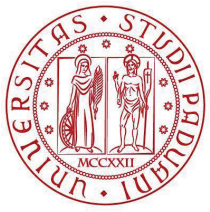


UNIVERSITÀ DEGLI STUDI DI PADOVA

UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Scienze Economiche ed Aziendali "Marco Fanno"

SCUOLA DI DOTTORATO DI RICERCA IN ECONOMIA E MANAGEMENT
CICLO XXIV

ESSAYS IN ECONOMICS OF EDUCATION

Direttore della Scuola: Ch.mo Prof. Giorgio Brunello
Supervisore: Ch.mo Prof. Giorgio Brunello

Dottoranda: Elena Crivellaro

31 Luglio 2012

Per Elisabetta e Carlo

"If we knew what it was we were doing,
it would not be called research, would it?"
Albert Einstein (1879-1955)

Contents

List of Tables

List of Figures

Acknowledgements

First and foremost I wish to express my deepest gratitude to my supervisor, Prof. Giorgio Brunello in Padua. He has supported me throughout my thesis with patience and knowledge whilst allowing me the room to work in my own way. His suggestions have been fundamental in the choice of the research topic and throughout all the stages of this thesis. The expertise and the opportunities he offered me represented a unique value added not only to this work but also to my personal growth. Additionally, he has given me the opportunity of spending two years in London, at the London School of Economics as a visiting PhD student.

I want to thank my other coauthors, Doctor Lorenzo Rocco, Doctor Sandra McNally, Doctor Francesca Cornaglia should be separately thanked. She not only was my coauthor during my visiting at LSE but she has offered me a professional support afterwards, as well.

I am grateful to Prof. Steve Pischke, who "adopted" me during my visiting period at LSE and for his example of dedication and his sincere interest in my ideas and projects.

A special word of gratitude goes to Prof. Cesare Dosi: he supervised my first work as an undergraduate. He kept on supporting me even in the years after my graduation, being for me a reference for all my choices, including the one of enrolling in a Ph.D. in Economics.

It is difficult to describe the importance of the support received from my classmates of the PhD in Economics at the University of Padua and at LSE. Sharing with them the burden of the exams as well as the doubts and the satisfactions of the academic research has contributed in making my PhD a wonderful experience. I am also grateful to all the professors and colleagues in the Economics Department of the University of Padua and London School of Economics for providing a stimulating

environment in which to learn, grow, and enjoy academic research.

I wish to express a special thanks to my family: my parents, my brother and my grandmothers, who unfortunately can not see this work finished. Their unconditional love and encouragement represented the real strength of all my choices. My warmest gratitude goes to my parents for having always supported and sustained me in all my choices, for their ability of interpreting my moods, their enthusiasm in listening to academic stories, for helping me in all the situations, for being always a reference point and an infinite source of trust.

Last but not least, a special thank goes to all my friends spread out all over the world, but always present. Without their support, I would not have been able to simultaneously explore economics and myself.

Introduction

"Education enhances one's ability to receive, decode,
and understand information,
and that information processing and interpretation
is important for performing or learning to perform many jobs."
Richard R. Nelson and Edmund S. Phelps (1996)

Education as a way of increasing human capital is considered to be a basic factor in the growth process of the aggregate economy, according to the perspective that social returns to education extend beyond private returns. Returns to investment into human capital and educational processes are very crucial issues to analyze.

The leading theme in my research is the interest in the microeconomic aspect of human development both at theoretical and empirical level. In fact, this PhD thesis considers various aspects of the education system and its links with the labour market. It is composed by three chapters, each one corresponding to a self-contained paper, applying different methodologies (theoretical and empirical) and different perspectives. The first two chapters focus on the returns to education justified by the importance accorded to the investment in human capital as an explanation of wage differentials. The first chapter deals with the returns to higher education in 12 European countries, and with the evolution of wage inequality, in Europe. The second chapter provides empirical evidence of the returns to education acquired under communist regime. The third chapter explores the relationship between mental health and education decisions. Its aim is to examine how mental health predicts academic success.

More in detail, the first chapter investigates the evolution of the returns to higher education and of the college wage premium in Europe over the last 15 years. While

there has been intense debate in the empirical literature about the evolution of the college wage premium in the US, its evolution in Europe has been given little attention. This paper focuses on how does this evolution affect wage inequality and how does this evolution differ across age cohorts, in 12 European countries, using ECHP and EU-SILC data. Additionally, the paper explores whether there are cross-country differences in returns to education, and whether these are mainly driven by international differences in labour-market settings. The period analyzed is a period in which higher education participation rate increased dramatically: graduate supply considerably outstripped demand which ought to imply a fall in the premium. I use cross country variation in relative supply, demand and labour market institutions to look at their effects on the trend in the college wage gap. An important contribution to the literature is that I address possible concerns of endogeneity of relative supply by an instrumental variable strategy. I find some evidence of significant cross country differences in the level and the growth of the college wage premium. Results show a significant decline of college returns in countries with higher relative supply of skilled workers and a marked fall in college returns for recent cohorts for both men and women in all European countries. The estimated growth in the wage gap appears negatively correlated to changes in relative supply and positively correlated with the relative demand index. Institutional constraints also have a role in determining wage inequality.

In the second chapter, together with my coauthors Giorgio Brunello and Lorenzo Rocco, using data for Germany and 23 other economies in Eastern and Western Europe, we estimate the monetary returns to education acquired under communism more than 10 years after the fall of the Berlin Wall. We show that, in the 2000s, Eastern European workers who completed their education under communism received in the 2000s similar returns to their education as did workers belonging to the same age cohorts who studied in Western Europe. This might suggest that education under communism is still as valuable as education attained in Western Europe. However, individuals educated under communism are more likely than their Western counterparts to be unemployed, retired and disabled, and therefore to earn lower or zero returns to their education. Moreover, when we allow the returns to pre- and post-secondary education to differ, we find that senior males who have attained

only primary or secondary education under communism are penalized in the post-transition Eastern European labour markets, and that those who have completed post-secondary education under communism enjoy, in these markets, higher payoffs to their education than similarly educated Western European individuals employed in the West.

In the last chapter, my coauthors Francesca Cornaglia and Sandra McNally and I, explore the relationship between mental health and education decisions. Mental health problems have been rising internationally. Although poor mental health has often been correlated with poor educational attainment and/or dropping out of education, there have been few longitudinal studies on this subject. It is crucial to understand the link between mental health and schooling success since mental health problems can affect human capital productivity having lifelong consequences. We address this issue using a recent english dataset. England is a very interesting country to undertake such an investigation because of both poor mental health and high drop-out rate of young people. The Longitudinal Study of Young People in England (LSYPE) allows us to measure mental health at age 14/15 and again at age 16/17. This is measured using the General Health Questionnaire (GHQ12), a screening instrument used to detect the presence of symptoms of mental illness and depression. We associate poor mental health with examination performance at age 16 and with the probability of being observed as being "not in education, employment or training" (NEET) at age 17/18. Our results show that "poor mental health" is associated with lower examination performance and with higher probability of being NEET. Decomposing the measure of poor mental health into its component parts: "anxiety and depression ", related to excessive worrying and difficulty controlling this worrying; "anhedonia and social dysfunction", related to reduced interest or pleasure in usual activities; and "loss of confidence or self-esteem", there is some evidence that loss of confidence or self-esteem drives the association between poor mental health and exam results for boys. For girls this factor is also important but the association is stronger for anhedonia and social dysfunction. Additionally, we investigate whether these associations are influenced by controlling for past behaviour. For example, mechanisms through which poor mental health might impact exam performance and the probability of being NEET include substance abuse and

playing truant from school. Results show that these mechanisms have a potential role to play in understanding the relationship between poor mental health and exam performance. From a policy perspective, this paper helps documenting the importance of the relationship between poor mental health, educational attainment and subsequent dropping-out behaviour, suggesting that there could be a causal mechanism. Hence, programmes aimed at improving the mental health of adolescents may be very important for improving educational attainment and reducing the number of young people who are "NEET".

Introduzione

L' istruzione viene considerata da molti economisti come un investimento in un bene molto speciale: il capitale umano. Come per tutti i tipi di investimento, è molto interessante, oltre che utile, valutarne il rendimento.

La tesi il cui titolo è "Essays on Economics of Education" è un compendio di tre articoli tra loro indipendenti che applicano metodologie diverse, sia teoriche che empiriche, e impiegano diverse prospettive. Il primo capitolo, "Returns to college over time: trends in Europe in the last 15 years", è un analisi dell'evoluzione del "college wage premium" e dei rendimenti dell'istruzione terziaria in Europa, cercando di trovare, in fattori di domanda-offerta ed istituzionali, le cause. Il secondo capitolo " Lost in transition? The returns to education aquired under communism in the first decade of the new millennium", analizza i rendimenti dell'istruzione acquisita durante il regime comunista a distanza di un decennio dopo la caduta del regime. Il terzo capitolo, "Mental health and Education decisions" investiga la relazione tra salute mentale e output scolastici (test scores e NEET) per un campione di adolescenti inglesi.

Più in dettaglio, il primo capitolo studia l'andamento nel tempo del rendimento relativo dell' istruzione terziaria in Europa come possibile causa della diseguaglianza salariale tra e all' interno di gruppi di lavoratori con diverse skills. Questa tematica ha interessato molti studiosi, soprattutto a causa della crescente diseguaglianza salariale che si è osservata negli Stati Uniti a partire dalla fine degli anni 80. Nel tempo questo filone letterario si è consolidato soprattutto nell'indicare quali sono le sue potenziali cause (Skill Bias Technical Change, Rendimento del capitale umano e Istituzioni del mercato del lavoro). La maggior parte degli studi riguarda però gli Stati Uniti o alcuni paesi europei considerati singolarmente, ma solo pochi studi utilizzano un'ottica comparata che permette di analizzare anche il ruolo di isti-

tuzioni che variano tra paesi e non solo nel tempo, e questi pochi studi sono ormai datati. Questo è il primo studio che analizza questo fenomeno in Europa, osservando un intervallo temporale piuttosto peculiare. Il periodo esaminato è infatti caratterizzato da un ingente aumento della partecipazione nell'istruzione terziaria, e conseguentemente, un forte aumento dell'offerta di laureati nel mercato del lavoro, in conseguenza anche della politica comunitaria (Lisbona, 2000). Il dataset utilizzato è stato creato unendo dati ECHP e EU-SILC per ottenere un intervallo temporale di 15 anni (dal 1994 al 2009) per 12 paesi europei. I paesi europei sono divisi in due sottogruppi, quelli con elevata offerta relativa di laureati e quelli con bassa offerta relativa di laureati, poichè è plausibile che nei due sottogruppi di paesi l'interplay tra domanda e offerta relativa di laureati e l'effetto delle istituzioni possano essere diversi. I risultati mostrano l'esistenza di un'effettiva diminuzione dei rendimenti dell'istruzione terziaria in molti paesi europei, e questa diminuzione è più netta nei paesi con elevata offerta relativa di laureati. Si nota anche che tali rendimenti sono minori per le corti più giovani. Per quanto riguarda i fattori che possono spiegare le cause dell'andamento del "college wage premium", i fattori di domanda e offerta sono molto rilevanti. Un ulteriore notevole contributo alla letteratura è l'utilizzo di una strategia di stima basata sulle variabili strumentali, per evitare i problemi dervianti dalla postenziale endogeneità dell'offerta relativa di laureati. I risultati mostrano che l'aumento di offerta relativa ha impatto negativo e significativo in tutti i gruppi di paesi. Fattori istituzionali come salario minimo e sindacati sembrano giocare un ruolo abbastanza importante, in particolare in paesi con minore offerta relativa di laureati, paesi che hanno annche subito recentemente maggiori cambiamenti alle istituzioni riguardanti il mercato del lavoro.

Il secondo capitolo tratta la tematica classica dei rendimenti dell'istruzione, in particolare, questo studio confronta i rendimenti del capitale umano accumulato in un'economia di mercato con quello accumulato nelle economie pianificate durante il regime comunista. L'istruzione accumulata in sistemi educativi centralizzati, dove la scelta dell'individuo non risente degli stessi benefici e costi delle economie di mercato, è generlamente considerata come meno adatta a un economia di mercato come quella dei giorni nostri, e dunque, si assume che i rendimenti di tale tipo di istruzione siano minori. Contrariamente alla letteratura precedente che utilizzava

prevalentemente un approccio pre-post per ogni singolo paese, questa ricerca utilizza un'ottica comparativa e considera contemporaneamente diversi paesi europei. I dati utilizzati sono GSOEP, per la Germania, e EU-SILC, per i restanti 23 paesi. La strategia di stima è basata su un "quasi-esperimento", utilizzando gli individui dei paesi dell'Ovest come gruppo di controllo, per stimare i rendimenti dell'istruzione accumulata nei sistemi scolastici "comunisti" prima della transizione all'economia di mercato. Per ovviare al tipico problema di endogeneità in cui si incorre nella stima OLS dei rendimenti dell'istruzione, fornendo quindi una stima distorta che differisce tra i paesi essendo diverso il ruolo dell'abilità nell'accumulo di capitale umano nei diversi sistemi scolastici considerati, viene utilizzata la tecnica proposta da Card e Rothstein (2007). I dati vengono aggregati in celle (definita per età, genere, paese e anno), differenziando per genere e introducendo opportune e plausibili ipotesi sulla struttura del termine di errore. Una volta rilassata l'ipotesi di rendimenti costanti per ogni anno di istruzione, vengono esaminati anche i diversi rendimenti per i diversi livelli di istruzione. I risultati mostrano che i meno istruiti delle economie dell'Europa dell'EST hanno sofferto relativamente di più la transizione all'economia di mercato sia in termini di maggiore disoccupazione che in termini di salari inferiori, a parziale conferma dell'inadeguatezza dei livelli scolastici inferiori all'università di quei paesi.

L'ultimo capitolo affronta una tematica meno standard: studia la relazione tra problemi di salute mentale e accumulo di capitale umano nel Regno Unito, paese con un triste primato sia in termini di benessere dei bambini e adolescenti, che in termini di individui che non stanno ricevendo un'istruzione, non hanno un impiego o altre attività assimilabili (tirocini, lavori domestici, ecc.), e che non stanno cercando un'occupazione (NEET - Not in education, employment of training). In questa ricerca si stima l'effetto della salute mentale sulla performance scolastica e sulla probabilità di essere NEET. Oltre a misurare l'effetto totale della salute mentale, misurata dal GHQ-General Health Questionnaire, si stima anche l'effetto di ogni singola componente della misura complessiva di salute mentale, riconoscendone la diversa importanza e significatività. Per l'analisi viene utilizzato un ampio studio longitudinale con tutti i vantaggi che esso comporta rispetto a quelli cross-section, molto più recente rispetto ad altri studi. I dati utilizzati sono presi dal LSYP,

che riguarda un campione molto vasto di adolescenti Inglesi (14/15 anni nel 2004, seguiti fino al 2008). La natura longitudinale e la presenza di informazioni sui comportamenti rischiosi (consumo di sostanze rischiose come alcool, sigarette e cannabis e truancy- assenza ingiustificata da scuola) permette anche di investigare meglio i potenziali meccanismi indiretti dell'effetto dei problemi mentali sulla performance scolastica. In particolare si evidenzia che mentre esiste un effetto indiretto della salute mentale sulla performance scolastica che passa attraverso i comportamenti rischiosi, questo non sembrerebbe esistere sulla probabilità di essere NEET.

Chapter 1

Returns to college over time: trends in Europe in the last 15 years. Stuck on the puzzle.

1.1 Introduction

Beside being a decisive factor on individuals' earnings (Mincer, 1974), education is one of the main determinants of personal success and development (Jencks, 1979). Making higher levels of education attainable to everybody can be seen as a way of reducing income inequalities, improving economic growth (Krueger and Lindahl 2001; Bils and Klenow, 2000) and increasing general levels of welfare through positive externalities (Acemoglu and Angrist 2001; Milligan, Moretti, and Oreopoulos 2004 and Lochner and Moretti 2004). In the last two decades there has been a huge increase in the average years of attained education and the proportion of young people enrolled into higher education has significantly risen in all developed countries. Over the period 1990-2005, undergraduate enrolment has increased by almost 50 percent in Sweden, Finland and Denmark, and by over 30 percent in the UK, Ireland, Italy, Spain and Portugal thanks also to the European policies (i.e. Lisbon 2000). This "boom" in education can be interpreted as a supply shock to European labour market and it is likely it has substantially affected the structure of wage differentials. Investing in educational resources for disadvantaged families

to provide equal access to successful early human development is fundamental, and we can look at these increasing rates as positive factors: certainly, a more educated society is a better one. However, it is important to investigate whether the increase in the education attainment, in particular, higher education, has affected equality. Higher education -post secondary education and college education, is an important political and social issue in developed countries and it is imperative to assess its return.

