and PIRLS," in Michael O. Martin and Ina V. S. Mullis, eds., *Methods and procedures in TIMSS and PIRLS 2011*, TIMSS & PIRLS International Study Center, Boston College. Available at http://timssandpirls.bc.edu/methods/pdf/TP_Sampling_Design.pdf, 2013.

Jung, Sven and Claus Schnabel, "Paying more than necessary? The wage cushion in Germany," *Labour*, 2011, *25* (2), 182–197.

Kahn, Lisa B., "The long-term labor market consequences of graduating from college in a bad economy," *Labour Economics*, 2010, *17* (2), 303–316.

Kalenkoski, Charlene Marie and Sabrina Wulff Pabilonia, "Parental transfers, student achievement, and the labor supply of college students," *Journal of Population Economics*, 2010, *23* (2), 469–496.

Kalwij, Adriaan S. and Mary Gregory, "A panel data analysis of the effects of wages, standard hours and unionization on paid overtime work in Britain," *Journal of the Royal Statistical Society*, 2005, *168* (1), 207–231.

Knaus, Michael and Steffen Otterbach, "Work hour mismatch and job mobility: Adjustment channels and resolution rates," *Economic Inquiry*, 2019, *57* (1), 227–242.

Kohaut, Susanne and Claus Schnabel, "Tarifliche Öffnungsklauseln – Verbreitung, Inanspruchnahme und Bedeutung," *Sozialer Fortschritt*, 2007, *56* (2), 33–40.

Kölling, Arnd, "Überstunden als dauerhafte Anpassung der Arbeitsnachfrage: Theoretische Überlegungen und empirische Überprüfung anhand von Daten aus dem IAB-Betriebspanel," in Jürgen Kühl, Manfred Lahner, and Joachim Wagner, eds., *Die Nachfrageseite des Arbeitsmarktes*, BeitrAB 204, 1997, pp. 205–226.

— **and Karen Lehmann**, "Arbeitszeitregelungen und Tarifbindung," in Lutz Bellmann, Knut Gerlach, Olaf Hübler, and Wolfgang Meyer, eds., *Beschäftigungseffekte betrieblicher Arbeitszeitgestaltung*, BeitrAB 251, 2001, pp. 105–133.

Kraft, Kornelius and Julia Lang, "The causes and consequences of adopting a works council," *Jahrbücher für Nationalökonomie und Statistik*, 2008, *228* (5-6), 512–532.

Kramarz, Francis, Pierre Cahuc, Bruno Crépon, Thorsten Schank, Oskar Nordström Skans, Gijsbert van Lomwel, and André Zylberberg, "Labour market effects of work-sharing arrangements in Europe," in Tito Boeri, Michael Burda, and Francis Kramarz, eds., *Working hours and job sharing in the EU and USA. Are Europeans lazy? Or Americans crazy?*, Oxford University Press, 2008, pp. 103–241.

Kriechel, Ben, Samuel Muehlemann, Harald Pfeifer, and Miriam Schütte, "Works councils, collective bargaining, and apprenticeship training – Evidence from German firms," *Industrial Relations: A Journal of Economy and Society*, 2014, *53* (2), 199–222.

Kroch, Eugene A. and Kriss Sjoblom, "Schooling as human capital or a signal: Some evidence," *Journal of Human Resources*, 1994, *29* (1), 156–180.

Krohn, Gregory A. and Catherine M. O'Connor, "Student effort and performance over the semester," *Journal of Economic Education*, 2005, *36* (1), 3–28.

Lackner, Mario, Rudi Stracke, Uwe Sunde, and Rudolf Winter-Ebmer, "Are competitors forward looking in strategic interactions? Evidence from the field," *IZA Discussion Paper No. 9564*, 2015.

Lee, Jungmin and Yong-Kwan Lee, "Can working hour reduction save workers?," *Labour Economics*, 2016, *40*, 25–36.

Lepinteur, Anthony, "The shorter workweek and worker wellbeing: Evidence from Portugal and France," *Labour Economics*, 2019, *58*, 204–220.

Leuven, Edwin, Hessel Oosterbeek, and Bas van der Klaauw, "The effect of financial rewards on students' achievement: Evidence from a randomized experiment," *Journal of the European Economic Association*, 2010, *8* (6), 1243–1265.