From an equality point of view, rising inequality in personal incomes is a well-observed phenomenon in many countries. Rising inequality can take two specific forms: more inequality within skill groups and across skill groups. Inequality within skill groups can be caused by increasing fragmentation of jobs, new technologies and reduced wage compression efforts of unions and governments. Inequality across skill groups is different: here the main determinants are supply and demand of skills. What is important is to pinpoint changing skill differentials, i.e. by educational groups, as these differentials are important incentives for skill formation, school enrolment and training efforts. In the US, skill differentials have increased a lot in the last two decades. Between 1961 and 1979, returns to a college education (compared to a high-school degree) have increased from 61% to 82% [1] , despite the huge increase in the number of college graduates. What happened in Europe is less clear. Rising returns have been observed for Portugal, Denmark and Italy, constant returns have been found in the UK and Germany, and falling returns for Sweden and Austria (at the beginning of 2000). Unfortunately, the majority of these evidences are until the end of 1990s, and afterward the phenomenon has not been studied much. What is going on nowadays? Are the returns still increasing in a period which has seen sheer expansion of the demand of higher education, leading as well to the establishment of new institutions in many developed countries? It is reasonable to assume that changes in educational participation rates across cohorts are likely to imply changes in the ability-education relationship as well. If the ability composition changes, this can have an impact on estimated returns to education. Using the simple supply and demand framework, an increase in the supply of highly educated workers would cause a decline in their wages. The demand for college can

[1] Katz and Murphy (1992)

be rising dramatically, but if the supply keeps up with the demand, college wages will not increase.

Furthermore, how can inequality be affected by these trends? There could be two possible ways of higher enrolment affecting wage inequality, going in two different directions:

1) Increasing the proportion of college educated workers puts more weights on a distribution of wages exhibiting higher mean and higher dispersion. This would increase wage inequality.

2) Increased skill heterogeneity may lower the average college wage premium, leading to lower wage inequality. Additionally, conditioned on a demand for highly educated workers which does not outstrip supply, the increased proportion of college-educated people puts downward pressure on the college wage premium, thus lowering wage inequality.

Still, the supply and demand framework alone, cannot account for empirical puzzles such as the one of the US. Thus, if these inequality trends are not primarily explained by market-driven changes in the supply and demand for skills, it is possible they can be clarified also by episodic institutional shocks. Changes in institutional factors such as the minimum wage have contributed to the evolutions in the wage differential between college and non-college educated workers.[2] Europe can be different in this case from the US: the presence of stronger institutions helped to moderate the changes in the European OECD countries.

This paper investigates the evolution of the returns to higher education and of the college wage premium in Europe over the last 15 years. I want to assess whether higher education pays in the labour market and to examine what is the trend in earnings inequality over the period under study. I explore along what dimension inequality is changing and what shifts in the demand and supply and/or changes in wage setting institutions are responsible for the observed trend. Furthermore, I analyze if there are cross-country differences in returns to education, and whether they are mainly driven by international differences in labour-market settings, such as institutional features of wage formation, labour-market regulations, and the tightness of the labour markets. Or whether these differences are connected to the relative

[2]See Fortin and Lemieux (1997) for a review of the effect of labor market institutions on the wage structure.

pervasiveness of public sector employment or to cross-country differences in levels of welfare-state protection.

As to my knowledge, this paper is the only recent study about the evolution of the returns to college and college wage premium in Europe, not focusing on just one specific country. While this topic has been widely addressed in the US, its evolution in Europe has been given little attention. An important contribution of this paper is the fact that it can examine the returns to education in the long run, in recent years, in Europe, as well as investigate the directions and the drivers of change. The period I am focusing on is very compelling, since it is a period in which higher education participation rate increased dramatically: graduate supply considerably outstripped demand which ought to imply a fall in the premium. Hence, I contribute by assessing the pattern of the college wage premium as a result of the recent expansion in graduation rates, being able to look at the returns to different cohorts. I use cross country variation in relative supply, demand and labour market institutions to look at their effects on the trend in the college wage gap. Another novelty of this paper is that I address possible concerns of endogeneity of relative supply by an instrumental variable strategy, this is something that has never done in the literature before. I observe a significant decline of college returns in countries with higher relative supply of skilled workers and a marked fall in college returns for recent cohorts for both men and women in all European countries. I find evidence that both market and non market factors matters in explaining wage inequality. More specifically, college wage premium appears negatively correlated to changes in relative supply and positively correlated with the relative demand index, in particular, in countries with higher relative supply of skilled workers, that present a stronger decline in the returns to college: wider relative supply lead to a decline in college wage premium. Institutional constraints, such as minimum wage and unions also matter.

The paper is organized as follows. Section 1.2 and 1.3 present a review of the literature and the theoretical framework. Section 1.4 presents the data used and describes the raw trends in wage changes, education differentials and wage inequalities. Section 1.5 is dedicated to the empirical framework. Section 1.6 shows the results of the trends in between- and within-education group wage inequality and the potential explanations for these evolutions. Section 1.7 concludes.

1.2 Literature review

Increasing returns to education has always been linked to changes in wage inequality (Levy and Murnane 1992[3], Katz and Autor 1999). Many contributions in the literature have noticed a growing college wage premium over time. Greater college premium implies greater inequality. The underlying causes of increasing inequality are highly debated among labour economists. There are two leading explanations, skill biased technical change (SBTC) and labour market institutions.

Many empirical studies found the SBTC to be the driving force behind widening earnings inequality: this conclusion stems from the observation that the relative supply of high skilled workers and the skill premium can only increase together if the relative demand for high skilled workers increase as well.

Many studies, focusing on the US, have noticed a growing college wage premium but the role of the supply of college graduates in determining changes in the returns to a college education has not been explored much. Katz and Murphy (1992) analyze the wage movements over 25 years, from 1963 to 1987, in the US, concluding that the rising in the relative demand for more skilled workers is "a key component of any consistent explanation for rising inequality and changes in the wage structure over the last 25 years".[4] Furthermore, they identify the fluctuations in the college/high school differential over that period, in the combination of growth of both relative supply of college graduates and demand for more educated workers. More recently, Taber (2001) prefers an explanation based on an increase in the demand for unobserved skills rather than one based on an increase in the demand for skills accumulated in college. His work suggests that the observed rise in the college premium in the 1980s is just a reflection of the increase in the return to unobserved ability: "rising returns to unobservable skills correlated with education is the main explanation behind the increased education wage differentials".[5] However, Chay and

[3]In an earlier contribution, Levy and Murnane (1992) present a set of hypotheses for explaining not only within-group inequality but also the growth of within-group variation over time. Their hypotheses include both supply and demand shifts for workers characteristics; the former consists in the changing characteristics of the labour force (including aptitude test scores, measures of ability to work with other people); as well as increasing returns to skill; the latter includes plant specific wage differentials within industry as well as changes in wage-setting institutions.

[4]Katz and Murphy (1992)

[5]Taber (2001).

Lee (2000) argue that the latter raise in unobserved ability accounts at most for 30 to 40 percent of the increase in the college premium. Similarly, Deschenes (2006) argues that most of the increase in the college premium is due to an increase in the return to schooling.

This evidence that, over the last decades of the 20th century, the US faced a simultaneous expansion of both college graduates and returns to education contradicts with the general law of demand and supply. The basic rule would suggest indeed that the price of a graduate worker should decrease when increasing its supply. This inconsistency has generated a large body of literature (Murphy and Welch, 1989, Card and DiNardo 2002, among others). However, surprisingly, the additional observation of a general decline in real earnings at college and lower educational levels has been mostly ignored when understanding this paradox. The study by Card and DiNardo (2002) is one of the firsts noticing a deceleration in the college wage premium, contrasting with the preceding decade. They provide evidence that increasing education can lower wage inequality. Card and Lemieux (2001), using a model with imperfect substitution of workers with similar education but belonging to different age cohorts, find that own cohort supply of college-educated graduates negatively impact the college wage premium: they show that the rise of the premium is confined to rise for younger workers which can be driven by falls in the growth of educational attainment that began with cohorts born in the 1950's. Lemieux (2006) investigates the change in wage inequality and wage structure, showing that in the US, between 1973 and 2005, returns to post secondary school increased sharply whereas returns to lower levels of education remained unchanged. Using quantile regressions he shows that the return for post secondary education has increased more in upper quantiles (like the 90th).

On the other hand, other researchers have argued that skill biased technological change can not explain alone the increase in wage inequality during the '80s. Acemoglu (2003) argues that the relative supply and demand framework does not provide an entirely satisfactory explanation of the behaviour of skill premia across countries. Giving space to labor market institutions to play an important role in the story.

The alternative explanation attributes international differences in wage inequal-

ity across skill groups to differences in labor market institutions. Several explanations for the rise in wage inequality focus on changes in wage-setting institutions.[6] 'Institutions' are non competitive forces acting on the labour market, such as labor unions, minimum wage, product and labour market regulations, taxes and subsidies and social norms. All these factors can affect the shape of wage distribution, including earnings inequality.

The two institutions that have received more attention in the US are labor unions and the minimum wage. DiNardo, Fortin, and Lemieux (1996) find that, in addition to supply and demand factors, also changes in labour market institutions -namely de-unionization and declining minimum wages - are important in explaining wage inequality. Lee (1999), using variation in the minimum wage across regions, shows that not only minimum wage is negatively correlated with rising inequality at the top end of wage distribution, but also it can explain much of the increase in the dispersion at the lower end of wage distribution. Goldin and Katz (2007) use a supply, demand and institutions framework to understand the returns to education in the US, in the past century, combining the usual supply-demand framework with institutional rigidities and alterations.

Concerning Europe, few are the studies on the evolution of college wage premium and skill differentials. The majority of the studies dealing with the returns to education in Europe focus on both standard returns to education and on single countries. Recent evidence of the impact of the increasing supply of graduates on their wage are available for the UK (Walker and Zhu, 2008; Chevalier and Lindley 2009). In particular, Walker and Zhu (2008), are interested in how the college premium has varied across time, across subjects of study, across the wage distribution and across two different cohorts. They show that up to 2000 there is almost no evidence of declining returns to college following the surge in participation in higher education, however, beyond 2002 they find suggestive evidence of modestly declining wage premia for graduates. Furthermore, very few are the studies dealing with the relationship between wage inequality and education. Harmon, Oosterbeek, and Walker (2003), use UK data and find that the returns to schooling are higher for those at the very top of the wage distribution compared to those at the very bottom.

[6]Bluestone and Harrison (1988) offer an extensive discussion of the possibilities.

Martins and Pereira (2004) have provided empirical evidence that in fifteen European countries during the mid 1990s, returns to education at the upper quantiles significantly exceeded those at lower quantiles, that is increasing education increases within wage inequality: in 15 European countries, more skilled workers (individuals receiving higher hourly wages conditional on their characteristics) are associated with a stronger education-related earnings increment. Leuven, Oosterbeek, and van Ophern (2004) use data on cognitive skills to look how well cross-country differences in supply and demand can explain differences in skill differentials. Concerning the institutional literature, Machin (1997) and Dickens, Machin, and Manning (1999) for the UK, find that, respectively, higher union density and higher minimum wages reduce wage inequality. Manacorda (2004), in Italy, and Edin and Holmlund (1995), in Sweden, find that wage setting institutions are important for wage inequality. Koeniger, Leonardi, and Nunziata (2007) , with panel data on institutions in OECD countries, assess the quantitative relationship between institutions and male wage inequality. Their findings show that labour market institutions matter: employment protection index, unemployment benefit, union density and the minimum wage are significantly negatively associated with wage inequality within countries. An interesting study combining SBTC and institutions is Brunello, Comi, and Lucifora (2000). They look at the evolution of the college wage gap in 10 European countries from the early to mid 1980s to the mid to late 1990s, finding significant cross country differences in the level and dynamics of the gap. In particular they find negative correlation between wage gap and relative supply and positive correlation both with the long run rate of productivity growth and with an index of between industry demand shocks. Among the relevant institutional factors, the find declines in union density, in the centralization of the wage bargain and in employment protection measures to have lead to a faster growth in the college wage gap. Barth and Lucifora (2006) investigates the effect on the wage structure of the boom in education in Europe, estimating a model with supply and demand for types of workers. Their findings suggest that the educational boom matched the demand shifts due to skill boas technical change, and they find no evidence supporting the hypothesis of skill erosion within college graduates.

1.3 Theoretical framework

Following the conventional conceptual framework of this literature[7], I model the relative wage dynamics as a combination of supply and demand factors and labour market institutions.

From a theoretical perspective there is the need to account separately for the relative wage of two types of workers. Consider an extended version of the CES production function with two labour inputs that are imperfect substitutes: low educated (or unskilled) and high educated (or skilled). Assume that firms in each economy use the following simple production function where output depends on employment:

$$Y_{ct} = e^{\phi_{ct}} N_{ct} \tag{1.1}$$

with Y being the total output produced, N the employment in efficiency units, c the country, t the time and ϕ a country and time specific productivity shock, a parameter denoting total factor productivity.

Employment is made by two groups of workers, skilled and unskilled labour, which are employed according to

$$N_{ct} = [(e^{\alpha_{lct}} L_{ct})^{\rho} + (e^{\alpha_{hct}} H_{ct})^{\rho}]^{\frac{1}{\rho}} \tag{1.2}$$

α is an efficiency parameter indicating the productivity of a particular type of worker (L,H) in country c at time t, it is an index of the technological efficiency of a worker as it is factor augmenting technical change parameter capturing changes in input quality over time. H_{ct}, L_{ct} are the quantities employed of college equivalent (skilled labour) and high school equivalent (unskilled labour).

It is assumed that the economy is at full employment, that means the total effective aggregate labor supply of each labor group is employed in the industries of the economy. Another assumption is that H_{ct}, L_{ct} are exogenous. That is the

[7]In their paper, Katz and Murphy (1992), used a demand and supply of skills framework to analyze the change in wage inequality over time. The same framework has then been used by Katz and Autor (1999), Goldin and Katz (2007) and Leuven, Oosterbeek, and van Ophern (2004) to look at differences in skills groups across countries. All these studies focus exclusively on demand side modeling

aggregate supply does not depend on its relative average wage.

$\rho = 1 - 1/\sigma$, is a time-invariant production parameter, where σ is the aggregate elasticity of substitution between labour inputs. The low quality and high quality workers are gross substitutes if $\sigma > 1$ and $\rho > 0$, whereas they are gross complements if $\sigma < 1$ and $\rho > 0$.

Skill neutral technological progress raises both $e^{\alpha_{lct}}$ and $e^{\alpha_{hct}}$ by the same proportion. Whereas, skill biased technical changes involve the increase of $\frac{e^{\alpha_{hct}}}{e^{\alpha_{lct}}}$

Competitive labour markets are assumed, so college equivalent and high school workers are paid their marginal products, then profit maximization with respect to N_{ict} (with $i = L, H.$) yields to

$$w_{ict} = e^{\phi_{ct} + \alpha_{ict}} \left[\frac{N_{ict}}{N_{ct}} \right]^{\rho - 1}$$

where w_{ict} is the real wage for labour input i in country c at time t.

In other terms, efficient utilization of different skill groups requires that the relative wages are equated to the relative marginal products.

The relative wage of high skill to low skill workers can be written as

$$w = \frac{w_{ct}^{H}}{w_{ct}^{L}} = \left(\frac{e^{\alpha_{hct}}}{e^{\alpha_{lct}}} \right)^{\frac{\sigma - 1}{\sigma}} \left(\frac{H_{ct}}{L_{ct}} \right)^{-\frac{1}{\sigma}} \tag{1.3}$$

which is equal to:

$$lnw = \rho \left(\frac{\alpha_{hct}}{\alpha_{lct}} \right) - \frac{1}{\sigma} ln \left(\frac{H_{ct}}{L_{ct}} \right) \tag{1.4}$$

The relative wage of different educational groups is generally used as a measure of between groups inequality. $\left(\frac{H_{ct}}{L_{ct}} \right)$ represents the relative supply of skilled versus unskilled labour, and $\left(\frac{\alpha_{hct}}{\alpha_{lct}} \right)$ the skill bias technological change. This can be rewritten as

$$ln \left(\frac{w_{ct}^{H}}{w_{ct}^{L}} \right) = \frac{1}{\sigma} \left[D_t - ln \left(\frac{H_{ct}}{L_{ct}} \right) \right] \tag{1.5}$$

where D_t indexes relative demand shifts which favor high skilled workers and it is measured in log quantity units.

Equation (1.4) can lead to a very simple and intuitive demand-supply interpretation. Given a skill bias technical change, the substitution effect is such that the skill

premium increases when there is a scarcity of skilled relative to unskilled workers.

Consequently, $-\frac{1}{\sigma}$ represents the slope of the relative demand of skilled versus unskilled workers: the impact of changes in relative skill supplies on relative wages is inversely related to the magnitude of aggregate elasticity of substitution between two skill groups. That is, the greater is σ, the smaller is the impact of shifts in relative supplies on relative wages, that means the fluctuations in the demand shifts must be greater to be able to explain changes in the relative wages.

Relative demand changes can be due to shifts in product demand, SBTC and non-neutral changes in the relative changes in relative prices/quantities of non-labour inputs, so marginal productivity and elasticity.

The relative demand is shifted by the bias of the technological change:

$$\frac{\partial lnw}{\partial \left(\frac{\alpha_{hct}}{\alpha_{lct}}\right)} = \frac{\sigma - 1}{\sigma}$$

This means that, given the relative supply, if there is skill biased technological change (i.e. technological shock shifting the demand line outwards) the wage premium will increase.

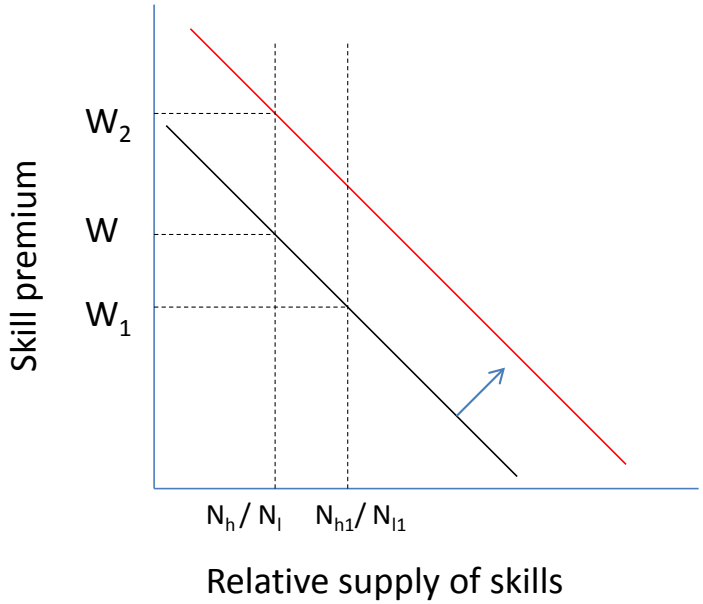
Similarly, for a given "skill bias", $\left(\frac{\alpha_{hct}}{\alpha_{lct}}\right)$, an increase in the relative supplies $\left(\frac{H_{ct}}{L_{ct}}\right)$ lowers relative wages with elasticity σ.

Figure 1.1 shows how an increase in the supply (from N_h/N_l toN_{h1}/N_{l1}) reduces the skill premium (from w to w_1) and how a skill biased technological shock (outwards shift in the demand line), given the supply, increases the skill premium (from w to w_2).

Following the reasoning above, the evidence of a negative relationship between college premium and relative supply of skills in the recent period in Europe can be interpreted as an increase in the relative supply of college skills, under the assumption of stable demand's conditions.

The main assumption of this model is that the supply of skills is predetermined. This setting assume market clearing, meaning that there is no unemployment. This is an assumption that can be criticized, however this is standard in this literature. In short, there are the main forces that operates in this framework: the relative supply and the relative demand of more-educated workers. When these two forces

Figure 1.1. Skill premium and relative supply of skills.



fail in explaining the wage differentials, the pattern can be reconciled by institutional factors such as change in union density/strenght and wage setting policies. Labor market institutions, indeed, differently alter the outside option of skilled and unskilled workers thus affecting wage differential as well as relative labor demand.

Moreover, it is reasonable to assume that, in a period of accelerating education expansion, educational premia are likely to twist reducing inequality among young workers relative to the old (the opposite should be true if the education expansion is decreasing).[8] What is important in this framework, in addition to the level of educational supply, is its rate of change.

Assuming that there can be differences on the level of wages depending on age, that means that age cohorts are imperfect substitutes in production, a common way to combine them is as CES aggregate. In each country, we thus have:

[8]The intuition is the following: when education levels are arising, younger cohorts are relative more educated than older, when education levels stagnate, this implies that the pattern of educational differentials across cohorts twists.

$$H_t = (\sum_J \delta_j H_{jt}^{\eta})^{1/\eta}$$

and

$$L_t = (\sum_J \beta_j L_{jt}^{\eta})^{1/\eta}$$

with $\sigma_A = 1/(1-\eta)$ is the elasticity of substitution between different age cohorts, δ, β efficiency parameters assumed fixed, j indices the age groups and H_{jt}, L_{jt} are age groups specific supply by education in each time period.

The aggregate output is again function of total skilled and unskilled supply, and some technological parameter, simplifying (1.1) :

$$Y_{ct} = [e^{\alpha_{Hct}} H_{ct}^{\rho} + e^{\alpha_{Lct}} L_{ct}^{\rho}]^{\frac{1}{\rho}} \tag{1.6}$$

Under the general assumption the the economy is competitive and that wages are paid their marginal products[9], then

$$\frac{\partial Y_{ct}}{\partial H_{jct}} = \frac{\partial Y_{ct}}{\partial L_{ct}} \times \frac{\partial L_{ct}}{\partial L_{jct}}$$

Writing the relative wages of skilled versus unskilled workers in the same cohort, we get:

$$ln\left(\frac{w_{jct}^H}{w_{jct}^L}\right) = \left(\frac{\alpha_{hct}}{\alpha_{lct}}\right) + (\rho - \eta) ln\left(\frac{H_t}{L_t}\right) + ln\left(\frac{\beta_j}{\delta_j}\right) + (\eta - 1) ln\left(\frac{H_{jct}}{L_{jct}}\right) \tag{1.7}$$

Therefore, the relative wage ratio for cohort j, depends on the age specific efficiency parameters β_j, δ_j and on the relative supply in the given cohort $\left(\frac{H_{jct}}{L_{jct}}\right)$, in addition to the technology parameters and the aggregate supply.

Rearranging, equation (1.7) can be rewritten as:

[9]Efficient utilization of skill groups further requires that relative wages across skill groups are equated with relative marginal products.

$$ln\left(\frac{w_{jct}^{H}}{w_{jct}^{L}}\right) = \left(\frac{\alpha_{hct}}{\alpha_{lct}}\right) - \frac{1}{\sigma} ln\left(\frac{H_t}{L_t}\right) + ln\left(\frac{\beta_j}{\delta_j}\right) - \frac{1}{\sigma_A}\left[ln\left(\frac{H_{jct}}{L_{jct}}\right) - n\left(\frac{H_{ct}}{L_{ct}}\right)\right] \quad (1.8)$$

1.4 Data and aggregate trends

1.4.1 Dataset

I use a unique dataset, harmonizing the European Survey of Income and Living Condition (EU-SILC) and European Community Household Panel (ECHP), to assess the returns to college and wage inequality in Europe from 1994 to 2009. This paper is not the first one using ECHP and EU-SILC as a single data source.[10]

The EU-SILC is a collection of timely and comparable multidimensional microdata covering EU countries, starting in 2004, and conducted yearly until now (data available until 2009), for a total of six waves. It is based on nationally representative samples, which collects comparable cross sectional and longitudinal micro data on income poverty and social exclusion and contains information on income, housing, material deprivation, labour, health, demography and education.

The ECHP, precursor of the EU-SILC, started in 1994 and ended in 2001, thus consisting of eight waves. In the first wave in 1994, about 60,000 nationally representative households with approximately 130,000 individuals aged 16 years and over were interviewed in the 12 participating member states.[11]

One advantage of these data is that I have an overall period of 15 years in which I can observe a total of 12 European countries: Austria, Belgium, Germany, Denmark, Spain, Finland, France, Greece, Ireland, Italy, Portugal and United Kingdom. For each country in the sample, I only consider the sub-sample of individuals who reside in the country of birth (more than 94 percent of the total in 2009).

The reference sub-sample focuses on native male and female employees between 25 and 50 years old and are working. This age framework allows me to compare

[10] See for example Massari et al. (2012) and Goos et al.(2009)

[11] Austria, Finland and Sweden joined the ECHP project in 1995, 1996 and 1997, respectively. Sweden, Luxembourg and the Netherlands have to be excluded from the analysis because required information is missing.

the youngest college graduates with their non-graduates counterparts and to avoid selection bias due to retirement and pensions.

I use net annual earnings in the reference sub-sample of all wage and salary workers in the public and private sector. All measures of wages in the paper are adjusted and deflated using the Purchasing Power Parity PPP (base Euro 15=1) to take into account different cost of living and to allow for comparison among years.

To avoid bias from incorrect income data (outliers), I omit all employees whose net wages are below the minimum contribution level of the social Security System or above a certain threshold.

I define skilled workers whose with at least some higher education (i.e. tertiary or post secondary non tertiary education).

The two surveys record differently information about schooling and sometimes not even consistently through time. ECHP only displays information about the highest earned qualification, and provides an education variable in three levels: 3 broad levels (low -middle-high skills= low, secondary, college. They correspond to 0-2, 3 and 4-6 ISCED levels respectively. EU-SILC contains information on both earned qualifications (highest ISCED level achieved) and on ages at which individuals left school.

The construction of a consistent variable recording the entire length of the education path of workers across countries is problematic because of differences in schooling systems across the countries, and the lack of a record in the data. Because data on the actual years of schooling are not recorded in the survey, the measure of years of schooling used in these countries is a derived one. I have calculated the total number of years of education obtained by individuals in the following way: age in which the worker ended highest general education course minus starting education age according to the country of origin. Certainly this measure is problematic, it may introduce substantial bias since it can not take into account non-binding time frames for university degrees, or individuals dropping out of some degree, without finishing, to start a different one.

In order to keep the analysis as consistent as possible, the classification criterion applied is the highest educational qualification which is common to all countries and whose information is available in all datasets.

Therefore the three educational groups are defined as follows:

1) Low (primary or lower) education;

2) Intermediate (secondary) education;

3) High (post secondary-tertiary) education.

The advantage of this variable with respect to years of education is that it accounts for different duration of analogous school cycles.

In both the dataset there is no information about actual work experience or years of work interruption. Therefore, in the regressions I use potential experience conventionally defined as in Autor et a. (2008): Min{Age – Years of schooling- the age at which children start school; age-16}

The college wage premium is defined as the ratio of wage rates between college and high school graduates.

To control for aggregate labor supply and demand conditions, I use data from the OECD, EUKLEMS and ILO.[12] In particular, for the supply index, I use OECD data on the relative skill endowment, measured in terms of educational attainment. For the demand index I use data from EUKLEMS on the share of hours worked by skill workers relative to low skill workers. The institutional data are provided by OECD and ILO. These are yearly data which do not depend on the skill level, measuring wage bargaining institutions, strictness of employment protection legislation, minimum wage, union density and public sector employment.

1.4.2 Descriptive statistics

Tertiary education attainment more than doubled in most European countries, over the last decades. The strong increase in participation rates in Europe is evident: Figure 1.2 shows the evolution of higher education (post secondary and college) in Europe over the last 15 years. In particular, it shows the percentage of people aged 20-50 achieving post secondary education from 1994 to 2009. The trend is strongly increasing for both men and women, with women presenting a more marked increase. The slight decline in 2008 and 2009 can be due to the fact that some interested people may still be in education.

[12]Detailed information can be found in the data appendix A1.

Figure 1.2. Evolution of higher education

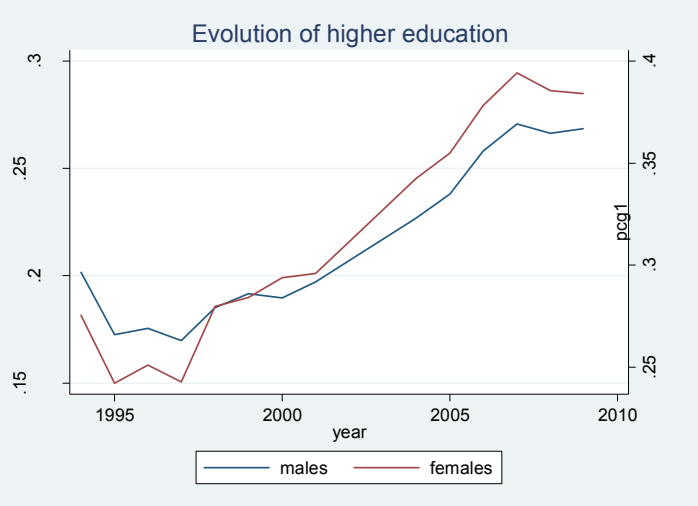


Figure 1.3. Increasing trend in higher education by cohorts

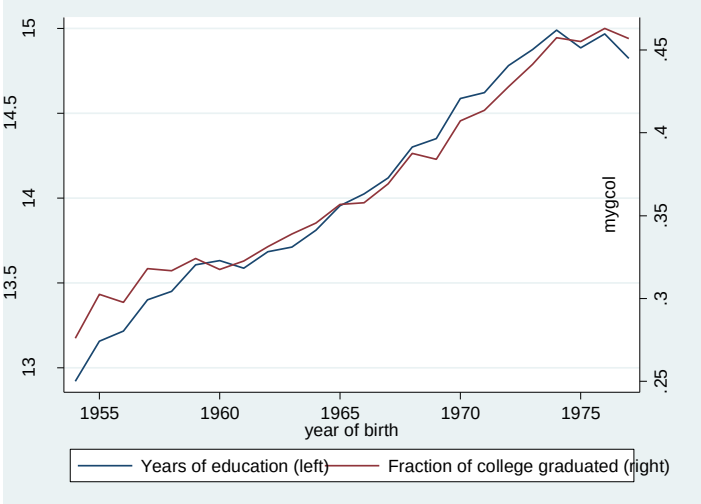


Figure 1.3 shows the recent history of the percentage of each cohort currently undertaking higher education. The figure confirms the increasing trend in education attainment in Europe over time, showing that the average years of education achieved and the fraction of college graduates have increased by age cohorts. For people born in 1955 the average number of years of education completed was almost 13.5 year, and the percentage of higher educated of that cohort was 30%; these numbers are almost 15 and 45% for the 1975 cohort.

The sample I am using differs by countries in population and income shares of each educational group. Over the period, mean real income by educational group changed differently across countries and educational groups. However, the trends in the education patterns (generally increasing) are pretty similar in many European countries. Namely, I differentiate between countries with high (initial) relative supply of graduates and countries with low (initial) relative supply of graduates, measured at the beginning of the period. Denmark, Finland, Ireland, Spain, France and Belgium are countries that were experiencing high percentage of people achieving higher education in the '90s. On the other hand, countries, such as Italy, the UK, Portugal, Germany, Greece and Austria, had lower graduate rates at the beginning of the period analyzed. These countries are divided according to the ratio of college graduates over high school graduates. This is a measure of the relative supply of graduates in each country. Looking at the values of this ratio in 1994, I divide into two regions: high and low relative supply of graduates countries. Countries characterized by a lower stock of high educated individuals experienced even higher growth in attainment levels, thus suggesting a catching-up phenomenon.

These aggregate patterns hide significant heterogeneity across countries.These two set of countries are thus very likely to have faced different evolutions in the educational attainment, as well as different evolution (different saturation times) of the demand for these type of workers. Additionally, these two set of countries differ for different level and degrees of labour market institutions.

In table 1.1 descriptive statistics of education and income in different regions and by different years are shown. The percentage of people achieving different degrees, together with the average years of education achieved and the log of wages are shown for both men and women in the two regions: high and low relative supply countries.

Table 1.1. Descriptive statistics.

	High relative supply		Low relative supply	
	ECHP	EUSILC	ECHP	EUSILC
Panel A: Males				
College	35.44%	34.24%	16.92%	24.78%
Secondary	34.77%	42.79%	39.28%	43.75%
Low	29.79%	22.97%	43.80%	31.47%
Years edu	12.73	13.76	12.15	12.94
Log wage	9.58	10.01	9.21	9.73
Panel B: Females				
College	44.97%	45.55%	23.13%	34.80%
Secondary	33.87%	39.13%	41.44%	42.87%
Low	21.16%	15.32%	35.43%	22.33%
Years edu	13.19	14.52	12.54	13.42
Log wage	9.29	9.73	8.94	9.43

As said before, the recent rapid expansion of higher education rates, has some shadows. Firstly, to assess whether the increase in participation was beneficial or not, it would be interesting to answer the following question: Has this increase in highly educated people flooded the labour market that the wage premium for higher education has been significantly reduced? A closely related issue is the possibility that this expansion has digged deeper into the distribution of students' abilities given the possibility to weaker and less able students to access higher education, thus resulting in less productive graduates than the ones of earlier cohorts. Moreover, a concern about school and teacher quality can arise. Indeed, this can be caused by a reduction in the average productivity of the recent cohorts of graduates as well.

All this points would suggest that the recent expansion may have resulted in lower returns in particular at the bottom of the wage distribution where less able individuals might be expected to be concentrated.

1.4.3 Relative wage changes and educational differentials

From the descriptive table in the appendix -see table A1, it is evident that younger cohorts have, on average, lower real wage rates, reflecting a combination of both age differences and of the overall decline in average real earnings in Europe.

Older male and female cohorts have higher earnings with respect to younger cohorts, however this can be a consequence of the life-earning profile. An interesting feature of Table A1 relates to the differences across cohorts in educational attainment. Average education displays a rising inter-cohort trend for the cohorts born before 1950, followed by a decline for those born in the 1950s and early 1960s. This pattern is documented and analyzed by Card and Lemieux (2001).