Light, Audrey and Wayne Strayer, "Determinants of college completion: School quality or student ability?," *Journal of Human Resources*, 2000, *35* (2), 299–332.

Löfgren, Curt and Henry Ohlsson, "What determines when undergraduates complete their theses? Evidence from two economics departments," *Economics of Education Review*, 1999, *18* (1), 79–88.

Lokshin, Michael and Sergiy Radyakin, "Month of birth and children's health in India," *Journal of Human Resources*, 2012, *47* (1), 174–203.

Mankiw, N. Gregory, "Imperfect competition and the Keynesian cross," *Economics Letters*, 1988, *26* (1), 7–13.

Metcalfe, Robert, Simon Burgess, and Steven Proud, "Student effort and educational attainment: Using the England football team to identify the education production function," *CMPO Working Paper No. 11/276*, 2011.

Michaelis, Jochen and Benjamin Schwanebeck, "Examination rules and student effort," *Economics Letters*, 2016, *145*, 65–68.

Miller, Patrick and Martin Plant, "Truancy and perceived school performance: An alcohol and drug study of UK teenagers," *Alcohol and Alcoholism*, 1999, *34* (6), 886–893.

Miller, Paul and Charles Mulvey, "Trade unions and the distribution of paid overtime," *Journal of Industrial Relations*, 1991, *33* (2), 220–233.

Moksnes, Unni K., Audhild Løhre, Monica Lillefjell, Don G. Byrne, and Gørill Haugan, "The association between school stress, life satisfaction and depressive symptoms in adolescents: Life satisfaction as a potential mediator," *Social Indicators Research*, 2016, *125* (1), 339–357.

Mühlenweg, Andrea M. and Patrick A. Puhani, "The evolution of the school-entry age effect in a school tracking system," *Journal of Human Resources*, 2010, *45* (2), 407–438.

Müller, Torsten and Thorsten Schulten, "Germany: Parallel universes of collective bargaining," in Torsten Müller, Kurt Vandaele, and Jeremy Waddington, eds., *Collective bargaining in Europe: Towards an endgame*, European Trade Union Institute (ETUI), 2019, pp. 239–265.

Mundlak, Yair, "On the pooling of time series and cross section data," *Econometrica*, 1978, *46* (1), 69–85.

Musch, Jochen and Roy Hay, "The relative age effect in soccer: Cross-cultural evidence for a systematic discrimination against children born late in the competition year," *Sociology of Sport Journal*, 1999, *16* (1), 54–64.

Nash, Roy, "Class, 'ability' and attainment: A problem for the sociology of education," *British Journal of Sociology of Education*, 2001, *22* (2), 189–202.

Non, Arjan and Dirk Tempelaar, "Time preferences, study effort, and academic performance," *Economics of Education Review*, 2016, *54*, 36–61.

Oberfichtner, Michael and Claus Schnabel, "The German model of industrial relations: (Where) does it still exist?," *Jahrbücher für Nationalökonomie und Statistik*, 2019, *239* (1), 5–37.

Odermatt, Reto and Alois Stutzer, "(Mis-)predicted subjective well-being following life events," *Journal of the European Economic Association*, 2019, *17* (1), 245–283.

OECD, "Education at a Glance 2013: OECD Indicators," *OECD Publishing*, 2013.

—, "OECD Employment Outlook 2014," OECD Publishing, Paris. doi:10.1787/empl_outlook-2014-en, 2014.

—, "OECD Employment Outlook 2018," OECD Publishing, Paris. doi:10.1787/empl_outlook-2018-en, 2018.

Oettinger, Gerald S., "The effect of nonlinear incentives on performance: Evidence from Econ 101," *Review of Economics and Statistics*, 2002, *84* (3), 509–517.

Oswald, Andrew J. and Ian Walker, "Rethinking labor supply: Contract theory and unions," *The University of Sydney, Working Papers in Economics No. 207*, 1994.

Pannenberg, Markus, "Long-term effects of unpaid overtime," *Scottish Journal of Political Economy*, 2005, *52* (2), 177–193.

— **and Gert G. Wagner**, "Umfang und Kompensation von Überstunden – Eine vergleichende Analyse für Westdeutschland und Großbritannien," in Lutz Bellmann, Knut Gerlach, Olaf Hübler, and Wolfgang Meyer, eds., *Beschäftigungseffekte betrieblicher Arbeitszeitgestaltung*, BeitrAB 251, 2001, pp. 73–94.