Figure 1.4. Evolution of college wage premium

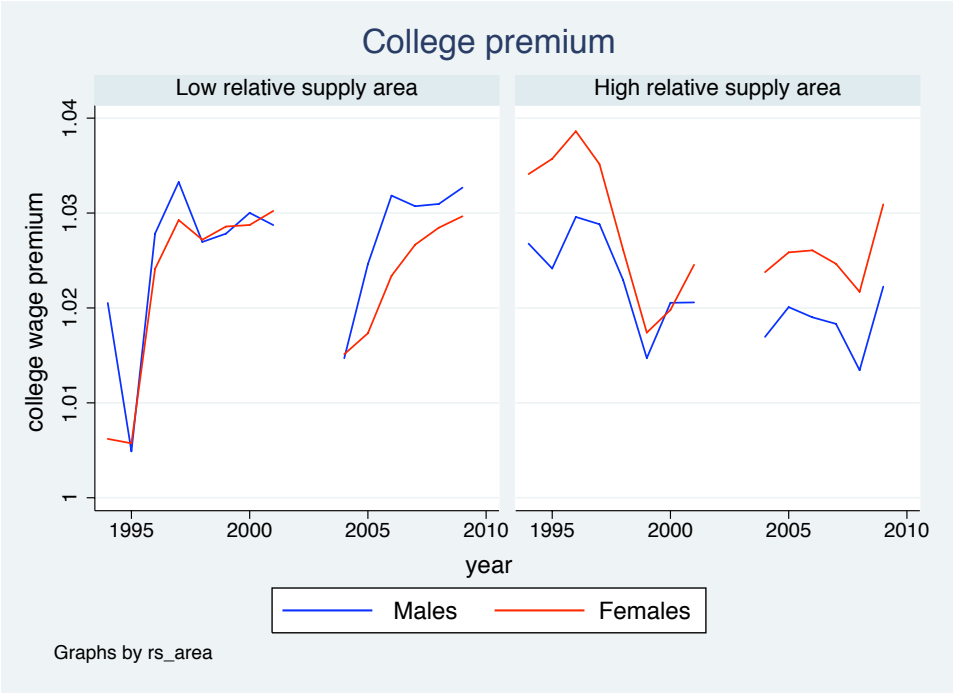
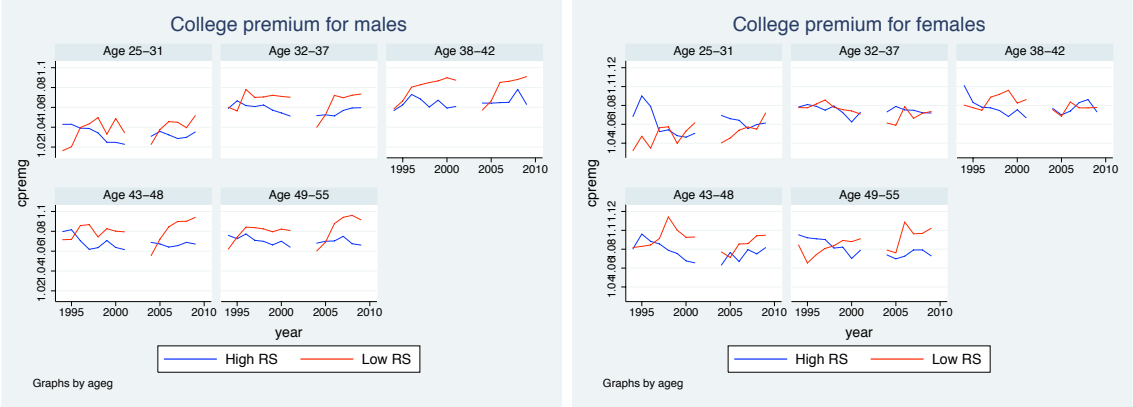


Figure 1.4 show that college wage premium has evolved very differently among countries with high and low relative supply of graduates. College wage premium is calculated as the ratio between the log wage of college graduates and high school graduates. The level of the college wage premium is always positive, being a measure of the higher rewards for the more educated, with high relative supply countries falling down heavily over time and low relative supply countries experiencing a growing trend. The trend is very similar for men and women in both set of countries. The pattern observed in the high relative supply countries would suggest that the huge influx of college graduates has saturated the demand for this type of workers, reducing continuously their potential comparative advantage, and generating in this

Figure 1.5. Evolution of college wage premium by age cohorts



way people that, despite having a degree, are not that different from their high school graduate peers. This is not the case in low relative supply countries: it seems to be the case that in this set of countries there is still an unsaturated demand for skilled workers.

Nevertheless, the evolution over time of the college wage premium can be due to both, different dynamics of cohort-specific relative wages, and changes in the composition of employment by cohort. This means that the relative wage can vary across cohorts and, more specifically, younger cohorts can experience higher wage gaps. For this reason, it is interesting to look at the evolution of the college premium by different cohorts. In figure 1.5 individuals are grouped by level of educational attainment, cohort and country.

The figure on the left shows the cohorts evolution for men. Quite interestingly, the differences between cohorts and regions are striking: firstly younger cohorts are always showing much lower premia with respect to the oldest ones. Additionally, high relative supply countries are showing a declining college premia over time for each cohort considered, on the contrary, the low relative supply countries are experiencing an increasing trend. The situation is less evident for females: only the oldest cohorts in low relative supply countries the premium is increasing and is higher than in high relative supply countries.

1.4.4 Wage inequality

As Ashenfelter and Rouse (2000) state ‘‘The school is a promising place to increase the skills and incomes of individuals. As a result, educational policies have the potential to decrease existing, and growing, inequalities in income’’.[13]

This line of thought carries with it the presumption that new highly educated cohorts will benefit from such levels' traditionally high returns. However, this approach does not consider whether such levels are characterized by reasonably concentrated or disperse returns. If the latter situation turns out to be the most representative, then one should acknowledge the existence of potential problems relating to within-levels inequality of educational policies intended to fade wage dispersion. Moreover, the scarce evidence available suggests that ‘‘differences in the extent of earnings inequality among high income countries are heavily influenced by rewards for educational attainment’’.[14]

Table 1.2 shows the trend, in the microdata, the age (experience) premium and the education premium, both measures of between-wage inequality. The former is the ratio between the earnings of ‘younger’ (25-30) and ‘older’ (45-50) workers, the latter is the ratio of the earnings of university graduates to the earnings of high school graduate. Concerning the age premium, Panel A, for countries with high relative supply, specifically for males with college degree, the trend is slightly increasing, although it is decreasing for non college degrees and for both categories in countries with low relative supply. For females both with and without college education, in both regions, the evolution is more stable even if declining in high relative supply countries and increasing in the low relative supply area. The trend in the education premium, Panel B, seems to be pretty stable for females in low relative supply countries, decreasing for both men and women of different age groups in high relative supply of graduates countries and increasing, for the old age cohorts, in low relative supply countries.

[13] Ashenfelter and Rouse (2000, p. 111)
[14] Sullivan and Smeeding (1997).

Table 1.2. Between group inequality: Age and education premia.

	High relative supply		Low relative supply	
	ECHP	EUSILC	ECHP	EUSILC
Panel A: Age premium				
MALES				
college	2.04	2.25	1.67	1.50
non college	2.15	1.90	1.60	1.48
FEMALES				
college	1.92	1.72	1.49	1.62
non college	2.19	1.93	1.85	1.61
Panel B : Education premium				
MALES				
Age <=28	1.24	1.14	1.25	1.36
Age 29-34	1.43	1.25	1.50	1.44
Age 35-49	1.54	1.45	1.68	1.69
Age 40-45	1.60	1.58	1.62	1.77
Age 45+	1.67	1.64	1.69	1.75
FEMALES				
Age <=28	1.38	1.32	1.24	1.37
Age 29-34	1.45	1.36	1.38	1.42
Age 35-49	1.50	1.41	1.45	1.46
Age 40-45	1.54	1.47	1.59	1.56
Age 45+	1.62	1.55	1.53	1.63

1.4.5 Labour market institutions

Institution is a system of laws, norms or conventions resulting from a collective choice, and providing constraints or incentives which alter individual choices over labor and pay. Institutions create a wedge between the value of the marginal job for the firm and the wage. Traditionally in the literature, the institutional features that are considered important for wage formation are: unions and bargaining institutions, wage regulation and welfare benefits, and labour market policies. A common finding of the studies that have investigated the effects of institutions on wage dispersion is that the interactions between supply, demand and institutions can take several routes altering both the between as well as the within structure of wages (see for example, Brunello, Comi, and Lucifora (2000) and Barth and Lucifora (2006)).

In investigating the evolution of wage inequality, I use institutions as another potential explanation of the trend in the college wage gap.[15] I use union density as

[15]Detailed information on institutional data used in the empirical analysis can be found in

a measure of wage bargaining institution. The data on Employment Protection Legislation index the set of rules and procedures governing the treatment of dismissals of workers employed on a permanent basis. Statutory minimum wage is conventionally defined as the ratio between the official minimum wage and the median wage.[16] Table A2 contains summary statistics of the institutional variables.

It is necessary to have time varying information on institutions. Indeed, the effects of institutions in regulating wages might change over time because of market deregulation, depletion of workers' guarantees, deunionisation and decentralisation of collective bargaining.

Generally, institutions are pretty stable, in the sense that do not change much over time. However, in the period analyzed there has been sufficient labour market related reforms. The two regions analyzed differ by institutional settings as well. Namely, countries with higher relative supply of graduates seem to be also more protective: the employment protection index is higher, as well as the union density and the minimum wage. And countries with lower relative supply are the ones which implemented more reforms during the period. All the differences are significant. These countries present lower inequality (lower Gini coefficient), and slightly higher employment rate. Concerning the demand of graduate workers, there is a lot of heterogeneity across countries, however on average it seems that there are no big differences among the two regions. Reforms actually implemented in EU countries in recent years with the goal of fighting unemployment did not increase or reduce employment protection or increased the generosity of unemployment benefits for everybody.

1.5 Empirical framework

In the empirical exercise, I first take a long run perspective and analyze the effect of having college or high school degrees on the net wages over time. In order to obtain some simple evidence on the form of the relationship linking earnings and schooling, I estimate an unrestricted regression of log wage on a set of dummy

appendix A1.

[16]It is to be noticed that not all the countries in our sample have an official minimum wage: Austria, Germany, Denmark, Finland and Italy do not have an official minimum wage.

variables for each schooling level available in the data. To investigate the potential sources of inequality I estimate regression models for the college wage gap that extend the basic specification in equation 1.5. I address the issue of the potential endogeneity of relative supply in the college wage premium equation with an IV strategy. Furthermore, I run quantile regression estimates to address the relation between schooling and wage inequality. Quantile regressions are used to consider the differences through income distributions in education premia between different groups of individuals.

1.5.1 Returns to college

In the first part of the empirical analysis I focus on the evolution of returns to college over time. Ordinary least squares methods are applied to standard Mincerian earnings function where the education variable, instead of being measured by the number of years of education completed, takes the form of set of dummy variables indicating the type of degree completed. The equation of interest becomes the following:

$$Y_{icat} = \alpha + \beta_1 College_{icat} + \beta_2 Secondary_{icat} + \beta_3 EXP_{icat} + \beta_4 EXP^2_{icat} + \lambda_{at} + \theta_{ct} + \gamma_c + \tau_t + \chi_a + u_{icat}$$

for the individual i, in country c, of the cohort a, measured at time t. where $College_{icat}$ or $Secondary_{icat}$ are dummies indicating whether having completed college or high school degree, the baseline is no degree.

Looking at different cohorts, allowing them to be imperfect substitutes in production, since the education variables vary in term of education quality-value, across states and over time, I collapse the individual level data at the cohort level, country, survey year. The aggregation of single birth year cohorts into 7-year birth cohorts ensures large enough samples when the cohorts are followed on a year-to-year basis. Moreover, this definition is fine enough to group individuals who attended elementary and secondary school together, and that were subjected to similar influences from the educational and economic environments (for example school quality and expected gains to an additional year of education). I work with the cell means of

the log annual net earnings and the other variables (weighted by the corresponding cell sizes), to explore whether there are differences among people of the same age in different points in time.

The cell level model on which cohort estimates are based on is the following:

$$\bar{y}_{cat} = \alpha + \beta_1 E\bar{D}U_{cat} + \beta_2 \bar{X}_{cat} + \mu_{ct} + \lambda_{at} + \theta_t + \gamma_c + \chi_y + u_{cat} \tag{1.9}$$

where $E\bar{D}U$ is a vector containing the dummies variable for different degrees. To account for group specific error components, I cluster standard errors at country, gender and wave level.

1.5.2 The sources of rising inequality

In section 1.3 I have presented the theoretical model on which I draw to analyze the leading proximate causes of overall and between-group wage inequality.

Taking the supply, demand and institutions framework to the data, recalling from equation (1.5)

$$lnw = \rho \left(\frac{\alpha_{hct}}{\alpha_{lct}} \right) - \frac{1}{\sigma} ln \left(\frac{H_{ct}}{L_{ct}} \right) \tag{1.10}$$

This equation suggests an explanation of relative wage movements made of both market factors and institutional factors.

Supply is assumed to be observable, the unknowns are the elasticity of substitution and the skill bias technical change that can be both seen as demand shifts. As frequently done in the literature, to control for changes in the demand conditions, I proxy the shift D_{ct}, with a demand index [17], time trends and a measure of technology -R&D intensity.[18]

The idea is that all these measures increase relative productivity in the skill intensive sectors, I thus expect a positive coefficient in my estimations.

To check which are the potentially relevant institutional factors, I include controls for union density, minimum wage, employment protection, Gini index and a measure

[17] This demand index is similar to the demand index used by Katz and Murphy (1992) which is based on the changes in the relative employment.

[18] Ratio of R&D expenditure over value added in the manufacturing sector measured every year in each country.

of the public sector employment.[19]

The model I estimate is the following:

$$ln\left(\frac{w_{ct}^{H}}{w_{ct}^{L}}\right) = \gamma_0 + \gamma_1 D_{ct} + \gamma_2 ln\left(\frac{H_{ct}}{L_{ct}}\right) + \gamma_3 X_{ct} + \tau_t + \mu_c + \varepsilon_{ct} \tag{1.11}$$

where X_{ct} is a vector of labour market institutions and γ_2 provides an estimate for $1/\sigma$. I control for country fixed effects, time fixed effects and interaction between country and time fixed effects, as well. To get efficient estimates standard errors are clustered at country, cohort and wave level.

Since the focus of this paper is on which is the role of the supply in the evolution of college wage premium, I will conduct separately the analysis for the two set of countries. Certainly, the evolution of the relative supply trend has differed in the two set of countries, therefore, I expect differences in the growth of the college wage premia as well. The model above suggests that the competitive wage of a particular type of worker depends positively on the average rate of technical change (α)- meaning a positive effect on the wage ratio of SBTC, negatively on their relative supply change and positively on their relative product -demand shift (that is associated to the technical change).

Concerning institutional factors, the effect is quite complex. The impact of institutions is generally concentrated in specific parts of the wage distribution. Institutions may affect wage differentials in various ways, depending as well on the elasticity of labour supply and across demographic groups. Moreover, institutions have different effects across industries by changing the incentives for capital investment. and thus affecting indirectly wage inequality. In turn, all the institutions I am exploring tend to compress wages. They improve the outside option of employers or unions more for low skilled groups, strengthening their bargaining position and compressing the skill wage differentials. Concerning unionism, unions increase the wage rates of their members above the level they would achieve in the absence of representation, thus they would favor the low skilled workers inducing inequality to decline. The problem with this argument is that it ignores the effects of union wage policy on non-union wages. If a set of jobs usually performed by a particular

[19]Detailed information on the sources of the institutional data is contained in the Appendix A1.

type of labour is unionized and the employer forced to pay higher wages, the supply of labour to all other jobs done by that type of labour will increase together with a reduction in wages. Therefore, it is not clear if the average wage for the group rises or falls with the increase in union representation. Additionally, it can be that workers with white collar jobs, at the higher end of the wage distribution are very unionized - for example, this is the case of some professional orders in Italy, leading thus to an unclear effect of unions on the wage premium. Minimum wage is another institution which mostly concerns lower skilled workers: a binding minimum wage increases the relative wages of unskilled, thus reducing wage inequality. Minimum wage can impact the wage distribution in several ways: firstly, avoiding employment of workers with productivity lower than the minimum wage. Secondly, preventing firms from pushing down wages for workers with low bargaining power and reducing the heterogeneity at the bottom. Additionally, a minimum wage increase leads to an increase in wages for workers paid at the minimum wage level, a weaker increase for workers with wages slightly above the minimum wage (spill-over effects) and little or no effect on high-paid workers (Charnoz, Coudin, and Gaini, 2011). In summary, the presence of a statutory minimum wage by setting an explicit threshold for the lowest wage rate paid tends to reduce wage dispersion. Thanks to its regressive nature, such measure is likely to have a stronger effect at the bottom of the wage distribution rather than at the top. Employment protection policies are often associated with a more compressed wage structure. Following Boeri and Jimeno (2005), I expect Employment protection to protect unskilled workers more than skilled workers, having thus a negative effect on the wage ratio. There is a potential trade-off between EPL and unemployment benefit which may be explained by conflicting interests of insiders-outsiders and low-high skilled. More educated labour force leads to more unemployment policies and less job protection, that is why I assume that EPL is a more favourable measure for low skilled workers.

In turn, accepting the hypothesis that the effects of institutions on the outside option of workers are mostly in favor of the unskilled, then I expect a negative effect of the aggregate institutional measures on the relative wage.

In addition to this standard set of labour market institutions, I add a measure of the public sector pervasiveness -relative percentage of the population working

on the public sector. Public sector employment is perceived as safer and offering more benefits, for this reason, more risk averse individuals sort into public sector employment.[20] However, it seems to be the case that workers at the lower tail of the wage distribution benefit more from public sector employment than workers at the upper tail of the wage distribution. Actually, there is evidence that there can be a wage penalty for highly qualified employees - see for example (Melly, 2005). The idea is that public sector employment may have acted to offset the widening wage inequality seen in recent years and to narrow the college wage premium.

I also control for type of contract: whether full time or part time contract and whether permanent or fixed term employment. These are measures that are somehow related with job stability and job protection and can thus be relevant in assessing wage inequality. Since it is plausible that both market and institutional factors alter the wage distribution both across skill groups and across age groups, data are aggregated by country, year of the survey and age group.

This model, including cross country differences in the role of labour institutions, does a reasonable job accounting for trends in skill premium, however some questions rest unsolved. [21]

The main general concern of this model is that relative skill supply are predetermined, thus labour supply of each group is inelastic. In particular nowadays, this assumption may not hold. In this sense, a first issue to address is, indeed, the one of immigration. It is likely that, since immigrants, on average, are less educated than natives, changes in immigration flows during years affected the relative skill supplies, having as well an impact on college wage premium. Hence, it is important to understand how much of the change in skill supplies have come from changes in immigration and how much is stemming from changes in the native population. The first and most common presumption is that immigration greatly increases the premium to skill, as immigrants increase the supply of less educated people. However, following the reasoning of Goldin and Katz (2009), immigration is found (in the US) not to be so relevant in determining the relative skill supplies having a modest

[20] This is shown to be the case in Germany by (Pfeifer, 2011)

[21] First of all, is technology or relative supply really exogenous? There could be, indeed, trade induced demand or a supply-induced demand. Another potential issue that should be consider is the polarization/ non-monotonicity of jobs. The phenomenon for which middle skilled group is losing demand to both high and low skilled.

impact on the wage premium. The main reason can be found in the change of the educational distribution of more recent migrants: in the recent period immigrants can be distributed at both the very top or the very bottom of the educational ladder.[22] To avoid problems stemming from the possible misreporting of educational information about migrants, I select my sample on native people. However, in many European countries, in particular in many countries belonging to the subgroup of the "low relative supply countries" - i.e. Spain, Italy, UK, migration is a very important and massive phenomenon, it is possible, that it has an effect on the relative supply of college graduates and thus on college wage premium.[23]

Previous literature focuses on the relation between relative supply and college wage premium without considering the potential endogeneity of the relative supply. Without taking this issue into consideration, there is the risk that OLS estimation of the effect of relative supply on college wage premium is inadequate ($\hat{\gamma}_2$ is biased). Theoretically, the bias is negative ($\underset{n\to\infty}{\text{plim}}\,\hat{\gamma}_2 < \gamma_2$) if the errors are negatively correlated or if relative supply is measured with error, and positive otherwise. The assumption that the relative supply of workers is predetermined is plausible in the very short run. Whereas, it is reasonable to think that, in long run, the fraction of workers that chooses to become more educated responds both to innovations that increase the relative demand for more educated labour and to innovations increasing ability premia.

From the individual point of view, given the existing set of possibilities to access education, a worker choose whether to undertake education and to which extent, according to which choice yields him higher lifetime earnings (i.e. according as well to the relative wages he/she expects). Thus, a significant relationship between education attainment, hence relative supply, and some individual outcome may simply result from some unobserved heterogeneity determining both variables. Similarly,

[22] Goldin and Katz (2007), they found that immigration had only a minor impact on the growth in the relative supply of the college graduates and a moderate impact on the high school graduates workers relative to the supply in the 1980-2005 period.

[23] To be sure my results, even if related only to native people, are not biased by the high proportion of migrants existing in some countries, I control for yearly immigration rate by country, and this does not change much the results. Additionally, as a further robustness check, I have controlled for relative migration (i.e. share of college graduate migrants over non-college graduates migrants.) for the countries for which these date are available. Results are in line with previous findings.

the concern can refer to some unobserved country-specific factor that shifts the relative demand for skilled workers, leading to higher relative wages and higher relative employment and confounding the estimation of the inverse substitution elasticity. To overcome these concerns, I use as instrumental variables for the aggregate relative supply ratio, data on the reforms affecting the university system. In particular, I use measures of university autonomy and access, and information on student financing such as financial support.[24] This empirical strategy exploits the differences across countries in the accessibility to tertiary education that are due to changes in institutions and legislations.

1.5.3 Within wage inequality: quantile regressions

Ordinary least squares (OLS) regression estimates the effects of exogenous variables on the mean of the conditional distribution of the dependent variable. On the contrary, quantile regression models allow to characterize the entire conditional distribution of the dependent variable and it allows me to investigate if returns to higher education—and the evolution over time—are dissimilar at different quantiles of the distribution.

This method becomes very useful to investigate the progress of the impact of schooling on within-levels wage inequality. Quantile regressions are able to compare returns to secondary education and to college for the skilled and unskilled workers, conditional on their schooling and experience.

The quantile regression model is the following:

$$lnw_i = x_i\beta_\theta + u_{\theta i} \tag{1.12}$$

With $Quant_\theta(lnw_i|x_i) = x_i\beta_\theta$

Where $x_i$and β_θ are the vector of exogenous variables and the vector of parameters respectively.

$Quant_\theta$ indicates the $\vartheta_j th$ conditional quantile of ln$w$ given x. The $\vartheta_j th$ regression quantile,$0 < \vartheta_j < 1$, is the solution of the following minimization problem:

[24]The data used have been kindly provided by Daniele Checchi, Elena Meschi and Michela Braga, who in Braga, Checchi, and Meschi (2011) have constructed a dataset on school reforms occurred in the last century in 18 countries in Europe. See appendix A1 for details about the data.

$$min_{\beta \in R^k} \left\{ \sum_{i:lnw_i \geq x_i\beta} {}_{\theta i}\theta |lnw_i - x_i\beta_\theta| + \sum_{i:lnw_i < x_i\beta} {}_{\theta i}(1-\theta)|lnw_i - x_i\beta_\theta| \right\}$$

That can be also written as:

$$min_{\beta \in R^k} \left\{ \sum_i \rho_\theta |lnw_i - x_i\beta_\theta| \right\}$$

with $\rho_\theta(\varepsilon)$ is the check fuction defined as $\rho_\theta(\varepsilon) = \theta\varepsilon$ if $\varepsilon \geq 0$ or $\rho_\theta(\varepsilon) = (\theta - 1)$ if $\varepsilon < 0$.

Basically, this technique provide pictures of different points of a conditional distribution. Since it is very informative knowing if the relationship between the regressors and the independent variables varies across its conditional distribution, this methodology has been used in the returns to education literature to assess the possible impact of schooling upon inequality, through its within-levels inequality component. The rationale goes as follows: If the earning increments that stem from schooling (a certain degree) were the same across the wage distribution, then this would mean that schooling (the degree) would not impact upon within-levels wage inequality. This is a consequence of the fact that distributions of wages conditional on different levels of schooling (degree) would differ only on their locations and not on their dispersions. However, it may be the case that these dispersions do indeed vary across educational levels, thus resulting in an impact of schooling upon the wage distribution, through its within-levels channel. I will test this last possibility by using quantile regression estimates of different returns for different degrees.

The empirical results are obtained regressing:

$$lny_i = \alpha_\theta + \beta_{\theta 1} College_i + \beta_{\theta 2} Secondary_i + \delta_{\theta 1} EXP_i + \delta_{\theta 2} EXP^2 + u_i \qquad (1.13)$$

where θ is the quantile being observed.

1.6 Results

In this section the results of the empirical analysis are shown. In the first two subsections I will present the evidences of the evolution of the returns to college, for the entire sample and by age cohorts. The third subsection deals with the potential sources of inequality. Finally, the last subsection repeats the analysis using quantile regressions in order to focus on the evolution of within inequality.

1.6.1 Returns to college results

Table 1.3 shows the results for each region and each dataset, separately for males and females. In this table year effects are shown. Panel A of table 1.3 covers the period from 1994 to 2001, ECHP dataset, whereas panel B covers the period from 2004 to 2009, EU-SILC data. All results stem from from separate regressions for men and women of the log annual net wage on education categories, a quadratic in experience, interactions between education and time, country and time, country, time and age cohorts fixed effects (See section 1.5). Errors are clustered at country, cohort and wave level. The baseline education category is low educational attainment (i.e. ISCED level 1-2). The log of wages of each education group presents trends which differ across the education groups, gender and regions. In general, simple returns to post secondary education have continuously decreased over time for both males and females. The decline is significant and more marked for high relative supply countries. Furthermore, the fall is much clearer in the first half of the period analyzed (1994-2001) for both men and women. However, this is relative to low educated people. When considering the college wage premium - the difference between college and secondary school graduates, to have an idea of its evolution, returns to secondary school should be considered as well. Concerning the evolution of secondary[25] school degree, it seems that, on average, with the exception of women in low relative supply countries, the returns to secondary school degree have remained quite stable over the period analyzed. This can be seen as a confirmation of the observation of the declining college wage premia in high relative supply countries. The inequalities between education groups-adjusted for the level of experience- are

[25]coefficients are omitted for simplicity, but the full table is available upon request.

therefore decreasing over the period. This decline in between-education group inequalities can be observed by examining the degree premiums relative to no degree (see Figure 1.4). For women the decline is less evident but it is still noticeable in high relative supply countries: college returns are declining significantly, even more strongly than for men, in the first half of the period, while this decline is less strong in the second half (EU-SILC data). For women in low relative supply countries, it seems that the returns to both college degree and secondary schooling are more or less stable across waves.

Table 1.3. OLS estimates of the returns to higher education for workers aged 20-55 (1994-2009).

Panel A: ECHP				
	(1)	(2)	(3)	(4)
	MALES		FEMALES	
VARIABLES	High relative supply	Low relative supply	High relative supply	Low relative supply
college	0.370***	0.366***	0.531***	0.465***
	(0.029)	(0.032)	(0.037)	(0.041)
secondary	0.178***	0.166***	0.284***	0.324***
	(0.019)	(0.021)	(0.034)	(0.033)
College 1995	0.019	-0.041	0.029	-0.009
	(0.038)	(0.044)	(0.051)	(0.054)
College 1996	-0.015	-0.048	-0.033	-0.069
	(0.036)	(0.040)	(0.046)	(0.051)
College 1997	-0.033	-0.049	-0.038	-0.000
	(0.036)	(0.041)	(0.045)	(0.053)
College 1998	-0.068*	-0.037	-0.081*	0.041
	(0.037)	(0.043)	(0.046)	(0.055)
College 1999	-0.072*	-0.038	-0.141***	0.033
	(0.037)	(0.042)	(0.049)	(0.054)
College 2000	-0.109***	-0.032	-0.137***	0.016
	(0.036)	(0.045)	(0.046)	(0.055)
College 2001	-0.118***	-0.022	-0.111**	0.055
	(0.035)	(0.043)	(0.046)	(0.053)
Observations	62,512	51,885	51,166	36,950
R-squared	0.429	0.507	0.314	0.371
Panel B: EUSILC				
college	0.306***	0.304***	0.458***	0.433***
	(0.021)	(0.023)	(0.027)	(0.027)
secondary	0.118***	0.170***	0.180***	0.298***
	(0.021)	(0.014)	(0.024)	(0.020)
College 2005	-0.026	0.044	-0.004	0.023
	(0.028)	(0.033)	(0.037)	(0.038)
College 2006	-0.044	0.021	-0.021	0.012
	(0.027)	(0.031)	(0.032)	(0.039)
College 2007	-0.061**	0.021	-0.017	-0.036
	(0.030)	(0.032)	(0.036)	(0.035)
College 2008	-0.087***	0.027	-0.057	-0.025
	(0.030)	(0.033)	(0.038)	(0.034)
College 2009	0.002	0.001	-0.005	0.040
	(0.025)	(0.031)	(0.033)	(0.035)
Observations	57,688	63,947	50,808	52,047
R-squared	0.354	0.357	0.266	0.241

Notes: each regression contains country fixed effects, year fixed effects, controls for age cohorts, interactions country and cohorts. Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

One interpretation of these OLS estimates is that the increase in the aggregate

relative supply of college workers can be responsive to the declines in relative wages.

1.6.2 Cohorts results

Changes in college/high school wage gap have diverged a lot over the last decades according to different age/experience groups. Drawing on Card and Lemieux (2001), to the extent that workers with similar education but different age/experience are imperfect substitutes in production, it is reasonable to expect age cohort specific relative supply to have an impact on the evolution of the college wage premium by age/experience. For this reason, to estimate the existence of cohort effects, I run different regressions for the college wage premium by different experience groups.

As said before, among the reasons behind the drop in the returns to college education (and education in general) there are demand and supply explanations, together with a non market one, that is a combination of institutional factors and economic cycle. Looking from the firm side, it is known that there is a reduced human capital investment after financial recessions: hiring on temporary contracts, offering no on-the-job training, lower education wage premia, lower incentives to investment also in formal education. Since in 2007 there has been the beginning of the financial crisis, it is reasonable to expect a massive drop in the wages for people entering the labour markets around this wrong moment, they somehow represent a lost generation.

To look at the evolution of the returns to college by cohorts in different points in time, I take the microdata, collapse them into cells defined by birth cohort, country and wave, separately by gender, weight by cell sizes, and estimate the college premium by cohort group. Table 4 and 5 provide a breakdown by cohort and by survey for the two regions analyzed, allowing the college premium to vary by cohort groups.

I split across three cohort groups in two subsample periods corresponding to the two datasets: People aged 43-50, the old, the middle aged: 34-42, and the young aged 25-33. I contrast these groups with the corresponding age balanced birth cohort groups in the EU-SILC subsample period 2004-2009, observed ten years later than individuals in the first period, who were born ten years later -i.e. at the same age as their 1994-2001 subsample counterparts). It is clear that the simple

analysis portrayed above masks important changes by cohort and region. Firstly, it is noticeable that returns are always lower, in absolute terms, for the young and higher for the old, no matter the region with high or low relative supply of graduates. Furthermore, there is evidence that returns have declined over time for older graduates in countries with high relative supply of graduates, for younger workers, returns to college are significantly lower than for the older workers, however they seem to be increasing over time. The coefficient of the returns to college for the EU-SILC dataset is higher and significantly different from the same coefficient measured 10 years earlier. However, also secondary school returns have increased quite a lot for the young, leading to an overall negative effect on the college wage premium. Vice versa, returns have hardly changed for both graduates and non graduates in region with lower relative supply of workers.

Table 1.4. The returns to higher education by cohorts. High relative supply countries.

	(1)	(2)	(3)	(4)	(5)	(6)
	Age 25-33		Age 34-42		Age 43-50	
VARIABLES	ECHP	EUSILC	ECHP	EUSILC	ECHP	EUSILC
college	0.268***	0.478***	0.477***	0.314***	0.502***	0.0927
	(0.0567)	(0.0946)	(0.0470)	(0.0905)	(0.0406)	(0.102)
secondary	0.163***	0.326***	0.168***	0.0741	0.114**	-0.0144
	(0.0586)	(0.0639)	(0.0513)	(0.0708)	(0.0537)	(0.0620)
gender	0.248***	0.237***	0.392***	0.383***	0.422***	0.360***
	(0.00978)	(0.0110)	(0.00810)	(0.00934)	(0.00865)	(0.00879)
Observations	918	720	918	720	816	640
R-squared	0.843	0.929	0.869	0.939	0.870	0.930
T-test of differences between College Eusilc and Echp						
[p-value]	[0.053]		[0.000]		[0.000]	
T-test of differences between Secondary Eusilc and Echp						
[p-value]	[0.056]		[0.082]		[0.112]	

Notes: each regression includes controls for experience and experience suqared, country dummies and year dummies. Clustered country by wave and year of birth standard errors within parentheses and p-values within brackets. *** p<0.01, ** p<0.05, * p<0.1

Table 1.5. The returns to higher education by cohorts. Low relative supply countries.

	(1)	(2)	(3)	(4)	(5)	(6)
	Age 25-33		Age 34-42		Age 43-50	
VARIABLES	ECHP	EUSILC	ECHP	EUSILC	ECHP	EUSILC
college	0.318***	0.206*	0.668***	0.612***	0.655***	0.487***
	(0.109)	(0.122)	(0.0872)	(0.122)	(0.0821)	(0.172)
secondary	0.220***	0.196***	0.300***	0.447***	0.506***	0.583***
	(0.0707)	(0.0699)	(0.0574)	(0.0796)	(0.0666)	(0.0859)
gender	0.240***	0.195***	0.335***	0.358***	0.336***	0.339***
	(0.0116)	(0.0147)	(0.0105)	(0.0105)	(0.0101)	(0.0112)
Observations	666	522	666	522	592	464
R-squared	0.888	0.929	0.902	0.928	0.894	0.930
T-test of differences between College Eusilc and Echp						
[p-value]	[0.483]		[0.006]		[0.370]	
T-test of differences between Secondary Eusilc and Echp						
[p-value]	[0.808]		[0.000]		[0.466]	

Notes: each regression includes controls for experience and experience suqared, country dummies and year dummies. Clustered country by wave and year of birth standard errors within parentheses and p-values within brackets. *** p<0.01, ** p<0.05, * p<0.1

1.6.3 The sources of rising inequality

Certainly, the different evolutions of wage distributions are also driven by different labour market structures in the countries analyzed, and to the dissimilar interactions between economic shocks and institutions. To investigate the proximate causes of the inequality, I regress the college wage premium on a set of variables including proxy for demand and supply and some institutional indicators. The idea is to identify which are the main drivers and whether they act in different way in different set of countries. The estimation results are presented in table 1.6 and 1.7, for high and low relative supply of graduates countries, respectively. All the standard errors are clustered by country, age cohort and wave to allow for any possible correlation in the unobservables of individuals of the same age in the same country.