Pellizzari, Michele and Francesco C. Billari, "The younger, the better? Age-related differences in academic performance at university," *Journal of Population Economics*, 2012, *25* (2), 697–739.

Peña, Pablo A., "Creating winners and losers: Date of birth, relative age in school, and outcomes in childhood and adulthood," *Economics of Education Review*, 2017, *56*, 152–176.

— **and Angela L. Duckworth**, "The effects of relative and absolute age in the measurement of grit from 9th to 12th grade," *Economics of Education Review*, 2018, *66*, 183–190.

Pfeifer, Christian, "Risk aversion and sorting into public sector employment," *German Economic Review*, 2011, *12* (1), 85–99.

Plante, Courtney N., Douglas A. Gentile, Christopher L. Groves, Adam Modlin, and Jorge Blanco-Herrera, "Video games as coping mechanisms in the etiology of video game addiction," *Psychology of Popular Media Culture*, 2019, *8* (4), 385–394.

Plug, Erik J. S., "Season of birth, schooling and earnings," *Journal of Economic Psychology*, 2001, *22* (5), 641–660.

Polachek, Solomon W., Thomas J. Kniesner, and Henrick J. Harwood, "Educational production functions," *Journal of Educational Statistics*, 1978, *3* (3), 209–231.

Ponzo, Michela and Vincenzo Scoppa, "The long-lasting effects of school entry age: Evidence from Italian students," *Journal of Policy Modeling*, 2014, *36* (3), 578–599.

Pusch, Toralf, "Bilanz des gesetzlichen Mindestlohns: Deutliche Lohnerhöhungen, aber auch viele Umgehungen," *Wirtschaftsdienst*, 2018, *98* (4), 252–259.

Reichert, Arndt R., Boris Augurzky, and Harald Tauchmann, "Self-perceived job insecurity and the demand for medical rehabilitation: Does fear of unemployment reduce health care utilization?," *Health Economics*, 2015, *24* (1), 8–25.

Rietveld, Cornelius A. and Dinand Webbink, "On the genetic bias of the quarter of birth instrument," *Economics & Human Biology*, 2016, *21*, 137–146.

Roberts, Chris, John Freeman, Oddrun Samdal, Christina W. Schnohr, Margreet E. de Looze, Saoirse Nic Gabhainn, Ron Iannotti, Mette Rasmussen, and the International HBSC Study Group, "The Health Behaviour in School-aged Children (HBSC) study: Methodological developments and current tensions," *International Journal of Public Health*, 2009, *54* (2), 140–150.

Sajaia, Zurab, "BIOPROBIT: Stata module for bivariate ordered probit regression," *Statistical Software Components*, 2008.

Schank, Thorsten, "Employment effects of longer working hours," *IZA World of Labor*, 2015, *216.*

_ **and Claus Schnabel**, "Betriebliche Determinanten des Überstundeneinsatzes," in Lutz Bellmann and Claus Schnabel, eds., *Betriebliche Arbeitszeitpolitik im Wandel*, BeitrAB 288, 2004, pp. 37–62.

Scheuer, Steen, "The impact of collective agreements on working time in Denmark," *British Journal of Industrial Relations*, 1999, *37* (3), 465–481.

Schmidt, Christoph M. and Klaus F. Zimmermann, "Work characteristics, firm size and wages," *The Review of Economics and Statistics*, 1991, *73* (4), 705–710.

Schneider, Helena and Sandra Vogel, "Tarifbindung der Beschäftigten in Deutschland: Eine Auswertung des Sozioökonomischen Panels," *IW-Report Nr. 15*, 2018.

Schulten, Thorsten, "German collective bargaining – From erosion to revitalisation?," in Martin Behrens and Heiner Dribbusch, eds., *Industrial relations in Germany*, Nomos Verlagsgesellschaft, 2019, pp. 11–30.

Schwandt, Hannes and Amelie Wuppermann, "The youngest get the pill: ADHD misdiagnosis in Germany, its regional correlates and international comparison," *Labour Economics*, 2016, *43*, 72–86.

Shek, Daniel T. and Xiang Li, "Perceived school performance, life satisfaction, and hopelessness: A 4-year longitudinal study of adolescents in Hong Kong," *Social Indicators Research*, 2016, *126* (2), 921–934.