Results show that together with demand and supply factors, also institutions can matter. The first column of tables 1.6 and 1.7 uses the original specification of

Katz and Murphy (1992) with only relative demand and supply measures included as explanatory variables. In what follows, I add in each column some measure of institutional constraints. In column 2, I add controls for minimum wage, employment protection legislation and union density. Column 3 includes a dummy indicator for having a full time contract, column 4 incorporates an alternative measure of the relative demand-R&D intensity. Finally, in the last column, I add the Gini index and the percentage of people working in the public sector. While in both regions, the coefficients for the relative supply variable are the ones expected, i.e. negative and significant, this is not the case for the relative demand index. The coefficient of the relative supply indicator is slightly higher in countries with lower supply of graduates (-0.008 vs. -0.014). In addition to this, countries with high relative supply of skilled workers present a higher and more significant relative demand indicator. Concerning the relative demand, high relative supply countries have positive and significant coefficients, although very low. Also using an alternative measure of demand (R&D intensity) gives the same result. This result is consistent with a naive SBTC story. This suggest that, despite the higher increase in the supply, these countries have still "space" for skilled workers since there is still a role for the relative demand to push their premium. For countries with lower relative supply, none of the demand measures appear to be a significant determinant of wage inequality.[26] The negative and significant coefficient of the dummy for male (gender) is not surprising. It is well known indeed that on average, there is much more selection for women into education rather than for men. A higher college wage premium for women is a common finding in the literature.

A compelling explanation for the evolution of between and within group wage inequalities is the role of institutions. The institution constraints' coefficients are expected to have mainly a negative sign which would suggest that these policies affect unskilled more that skilled workers. Minimum wage is not a significant determinant of wage inequality in high relative supply countries, whereas it is the case in countries with lower relative supply of graduates. A one percent increase in the

[26]To compare these results with others in the literature, referring to Autor, Katz, and Kearney (2008), I also included a time trend as a proxy for the demand for high skilled workers: a positive coefficient would be interpreted as a sign of SBTC. What I find is that the sign is not always positive neither significant, confirming the lower effect of the demand in contrast to the relative supply.

minimum wage lowers the college wage premium by around 3%. Employment protection legislation is significantly and negatively correlated with wage inequality in low relative supply countries but it looses significance in high relative supply countries. Union density does not seem to matter in high relative supply of graduates countries, however, although with a very low coefficient, it is negatively and significantly correlated with wage premium in low relative supply countries. Full time contracts seem to be good instruments to reduce wage inequality, in particular in high relative supply countries. Employment in public administration is negatively and significantly correlated with wage inequality, however the effect is slightly higher in low relative supply of graduates countries, countries in which the percentage of public employment is lower.

Consequently, it emerges that increases in the minimum wage, in full time contracts and employment protection also provide a valid explanation for the decrease in within-inequalities for the less-educated workers and the decreasing trend in lower-tail inequality over the period, regardless of educational level. Eventually, in addition to the supply and institutions story as an explanation for the declining evolution in college premium, another possible one is the economic cycle. Even if it has been shown that, during the Great Recession (2008-2009), there has been a much larger labor market response in the US rather than in Europe, the crisis has affected European labour market as well . Unemployment could also be a part of the story, as argued in Autor, Katz, and Kearney (2008): selection into unemployment could shift to the right the distribution of unobserved skills and of wages. However, adding unemployment rate and relative unemployment of skilled to unskilled people to the wage inequality regression does not change remarkably the results.[27]

As already said, this model is doing a good job in capturing the general trend, however it suffers for a potential endogeneity problem. Assessing the potential endogeneity of the relative supply, that is the relative share of the labour force with tertiary education relative to the share of the labour force with high school diploma, I instrument relative supply using the set of tertiary education institutions. Table 1.8 shows first stage estimates of the IV strategy for the relative supply: relative supply is regressed on the indicators measuring the variation in the tertiary

[27]Results are omitted but are available upon request.

Table 1.6. The college wage premium, age groups. High relative supply countries.

	(1)	(2)	(3)	(4)	(5)
Relative Supply	-0.00824***	-0.00933***	-0.0104***	-0.0101***	-0.00947***
	(0.00225)	(0.00251)	(0.00253)	(0.00252)	(0.00251)
Relative Demand	0.00997*	0.0133*	0.0109	0.00448	0.00718
	(0.00459)	(0.00662)	(0.00665)	(0.00691)	(0.00903)
gender	-0.00196***	-0.00198***	-0.00174***	-0.00168***	-0.00171***
	(0.000193)	(0.000198)	(0.000218)	(0.000217)	(0.000217)
Log Minimum Wage		-0.0111	-0.00954	-0.00528	0.0170
		(0.00917)	(0.00915)	(0.00919)	(0.0123)
EPS		0.000116	0.000111	-0.000237	0.000369
		(0.000441)	(0.000440)	(0.000450)	(0.000529)
Union density		-0.0000505	-0.0000444	-0.0000344	-0.00000639
		(0.0000327)	(0.0000327)	(0.0000326)	(0.0000342)
Full time			-0.00299**	-0.00319**	-0.00246*
			(0.00114)	(0.00113)	(0.00115)
R&D man.				0.000487**	0.000709***
				(0.000153)	(0.000168)
Gini					-0.00886
					(0.0112)
Public Emp.					-0.0385*
					(0.0157)
Observations	795	795	795	795	795
R^2	0.342	0.345	0.351	0.360	0.368

Notes: Controls for country and year fixed effects. Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.7. The college wage premium, age groups. Low relative supply countries.

	(1)	(2)	(3)	(4)	(5)
Relative supply	-0.0145***	-0.0117**	-0.0105**	-0.0103**	-0.0105**
	(0.00372)	(0.00380)	(0.00385)	(0.00385)	(0.00386)
Relative demand	0.00325	0.00608*	0.00609*	0.00580	0.000960
	(0.00278)	(0.00296)	(0.00295)	(0.00296)	(0.00348)
gender	-0.00229***	-0.00224***	-0.00209***	-0.00209***	-0.00212***
	(0.000154)	(0.000152)	(0.000174)	(0.000174)	(0.000173)
Log Min wage		-0.0328***	-0.0312***	-0.0300***	-0.0224*
		(0.00774)	(0.00778)	(0.00783)	(0.00877)
EPS		-0.000825	-0.000791	-0.00103*	-0.00116*
		(0.000477)	(0.000477)	(0.000509)	(0.000561)
union density		-0.000170**	-0.000135*	-0.000178*	0.0000591
		(0.0000595)	(0.0000626)	(0.0000701)	(0.000114)
Full Time			-0.00151	-0.00149	-0.00119
			(0.000850)	(0.000850)	(0.000859)
R&D man.				-0.000327	-0.0000930
				(0.000241)	(0.000288)
Gini					-0.0135
					(0.00853)
Public emp.					-0.0693*
					(0.0284)
Observations	620	620	620	620	620
R^2	0.425	0.446	0.449	0.451	0.458

Notes: Controls for country and year fixed effects. Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

education reforms, measured five years before. The underlying assumption is that, in order for these reforms to take action, being implemented, and to affect the relative supply, it take an average of five years.[28] Therefore, the level of tertiary education in a particular year, in a specific country is deemed to be affected by the level of institutional set-up of tertiary education five years before.

In all specifications, the instruments are shown to be good explanatory variables for aggregate relative supply, in both the two set of countries, as they are mostly significant at any conventional level. However the size and the relevance of the used instruments differs in the two set of countries. At the bottom of the table, we report the F-statistic of the excluded instrument. It oscillates between 20 and 120, well above the conventional threshold of 10 for strong instruments. Therefore, there should be no concerns about potential biases in the second stage due to the use of weak instruments.

The second stage results for high relative supply countries and for low relative supply countries are presented in table 1.9 and 1.10, respectively. I compare OLS and IV estimates of the college wage premium, where I replace relative supply with a set of instruments measuring country variation in the institutional set-up characterizing tertiary education. More specifically, column 1 and 2 show the baseline (Katz and Murphy) specification where college wage premium is regressed on a demand index and on a supply index. Columns 3 and 4 add labour market institutions such as minimum wage, EPS and union density as additional controls.[29] In all cases the estimated IV coefficient of relative supply are negative, strongly significant and larger in magnitude than the OLS. According to these estimates, the OLS coefficient of relative supply is -0.07 in the preferred specification in high relative supply countries, and -0.017 in countries with low relative supply of graduates. The IV estimates are substantially larger in both the set of countries and the specifications (-0.011 and -0.036 respectively for high and low relative supply countries), implying a positive bias. The Cragg-Donald Wald F-statistics confirms that instruments are strong predictors of the relative supply as we already know from the regressions in Table

[28]For this reason the sample observed is partially reduced a delimited to 2005, since the data on the tertiary education institutions arrive up to 2005.

[29]The richer specification -i.e. the one including the other controls used in the OLS estimations, such as the Gini coefficient, public employment, R&D intensity and full time contract, has been omitted since these variables do not appear relevant.

1.8. Additionally, in the IV estimates, the sign and the significance of the coefficients of the labour market institutions are very close to what has been found in the original OLS estimates. The most relevant institution is the minimum wage in countries with lower relative supply of graduates, this has a negative and significant effect - of a very similar size of the OLS one, on the college wage premium. A few conclusions can be drawn from these set of estimates. First, there is clear empirical evidence that being exposed to higher relative supply of graduates has caused a reduction in the college wage premium, that is the relative advantage of the relatively higher educated people. Second, the comparison between OLS and IV estimates suggest that the OLS estimates are upward biased.

Table 1.8. Relative supply equation: 1st stage

	High Relative Supply Countries		*Low Relative Supply Countries*	
Expansion of uni. accessibility	0.128^{***}	0.090^{***}	0.020^{***}	0.024^{***}
	(0.013)	(0.015)	(0.005)	(0.006)
Selectivity in uni. access	-0.020^{**}	-0.000	0.055^{***}	0.052^{***}
	(0.006)	(0.008)	(0.004)	(0.005)
Financial support	-0.027^{**}	-0.031^{***}	0.035^{***}	0.040^{***}
	(0.009)	(0.009)	(0.002)	(0.003)
Size of grant	0.076^{***}	0.062^{***}	-0.029^{***}	-0.022^{***}
	(0.006)	(0.007)	(0.004)	(0.004)
Loan to grant component	-0.034^{***}	-0.020^{*}	-0.036^{***}	-0.044^{***}
	(0.010)	(0.010)	(0.004)	(0.004)
Interest rate	-0.062^{**}	-0.083^{***}	0.034^{***}	0.042^{***}
	(0.020)	(0.019)	(0.008)	(0.008)
Index of university autonomy	0.056^{***}	-0.009	0.053^{***}	0.067^{***}
	(0.014)	(0.020)	(0.012)	(0.016)
Year FE	*No*	*Yes*	*No*	*Yes*
R-squared	0.332	0.429	0.665	0.695
Observations	545	545	450	450
F-stat	38.19	24.81	125.29	61.58
F-stat p-value	0.000	0.000	0.000	0.000

Notes. The dependent variable is relative supply of graduates. All regressions include a full set of year dummies. Robust standard errors in parenthesis. One star means 5% significantly different from zero, two stars 1%, three stars 0.1%.

Table 1.9. Assessing the endogeneity bias- High relative supply countries

	Baseline model		*+ Labour Market Institutions*	
	OLS	*IV*	*OLS*	*IV*
Relative Supply	−0.000	0.003	−0.007*	−0.011*
	(0.002)	(0.004)	(0.003)	(0.005)
Relative Demand	0.005***	0.005***	0.006***	0.006***
	(0.001)	(0.001)	(0.001)	(0.001)
gender	−0.002***	−0.002***	−0.002***	−0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
Minimum Wage			0.003	0.003
			(0.005)	(0.005)
EPS			−0.000	−0.000
			(0.000)	(0.000)
Union Density			−0.000***	−0.000***
			(0.000)	(0.000)
Cragg-Donald Wald F statistic		39.695		73.143
R^2	.277	.276	.311	.310
N	545	545	545	545

Notes: The dependent variable is college wage premium. Relative supply is instrumented by a set of indicators measuring tertiary education reforms: selectivity in university access, expansion of university access, financial support, increase grant size, loan component to grant component, interest rate and an index of university autonomy. All regressions include a full set of year, country and age cohort dummies. Robust standard errors in parenthesis. One star means 5% significantly different from zero, two stars 1%, three stars 0.1%.

Table 1.10. Assessing the endogeneity bias- Low relative supply countries

	Baseline model		*+ Labour Market Institutions*	
	OLS	*IV*	*OLS*	*IV*
Relative Supply	-0.018^{***}	-0.020^{***}	-0.017^{***}	-0.036^{**}
	(0.003)	(0.003)	(0.005)	(0.011)
Relative Demand	0.003^{***}	0.003^{***}	0.006	0.008^{*}
	(0.000)	(0.000)	(0.004)	(0.004)
gender	-0.002^{***}	-0.002^{***}	-0.002^{***}	-0.002^{***}
	(0.000)	(0.000)	(0.000)	(0.000)
Minimum Wage			-0.025^{**}	-0.027^{**}
			(0.010)	(0.010)
EPS			0.000	0.000
			(0.000)	(0.000)
Union Density			-0.000	-0.000^{*}
			(0.000)	(0.000)
Cragg-Donald Wald F statistic		230.866		17.596
R^2	.386	.385	.413	.3963
N	450	450	450	450

Notes: The dependent variable is college wage premium. Relative supply is instrumented by a set of indicators measuring tertiary education reforms: selectivity in university access, expansion of university access, financial support, increase grant size, loan component to grant component, interest rate and an index of university autonomy. All regressions include a full set of year, country and age cohort dummies. Robust standard errors in parenthesis. One star means 5% significantly different from zero, two stars 1%, three stars 0.1%.

1.6.4 Quantile regressions results

The divergence of the upper and the lower tail wage inequality and the convexification of the returns to education is a puzzle.

To look at the inequality within educational groups, in particular, college, I run quantile regressions. As already said in section 5, this technique allows me to look at different earning advantages of college degree at different deciles of the income distribution, so allowing to look at changes in wage distributions and heterogeneity of the skill premia.

Figure 1.6. Within group inequality -higher education. Quantile regressions

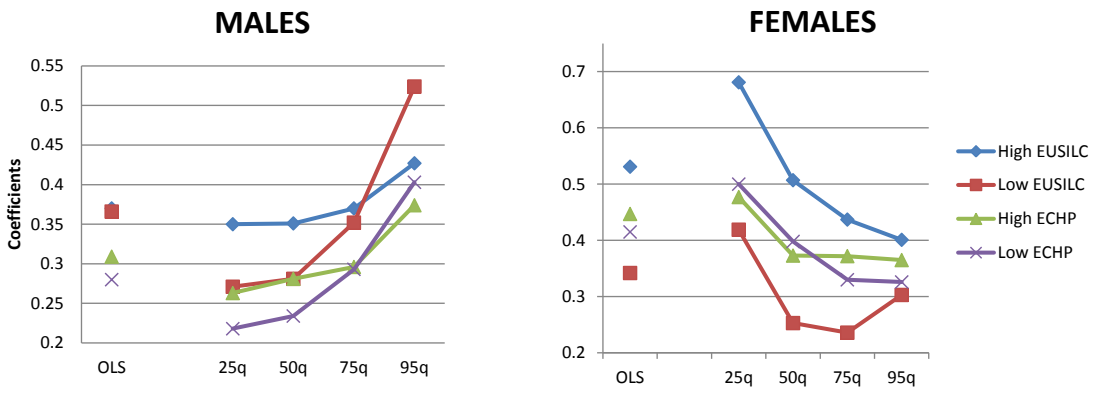


Figure 1.8 displays the log wage premium estimates for college degree for males and females in the two regions in the two dataset.[30] The premia of the high-skilled workers have increased for males over the distribution. For males, both in high and low relative supply countries, it is possible to notice a rise in upper tail inequality. Hence, despite the increase in the access to education, inequality is still increasing in both set of countries, however this increase is declining over time, especially for counties with higher relative supply. The gap between ECHP and EU-SILC estimates is reducing over time. Looking at the same estimates, focusing on the returns to secondary school, the increase in inequality is is much less marked (see figure A1). This is in line with previous findings (i.e. Martins and Pereira (2004))

[30]The same figure for high school graduates is shown in the appendix A2, Figure A1.

who fund increasing wage inequality within higher educated. This pattern is not observed for women, for them within-education group inequalities are decreasing over the distribution. Decreases are typically stronger for women in high relative supply countries. In particular, college returns are decreasing over the distribution and over time in low relative supply countries, for both college and secondary school degree. For countries with higher relative supply returns to both degree are still decreasing over the distribution but are slightly increasing over time. This is quite reasonable thinking of the fact that for women there is much more selection into education.

1.7 Conclusions

There has been much debate about the contribution of the increase of higher education participation to the widening wage inequality in the US. However, this has been less explored in Europe.

This paper aims at analyzing changes in the wage premium associated with a degree using a large European dataset obtained harmonizing two different sources. More specifically, I am interested in how the college premium evolved across time, across the wage distribution and across cohorts. I try to offer some insights into this topic by looking at the supply and demand for skills -in particular of graduates over time. I allow different education types to yield different returns in order to assess whether the decline in the returns to education is limited to specific skill groups. I analyze the effects of the recent strong increase in the value of the participation rates on returns to college and inequality in Europe, using cross country variation in relative supply, demand and labour market institutions to look at their effects on the trend in the college wage gap. I investigate the sources of inequality looking at both supply/demand and institutional components. As a final extension, to get a more comprehensive picture, I go through the inequality within education groups: quantile regressions allow me to look at the earning advantage of additional years of schooling at different deciles of the income distribution.

Although the literature does not provide much evidence that, on average, the college premium has shown any significant trend changes in recent years, despite the

large increase in the participation rates and in the flow of graduates into the labour market, my results show that there has been a fall in returns in the recent years, in particular for youngest cohorts, for both men and women. This fall has been more marked in countries with higher supply of skilled workers. I use harmonized micro data from two different sources (ECHP and EU-SILC) to construct a dataset which covers 15 years. I divide the countries into two different subgroups: countries with high relative supply of graduates at the beginning of the period analyzed (1994) and countries with low relative supply of graduates. The reason why I am doing so is that I argue the two set of countries, facing different evolution in the relative supply over time, have faced different evolutions in the college wage premium as well. Empirically, I find some evidence of a significant decline of college returns in countries with high relative supply of graduates and a marked fall in returns for recent cohorts for both men and women in all European countries. This decline is less evident in countries with low relative supply of graduates. A potential explanation of these findings is indeed the increase in the educational attainment over the period. The fall in the skill premium is intuitively the first outcome of a classic supply and demand effect. In particular, in high relative supply countries, i.e. countries with higher supply of skilled workers, it could be that the demand was not able to compensate for the increase in labour supply of skilled workers. To check for this I have looked at the potential sources of wage inequality, including supply and demand factors as well as institutional indicators. I address possible concerns of endogeneity of relative supply by an instrumental variable strategy. The estimates reveal important effect of the increased relative supply of the declining college wage premium. Additionally, institutional constraints such as Employment Protection Legislation, minimum wage and union density are relevant in explaining inequality. Finally, there is some empirical evidence on the role of education in reducing income inequality is not univocal. The main policy implication of these findings is that increasing accessibility to tertiary education in Europe, not only can lower the disparities among different education groups but it can, as well, lower the premia, possibly by the implied changes in ability composition across education groups.

Chapter 2

Lost in Transition? The returns to education acquired under communism in the first decade of the new millennium.

2.1 Introduction

The fall of the Iron Curtain ushered in drastic changes in the economic and education systems of Eastern Europe. The transition from a centrally planned to a market economy affected the lifestyles and the living standards of millions of Eastern Europeans, who faced the transition process being endowed with education acquired under communism and having experienced a completely different economic system and different incentive mechanisms. To what extent is this human capital still valuable in the 2000s? The existing empirical evidence on the returns to education during the transition period indicates that education acquired under communism has not suffered. This evidence, however, has two main limitations. First, empirical work in this area typically covers a single country and the transition years, which were affected by substantial economic turmoil and large labour reallocations. Relatively little is known about the period of EU accession, characterized by the progressive implementation of a functioning free market economy in Central and Eastern Eu-

ropean countries (CEE henceforth). Second, by failing to deal with the correlation between measured education and unobserved talent, most of this empirical research risks to produce inconsistent estimates of the returns to education. In this paper, we contribute to this literature in three directions. First, we estimate the returns to education earned during the first decade of the new millennium - a period when several Eastern European economies entered the European Union - by individuals who completed their education at least 10 years earlier under communism. Rather than using a pre/post comparison, as done by the empirical literature reviewed below, we evaluate these returns in a comparative perspective, by contrasting them with the returns earned during the same period of time by the same age cohorts who were educated and work in Western Europe.[1] While, Western education before 1989 was inspired and influenced by the presence of free market economies in the West, education under communism was designed on a broadly different set of values. One would therefore expect the former type of education to suit better the free market economies of Europe in the past decade, and ultimately to produce higher returns. Second, we use data both for East Germany and for seven CEE countries rather than for a single country, which gives a broader scope to our empirical results. Third, we explicitly recognize that education is correlated with ability, which is typically unobserved. Failure to account for this correlation yields biased estimates (see Card (1999)). We deal with the endogeneity of education in earnings regressions by using a strategy inspired by Card and Rothstein (2007) which consists of a) collapsing micro data into cells defined by gender, age cohort, country and year of the survey; b) differencing out the common gender components of average unobserved ability; c) capturing residual gender specific unobserved effects both with differences in parental education and labour market conditions and with country and time specific age effects, country, time, cohort and country by time dummies. Consistent with the observation that the human capital accumulated by individuals belonging to younger and older cohorts is heterogeneous, we derive our empirical earnings equations from a model where both male and female employment consist of several groups of imperfectly substitutable workers, who differ in

[1] An alternative comparison group would consist of the Eastern Europeans who have been entirely educated after the fall of the Iron Curtain. This group, however, is composed of individuals who are either still at school or in the early stages of labour market participation.

their age. The model allows for positive unemployment and generates for each cell defined by gender, age, country and time of the survey - a relationship between earnings, education and other covariates, which we estimate using both Eastern and Western European data. Differences in the estimated returns to education attained under communism and in the West are likely to be the joint outcome of differences in the type of education on the one hand and differences in labour market institutions and in the structure of the economy on the other hand. In an effort to isolate the former from the latter, we start our empirical investigation with Germany. Nowadays, more than 20 years since re-unification, East and West Germany share the same political and economic institutions and an integrated labour market, with unrestricted worker mobility. Although significant differences in income, wages and unemployment between the regions of Eastern and Western Germany of the country still remain, we believe that Germany is an almost ideal laboratory for our analysis. The key finding of this paper is that the returns earned between 2000 and 2009 by the Germans who completed their education under communism in the former GDR (German Democratic Republic) are not statistically different from the returns obtained by the Germans belonging to the same age groups who studied in the former FRG (Federal Republic of Germany). We investigate whether this result holds in a broader context by extending our comparison to 23 economies in Eastern and Western Europe and find that, reassuringly, it is confirmed. It is tempting to conclude from this that education acquired under communism is as valuable in the 2000s as education acquired in Western Europe. There are, however, two important qualifications. First, close to 30 percent of the individuals in the relevant age groups are not employed in Eastern Europe, compared to 20 percent in Western Europe. Since the unemployed, disabled, retired and engaged in domestic tasks or in the informal economy typically have lower monetary payoffs to their education, average returns to schooling which include these groups are bound to be less favourable for those educated under communism than the returns earned by the smaller group with a job. Second, our estimates rely on the important assumption that each year of schooling yields the same marginal return, independently of the level of education (primary, secondary or tertiary). When we remove this assumption and distinguish between pre and post-secondary education, we find that senior males who have at-

tained only primary or secondary education under communism earn significantly lower returns in the post-transition Eastern European labour markets than equally educated Western Europeans employed in the West.[2] On the other hand, there is evidence that males who have completed post-secondary education under communism enjoy higher payoffs in these markets than similarly educated Western Europeans who are employed in the West. In contrast, we find that senior females earn statistically similar returns in the East and the West for each level of attained education. We believe that the males versus females divide mirrors well the industry versus services divide. Before the fall of communism, low and medium educated males worked mainly in the industrial sector, and females were often employed in the service sector. While industry has been heavily affected by the transition from communist to free market economies, services largely benefitted and rapidly expanded. In summary, our answer to the question of whether education attained under communism is still valuable in the late 2000s is that it is as valuable as the one acquired and used in Western Europe by similar cohorts of individuals, but only for females, college educated males and those who are fortunate enough to be employed. The paper is organized as follows. Section 2 briefly summarizes the main economic effects of the transition from communism to market economies in CEE countries. Section 3 describes the key features of education under communism and Section 4 reviews the relevant literature. Methodological issues and proposed solutions are discussed in Section 5, the data are introduced in Section 6 and the results are presented in Section 7. In the concluding section we summarize and discuss our findings.

2.2 Transition from Communism

Under communism, planned economies in Eastern Europe and the Soviet Union invested a large share of their resources in the development of heavy industry, to support a fast process of industrialization and modernization of a production structure traditionally dominated by agriculture.[3] Workers of large metallurgical and

[2]By,equally educated we mean that they have completed the same years of schooling, independently of the quality of their education.

[3]Important exceptions were former Czechoslovakia and the GDR, where the industrialization process was well under way before communism (Berend and Berend, 2001).

mechanical factories were considered the elite of the proletarians and the vanguard of Marxism. To be able to balance the military power of the US, the focus on heavy industry continued for decades, sacrificing the production of consumption goods and the development of an advanced service sector. Under socialism, labour markets were characterized by complete job security and an egalitarian wage distribution. Wage grids were established, and differences between skilled and unskilled workers were kept small compared to Western standards (Munich, Svejnar, and Terrell, 2005). Wages were higher for blue collar workers, and workers in manufacturing were paid better than workers in services, in spite of lower average education. The fall of the communist regimes and the adoption of a market economy required a radical re-allocation of production factors away from the traditional industries. During the early stages of transition, CEE countries saw a dramatic fall of GDP and employment (Boeri and Terrell, 2002), followed by economic recovery (see Figure B.1 in Appendix B4). These countries lost 22.6 percent of their GDP in the initial phase of output decline, which lasted on average 3.8 consecutive years - only 2 years in Poland, 3-4 years in Hungary, Czech Republic and Slovak Republic and 5-6 years in the Baltic States (Bank, 2002). In most countries, unemployment stagnated at high levels for long time. The average unemployment rate was between 9 and 14 percent in the period between 1992 and 1999 and reached its peak at 14 to 16 percent 3 to 4 years after the beginning of the transition.[4] Real wages declined by about 20 percent during the initial two transition years and then remained stable or slightly increased since 1991. Dramatic changes occurred in the structure of employment between 1989 and 1998. On average, employment in agriculture remained stable, the employment share in industry felt by 10 percentage points and employment in the service sector expanded (Boeri and Terrell, 2002). During the same period, the share of agriculture on GDP remained at 14 percent, industry dropped from 45 to 33 percent and services increased from 41 to 53 percent (The World Bank 2002). From almost non-existent before the fall of the Iron Curtain, the private sector quickly reached 65 percent of all employment with a peak of 80 percent in Hungary by 1997 (Boeri and Terrell, 2002). In terms of GDP shares, CEE countries moved from 11 percent of GDP was generated by the private sector in 1990 to 68 percent in 1999.

[4]The Czech Republic was an exception, with initially low unemployment rates which peaked 9 years after the beginning of the transition at about 10 percent.

The downsizing of heavy industry caused a permanent reduction in the demand for unskilled labour. At the same time, the expansion of skill-intensive services (for example, finance, insurance and real estate, consulting, information services, tourism) dramatically increased the demand for more educated employees (Orazem and M.Vodopivec, 1997). Looking at the supply side, Lamo, Messina, and Wasmer (2010), suggest that the emphasis on vocational education in former communist countries slowed down worker mobility across sectors during the transition period, pushing many middle-age workers towards the exit strategy provided by early retirement programs put in place to facilitate the transformation of CEE economies. This pattern is still recognizable in the late 2000. Table 2.1 shows the percentage of individuals born between 1951 and 1964 who were retired, unemployed or disabled in Western and Eastern Europe during the years 2006 to 2008, more than 15 years since the fall of the Berlin Wall. The percentage is generally higher in CEE countries and approximately equal to 37 percent for those with less than upper secondary education.

Table 2.1. Percentage retired, unemployed or disabled, by educational attainment. Cohorts born between 1951 and 1964. By area and gender. Eastern and Western Europe (without Germany). 2006-2008.

		Males	Females
Western Europe			
	Less than high school	15.7	12.8
	High school	7.9	9.5
	College	4.8	5.2
Eastern Europe			
	Less than high school	36.9	37.5
	High school	21.7	24.3
	College	7.6	8.5

Source: our computations on EU-SILC data. The countries in Eastern and Western Europe are: Bulgaria, The Czech Republic, Estonia, Hungary, Poland, Romania, Slovenia, Slovakia, Austria, Belgium, Denmark, Cyprus, Ireland, Spain, France, Finland, Greece, Italy, Netherlands, Norway, Portugal, Sweden and the United Kingdom.

2.3 Education under communism

Communist nations have made huge efforts over the years to develop their own educational systems. During communism, education was provided free of charge, and stipends were often granted to students. The most well known achievement of communist education was the close to universal access to primary and lower secondary education, with enrolment rates much higher than in countries with comparable levels of development (Micklewright, 1999). Conversely, access to tertiary education was highly rationed. In the Czech Republic, for instance, only about half of the students seeking university admission were accepted (Filer and Munich, 2003).

Table 2.2. Educational attainment: by area and gender. Eastern and Western Europe (without Germany). 2006-2008.

		Males	Females
Western Europe			
	Years of schooling	12.62	11.98
	% with high school	43	40
	% with college	21	20
Eastern Europe			
	Years of schooling	12.46	12.36
	% with high school	76	69
	% with college	11	12

Source: our computations based on EU-SILC data. The countries in Eastern and Western Europe are: Bulgaria, The Czech Republic, Estonia, Hungary, Poland, Romania, Slovenia, Slovakia, Austria, Belgium, Denmark, Cyprus, Ireland, Spain, France, Finland, Greece, Italy, Netherlands, Norway, Portugal, Sweden and the United Kingdom.

Table 2.2 compares the educational attainment of individuals born in Eastern Europe1 between 1951 and 1964, who completed their education under communism, and those born in the West during the same period. While the share of individuals with high school education in CEE economies was relatively high for both genders (76 percent for males and 69 percent for females), the share of college graduates was relatively low (11 percent for males and 12 percent for females in the East against 20 percent or higher for both genders in the West). It is less recognized that education in communist countries was characterized by substantial disparities. The role of parental background in determining children education was as relevant and sometimes more relevant than in the West. In Hungary and Poland, for instance, the children of the highest social class were almost four times as likely as the average person to obtain an academic upper secondary or tertiary qualification. Partly because of this, the dispersion of standardized test scores in Eastern countries was similar to that observed in Western countries (Micklewright, 1999). Overall, the system in former communist countries encouraged students to select vocational curricula and to leave school after completing upper secondary education (Flanagan, 1998). According to Filer and Munich (2003), more than three quarters of Czech

and Hungarian secondary students were enrolled in vocational and technical tracks. Vocational education was extremely specialized, with over 300 separate curricula. In some cases, schools provided students "...with what was in effect merely firm-specific human capital for a local enterprise..." (Micklewright, 1999, p.346). High vocational specialization had the advantage of producing a smooth school to work transition, as skills developed at school were ready for use in industry, and the disadvantage of reducing worker adaptability - a potentially major limitation once firms started to be restructured during the transition to market economies (see Commander and Kollo, 2008).[5] This situation changed abruptly with the transition to a market economy, when the structure of incentives was progressively altered in favour of college and more general education.[6] At the onset of the transition, many doubted that the skills attained during the communist era would have been valuable in the new market economy: the school curricula were often outdated, instruction was based on frontal lessons, too little emphasis was placed on problem solving and independent thinking and too much on memorization (Bal, 2001). In addition, skills which are considered to be important in a market economy, such as strategic, knowing-how-to-learn and problem-solving skills were not well developed (Berryman, 2000). A comparison of the average quality of the human capital obtained by Eastern and Western Europeans by the mid to late 1990s can be carried out by using the International Adult Literacy Study (IALS), an international survey designed to evaluate in a comparative perspective the degree of "functional literacy" of adults. Functional literacy is defined as the ability to understand and employ printed information in daily activities, at home, at work and in the community - to achieve one's goals, and to develop one's knowledge and potential (OECD, 2000). Literacy affects the ability to take independent decisions, individual flexibility and the ability to adapt to new contexts, all valuable attributes in the new market economies emerging after the fall

[5]In Table B.1 in Appendix B4 we show the proportion of unemployed or retired by level and type of education (academic or vocational). The data are from IALS (International Adult Literacy Survey) and cover only a subset of Western and Eastern European countries. Results are broadly in line with those reported in Table 2.1, although IALS data were collected in the mid to late 1990s, before the end of the transition. Remarkably, the proportion of unemployed or retired does not differ much between individuals with academic and vocational education, especially among men.

[6]For East Germany see Riphahn and Trubswetter (2010). Between 1989 and 1997, the share of students enrolled in general education rose from 24 to 28 percent in Hungary, from 22.5 to 32.4 percent in Poland, from 17.8 to 22.1 percent in the Czech Republic and from 18.1 to 25 percent in the Slovak Republic (Berryman, 2000)

of the Berlin Wall. IALS considers literacy in three distinct domains: prose, document and quantitative literacy.[7] Table 2.3 reports average standardized test scores in the three domains[8] for the cohorts born between 1951 and 1964 in Eastern and Western Europe, who completed their education before the fall of the Berlin Wall. Scores were lower in Eastern than in Western countries, although not dramatically so, in all domains. Differences between genders were minor. The gap in average test scores was smaller in quantitative literacy than in prose and document literacy, a result coherent with the emphasis that communist schools placed on technical and vocational education.

Table 2.3. Standardized literacy test scores for Western and Eastern Europeans born between 1951 and 1964, by gender and domain of the test. Source: IALS

Region	*Domain*	*Females*	*Males*
Western Europe	Prose literacy	285	280
	Document literacy	283	290
	Quantitative literacy	282	296
Eastern Europe	Prose literacy	249	242
	Document literacy	250	253
	Quantitative literacy	264	270

Notes: Western Europe includes Ireland, Netherlands, Great Britain, Sweden, Belgium, Italy, Norway, Denmark and Finland. Eastern Europe includes Poland, Slovenia, the Czech Republic and Hungary.