Smith, Kristy L., Patricia L. Weir, Kevin Till, Michael Romann, and Stephen Cobley, "Relative age effects across and within female sport contexts: A systematic review and meta-analysis," *Sports Medicine*, 2018, *48* (6), 1451–1478.

SOEP v33, "Socio-Economic Panel (SOEP)," Version 33, data for years 1984-2016. doi:`10.5684/soep.v33.1i`, 2018.

Sprietsma, Maresa, "Effect of relative age in the first grade of primary school on long-term scholastic results: International comparative evidence using PISA 2003," *Education Economics*, 2010, *18* (1), 1–32.

Stinebrickner, Ralph and Todd R. Stinebrickner, "Working during school and academic performance," *Journal of Labor Economics*, 2003, *21* (2), 473–491.

_ **and** _ , "The causal effect of studying on academic performance," *BE Journal of Economic Analysis & Policy*, 2008, *8* (1).

_ **and** _ , "A major in science? Initial beliefs and final outcomes for college major and dropout," *Review of Economic Studies*, 2014, *81* (1), 426–472.

Stock, James H. and Motohiro Yogo, "Testing for weak instruments in linear IV regression," in Donald W. K. Andrews and James H. Stock, eds., *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, Cambridge University Press, 2005, pp. 80–108.

Torsheim, Torbjorn and Bente Wold, "School-related stress, school support, and somatic complaints: A general population study," *Journal of Adolescent Research*, 2001, *16* (3), 293–303.

Trejo, Stephen J., "Overtime pay, overtime hours, and labor unions," *Journal of Labor Economics*, 1993, *11* (2), 253–278.

van der Klaauw, Bas and Aico van Vuuren, "Job search and academic achievement," *European Economic Review*, 2010, *54* (2), 294–316.

Veliziotis, Michail, "Trade unions and unpaid overtime in Britain," *UWE Economics Working Paper Series No. 1304*, 2013.

Visser, Jelle, "ICTWSS: Database on institutional characteristics of trade unions, wage setting, state intervention and social pacts in 55 countries between 1960 and 2018," *Version 6.1 , November 2019, Amsterdam Institute for Advanced Labour Studies (AIAS), University of Amsterdam*, 2019.

von der Heiden, Juliane M., Beate Braun, Kai W. Müller, and Boris Egloff, "The association between video gaming and psychological functioning," *Frontiers in Psychology*, 2019, *10*, 1731.

Walwei, Ulrich, Lutz Bellmann, and Christoph Bellmann, "The German industrial relations system under pressure: Structure, trends and outcomes," in Daniel Vaughan-Whitehead, Youcef Ghellab, and Rafael M. de Bustillo Llorente, eds., *The New World of Work*, Edward Elgar Publishing, 2021, pp. 216–254.

Webbink, Dinand and Joop Hartog, "Can students predict starting salaries? Yes!," *Economics of Education Review*, 2004, *23* (2), 103–113.

Williams, Richard, "Using the margins command to estimate and interpret adjusted predictions and marginal effects," *The Stata Journal*, 2012, *12* (2), 308–331.

Wolpin, Kenneth I., "Education and screening," *American Economic Review*, 1977, *67* (5), 949–958.

Wolter, Stefan C., "Wage expectations: A comparison of Swiss and US students," *Kyklos*, 2000, *53* (1), 51–69.

Wooldridge, Jeffrey M., *Econometric analysis of cross section and panel data*, MIT Press, 2002.

WSI, "Tarifstatistik: Tarifbindung," Wirtschafts- und Sozialwissenschaftliches Institut, `https://www.wsi.de/de/tarifbindung-15329.htm`, 2022.

Zapf, Ines, "Individual and workplace-specific determinants of paid and unpaid overtime work in Germany," *IAB Discussion Paper No. 15*, 2015.

_ , *Traditionelle und moderne Formen der Arbeitszeitflexibilität. Arbeitsangebots- und -nachfrageseitige Faktoren von Überstunden und Arbeitszeitkonten*, Vol. 361, IAB-Bibliothek, 2016.

_ **and Enzo Weber**, "The role of employer, job and employee characteristics for flexible working time: An empirical analysis of overtime work and flexible working hours' arrangements," *IAB Discussion Paper No. 4*, 2017.