Substantial cross-country heterogeneity in test scores existed in both Western and Eastern Europe, as shown by Table 2.4, where average test scores in quantitative literacy are reported by country and level of education. We distinguish among four

[7]Prose literacy is the knowledge and skills needed to understand and use information from texts including editorials, news stories, brochures and instruction manuals. Document literacy is the knowledge and skills required to locate and use information contained in various formats, including job applications, payroll forms, transportation schedules, maps, tables and charts. Quantitative literacy is the knowledge and skills required to apply arithmetic operations, either alone or sequentially, to numbers embedded in printed materials, such as balancing a chequebook, figuring out a tip, completing an order form or determining the amount of interest on a loan from an advertisement (OECD, 2000).

[8]Western Europe in these data includes Ireland, Netherlands, Great Britain, Sweden, Belgium, Italy, Norway, Denmark and Finland. Eastern Europe includes Poland, Slovenia, the Czech Republic and Hungary.

levels of education: college, academic and vocational upper secondary and less than upper secondary.[9] As expected, test scores increased with educational attainment. With the exception of the Czech Republic, Eastern European countries performed worse than most Western countries at all levels of education and particularly at the lowest level (primary and lower secondary) and at the vocational upper secondary level.

Table 2.4. Standardized test scores - quantitative domain - for Western and Eastern Europeans born between 1951 and 1964, by gender, country of residence and level of education. Source: IALS

		Females				Males			
Region	Country	College degree	ISCED 3 academic	ISCED 3 vocational	Less than ISCED 3	College degree	ISCED 3 academic	ISCED 3 vocational	Less than ISCED 3
Western Europe	Ireland	301	281		236	321	297		253
	Netherlands	314	314	291	270	329	335	306	277
	Sweden	321	301		292	348	321		295
	Great Britain	307	290		253	338	290		267
	Belgium	316	293	280	248	333	313	288	264
	Italy	288	275	260	221	304	291	265	219
	Norway	326	287		276	335	302		283
	Denmark	313	305	296	264	333	337	309	287
	Finland	313	319	287	260	328	316	291	270
	West (average)	311	296	283	258	330	311	292	268
Eastern Europe	Poland	282	276	262	215	304	258	277	227
	Slovenia	299	270	235	199	307	275	249	198
	Czech Republic	327	322	318	275	345	327	331	288
	Hungary	305	280	273	236	313	297	273	231
	East (average)	303	287	272	231	317	289	283	236

Notes: for the comprehensive systems of Ireland, Sweden, Great Britain and Norway we report test scores in column *ISCED* 3 academic. Belgium includes only Flanders and Norway only the Bokmal region

2.4 Literature review

Two strands of literature are related to our study. The first examines the changes in the returns to schooling in CEE economies after the fall of the Berlin Wall. The second is broader, focuses on cohort effects in the returns to education and inves-

[9]Data do not allow the separation of academic and vocational education at this level.

tigates the hypothesis that, since education under communism is less appropriate for a market economy, it should receive a lower return than post-communist education. As summarized in Table B.2, there is a large body of empirical evidence documenting the increase in the return to schooling in CEE countries during the transition period. Selected examples of this literature are reviewed below. Fleisher, Sabirianova, and Xiaojun (2005), consider several contributions in the field and conclude that returns to education increased markedly during the transition, both in CEE economies and in Russia. Orazem and Vodopivec (1997), compare the wages of different skill groups in Slovenia before and after the transition, and find that returns to schooling increased during the early phases of the transition. Munich et al. (2005), obtain similar results studying the case of the Czech Republic. Andren, Earle, and Sapatoru (2005), estimate the impact of schooling on monthly earnings from 1950 to 2000 in Romania. Nearly constant at about 3 to 4 percent during the socialist period, returns to schooling increased steadily during the 1990s and reached 8.5 percent in the year 2000. Finally, Flabbi, Paternostro., and Tiongson (2008), use data from the International Social Survey Programme (ISSP) covering the period 1991-2002 and show that the estimated returns to schooling have increased mainly during the early transition stage, with limited changes during the period of late transition. Are these changes in the returns to education homogeneous or do they differ across different age cohorts? Card and Lemieux (2001), suggest that cohort effects in the college wage gap arise when workers belonging to different cohorts are imperfect substitutes in production. On the supply side, and conditional on demand, large cohorts of labour market entrants with a given education level can command lower earnings at entry because of stronger competition, which is more intense when the degree of substitutability across neighbour cohorts is higher. On the demand side, labour market conditions at the time of entry in the labour market matter, and tougher conditions prevailing at the beginning of a career may produce persistent negative consequences. For instance, Oreopoulos, Wachter, and Heisz (2008), estimate that young graduates entering the labour market during a recession suffer significant initial earnings losses which fade away only after 8 to 10 years. In CEE economies, cohort effects could have emerged both because of changes in labour supply by educational attainment and because of the differential exposure to the

transition, which radically modified the structure of labour demand by sector and by skills. Perhaps the most important determinant of cohort effects in the returns to education is skill-biased technical change. This type of progress increases the relative productivity of skilled labour and generates a continuous upward shift in its demand (see Acemoglu and Autor 2010, for a recent discussion and an extension of the original model).[10] The skill biased technical change hypothesis explains why the college wage premium has not declined over time in the US or the UK, in spite of the massive expansion of tertiary education observed earlier in the US and later in Western Europe (see Machin and McNally 2007 for a review of this large literature). Only a few studies attempt to estimate cohort effects in the returns to education in CEE countries. Jurajda (2005), uses Czech data and finds that the returns to one year of education in 2002 were close to 10 percent for the young generation aged 24 to 44 and equal to 8.7 percent for the older generation aged 45 to 61. Campos and Jolliffe (2007), study Hungary and argue that it is not formal education but experience acquired before the transition which is outdated. In the communist era, workers operated old technologies and followed procedures and regulations which disappeared in the subsequent market economy. They show that returns to general secondary, college and university education increased over time from 1986 to 2004, but were unchanged for vocational education. They also compare the returns to education earned by individuals aged less than or equal to 20 and more than 20 in 1986 - before the fall of the Berlin Wall -and in 2004, and find that returns are higher for the older age cohorts. Since the younger age group in 2004 was born in 1984 or later, it is quite likely that this group has been entirely educated after the fall of communism. The older age group includes instead both those who have studied entirely under communism and those who have experienced both systems. The higher returns earned by the older age group apparently suggest that pre-transition education is not obsolete in the modern market economy. A potential problem with this interpretation, however, is that the younger age group excludes most college

[10]The by now classical work by Katz and Murphy (1992), suggests that the long-term dynamics in the college wage gap between 1963 and 1987 in the US are consistent with a linearly increasing relative demand for college graduates, with fluctuations largely explained by changes in the relative supply. Card and Lemieux (2001), argue that the increase in college wage premium in the US, Canada and the UK has been largely due to the slowdown in the rate of growth of educational attainment that began with cohorts born in the early 1950s.

graduates, who typically graduate after age 20. Because of this exclusion, the lower returns found for the younger generation could be driven by self-selection out of education and into the labour market. Munich et al. (2005), use Czech data for the period 1991 to 1996 to estimate the wage effects of the number of years of communist and post-communist education and find that years of post-communist education have a lower return than years of education under communism. They argue that this evidence strongly contradicts the hypothesis that human capital acquired under communism is less appropriate for a market economy. Denny and Doyle (2010), estimate returns to formal education in the Czech Republic, Hungary and Slovenia in the mid Nineties on a sample of workers educated mainly under communism. They find that their estimates are not influenced by the inclusion of IALS test scores, that are measures of 'functional literacy', among the controls. Finally, Orlowski and Riphahn (2009), compare East and West Germany and find that, whilst returns to education are comparable, returns to experience are much lower in Eastern Germany than in Western Germany almost twenty years after re-unification. They argue that socialist labour market experience is of little value in the new market economy, but that schooling acquired in the East could still be a useful signal of innate individual productivity.[11] Similar results have been found by Chase (1998), for both the Czech Republic and Slovakia, and Flanagan (1998), for the Czech Republic. In summary, the available empirical evidence does not support the hypothesis that the education accumulated under communism is obsolete in a market economy. Yet this evidence is based on single CEE economies and often covers the period of early transition from communism, when intense economic restructuring could have altered the structure of returns. As argued by Machin and McNally (2007), there is a sharp contrast between the large body of empirical evidence documenting the rise in returns to education occurring during early pro-market reforms, and the paucity of studies that consider the late transition of the EU accession period. In the current paper, we fill this gap by estimating the returns to 'communist' education more than 15 years since the fall of the Berlin Wall.

[11] Education in the former GDR was very selective, and only about 10 percent of all students attaining grade 10 were admitted to high school and could attend the advanced school exam (Abitur) required to be admitted to a University (Riphahn and Trubswetter, 2010).

2.5 Methodological Issues and Proposed Solutions

It is well known that ordinary least squares estimates of the returns to education are distorted either because observed schooling is correlated with unobserved ability or because of measurement error in years of schooling. When we compare returns across countries, as we do in this paper, this distortion washes away only if the bias induced by measurement error and/or by the correlation between schooling and ability is constant across countries. One reason to believe that this is not the case in the current setup is that, while in Western economies the market influenced educational choice via relative returns to education, in Communist societies these returns had essentially no influence. In a sense, because market forces played no role in educational choice, there was a better match of talents to education in these societies, at least as far as higher education is concerned, and therefore a higher correlation between schooling and unobserved ability. In this section, we describe our proposed solution to this problem. In the first sub-section, we derive earnings functions from a theoretical model where workers of different age are imperfect substitutes. In the second sub-section, we introduce our estimation strategy.

2.5.1 The Model

Following Card and Lemieux (2001), we assume that firms in each economy produce output using the following production function

$$Y_{ct} = \bar{N}_{ct} \tag{2.1}$$

where Y is output, $\bar{N}$ is total employment in efficiency units, c the country and t the year. Total employment in efficiency units is the sum of male ($\bar{N}_m$) and female labour ($\bar{N}_f$)

$$\bar{N}_{ct} = e^{\theta_{fct}} \bar{N}_{fct} + e^{\theta_{mct}} \bar{N}_{mct} \tag{2.2}$$

where θ is an efficiency parameter, which varies by gender (m for males and f for females), country and time. Both female and male employment in efficiency units consists of k groups of imperfectly substitutable workers, who differ in their age

$$\bar{N}_{fct} = \left\{ \sum_{a=1}^{k} e^{\mu_{fcat}} \left[N_{fcat} \right]^{\rho_f} \right\}^{\frac{1}{\rho_f}} \tag{2.3}$$

$$\bar{N}_{mct} = \left\{ \sum_{a=1}^{k} e^{\mu_{mcat}} \left[N_{mcat} \right]^{\rho_m} \right\}^{\frac{1}{\rho_m}} \tag{2.4}$$

where μ is an efficiency parameter, a is age, $-\infty \leq \rho_f, \rho_m < 1$, and N is employment. Further assume that product prices are given in the international market, and normalized to 1. Define g as gender (f: females; m: males). Profit maximization with respect to employment yields the following first order condition

$$e^{(\mu_{gact} + \theta_{gct})} \left(\frac{N_{gact}}{\bar{N}_{gct}} \right)^{\rho_g - 1} = w_{gact} \tag{2.5}$$

where w is the real wage. By taking logs of (2.5) we obtain

$$lnw_{gact} = \mu_{gact} + \bar{\theta}_{gct} + (\rho_g - 1) ln \left(\frac{N_{gact}}{N_{gct}} \right) \tag{2.6}$$

where $\bar{\theta}_{gct} = \theta_{gct} + (\rho_g - 1) ln \left(\frac{N_{gact}}{\bar{N}_{gct}} \right)$.

Equation (6) is the demand for labour in the cell defined by gender, age, country and time. Labour market equilibrium requires that we characterize supply. Define relative supply as $\left(\frac{P_{gact}}{P_{gct}} \right)$, where P is the labour force. When the labour market is perfectly competitive, relative demand equals relative supply and we have that $\left(\frac{N_{gact}}{N_{gct}} \right) = \left(\frac{P_{gact}}{P_{gct}} \right)$ In the presence of frictions or wage bargaining, nonzero unemployment emerges and employment is equal to $N_{gact} = P_{gact} - U_{gact}$, where U is unemployment. Some rearrangement yields

$\frac{N_{gact}}{N_{gct}} = \frac{P_{gct}}{N_{gct}} \left[\frac{P_{gact}}{P_{gct}} (1 - u_{gact}) \right]$ where u is the unemployment rate. Taking logs and defining $\hat{\theta}_{gct} = \bar{\theta}_{gct} + (\rho_g - 1) ln \left(\frac{P_{gact}}{P_{gct}} \right)$, equation (6) becomes

$$lnw_{gact} = \mu_{gact} + \hat{\theta}_{gct} + (\rho_g - 1) ln \left(\frac{P_{gact}}{P_{gct}} \right) - (\rho_g - 1) u_{gact} \tag{2.7}$$

With wage bargaining or frictions, unemployment is positive and real wages are higher than in perfect competition. We assume that the efficiency parameter μ depends on education S - measured as the number of years of schooling and a resid-

ual component ω, which captures both unobserved talent and other un-observables such as school quality, which are likely to affect human capital and productivity (Hanushek and Woessmann, 2009)

$$\mu_{gact} = \psi_g S_{gact} + \omega_{gact} \tag{2.8}$$

Using (2.8) into (2.7) and letting the vector Y include the log of relative supply and the unemployment rate, we can write the wage equation more compactly as

$$lnw_{gact} = \beta_g Y_{gact} + \psi_g S_{gact} + \omega_{gact} + \hat{\theta}_{gct} \tag{2.9}$$

2.5.2 Empirical framework

The residual component in equation (2.9) includes age-invariant employment and labour force participation effects ($\hat{\theta_{gct}}$) and unobserved group characteristics such as average cognitive and non-cognitive abilities (ω_{gact}). To these we add selection effects, measurement errors and pure noise. Since some of these components are likely to be correlated with measured schooling, standard ordinary least squares estimates of eq. (2.9) are biased (see for instance Card, 1999). The patterns of self-selection into employment are likely to vary by gender and cohort. As discussed in Appendix B1, we control for these patterns by adding to (2.9) cell specific unemployment and activity rates. We also assume that unobserved abilities vary with observed parental background FB_{gac} and re-write (2.9) as

$$lnw_{gact} = \bar{\beta}_g Y_{gact} + \phi_g S_{gact} + \omega_{gact} + \hat{\theta}_{gct} + \nu_{gact} \tag{2.10}$$

where the vector Y now includes also parental background variables and activity rates. The error term ν_{gact} captures remaining cell-specific un-observables, which we further decompose as $\nu_{gact} = \chi_{act} + \zeta_{gact}$, a gender - invariant component, which includes for instance common environmental effects such as school quality - and a residual component ζ_{gact} . Following the approach proposed by Card and Rothstein, 2007, we "difference out" the gender-invariant component χ_{act} by taking gender differences between cells defined by the same birth cohort, country and time. Using Δ as the between genders difference operator (males minus females) and differencing

(2.10) by gender, we get

$$\Delta lnw_{act} = \bar{\beta}_f \Delta Y_{act} + (\bar{\beta}_m - \bar{\beta}_f) Y_{mact} + \phi_f \Delta S_{act} + (\omega_m - \omega_f) S_{mact} + \Delta\hat{\theta}_{ct} + \Delta\zeta_{act} \quad (2.11)$$

where the country by time effects $\Delta\hat{\theta}_{ct}$ can be captured in a flexible way by using country by time dummies. When differencing by gender, we take explicitly into account the possibility that the effect of schooling and other controls on log wages varies by gender. Equation (2.11) associates the gender difference in log average wages to the difference in average years of schooling and to the years of schooling attained by males in the same cell. If the effect of schooling on earnings varies by gender, this difference is picked up by the coefficient $(\omega_m - \omega_f)$. Although differencing removes common unobservables, gender specific unobservables still remain, and could be correlated with the change and level of average schooling. For this reason, we model the residual error component $\Delta\zeta_{act}$ as $\Delta\zeta_{act} = \kappa_{ct} + \kappa_{at} + \kappa_{ac} + \kappa_{act}$, where are country by time dummies, is defined as $\kappa_{at} = A * \kappa_t$, where A is age in the cell and κ_t are time dummies, and the term κ_{ac} is given by $\kappa_{ac} = \kappa_c + \kappa_a + \kappa_{ac} + \pi_1 GDP_{ac} + \pi_2 A * \kappa_c$, where κ_c and κ_a are country and birth cohort dummies, GDP is log real GDP per head at age 10, which is expected to capture the average economic environment faced by the age cohort at the time of education, and $A * \kappa_c$, are age by country effects. This set of dummies and additional controls affects the gender wage gap Δlnw_{act}. The use of country and country by year dummies allows us to control for country specific unobservables, both time variant and time invariant. By using birth cohort dummies, age by time dummies, GDP per capita at age 10 and age by country effects we also control in a flexible way the differential gender effects of potential experience and time of entry in the labour market, as well as the residual differential gender effects on selection into employment. Our maintained hypothesis is that, conditional on these controls, there is no residual correlation between the error term and the level and gender-difference in years of schooling. The identification assumptions required to obtain consistent estimates of the returns to schooling in this setup are

$E[\Delta S_{act}, \kappa_{act}] = 0$

$E[\Delta S_{act}, \kappa_{act}] = 0$

In the empirical implementation, we model relative supply as function of both relative population and the cell-specific activity rate. We also use the first lags of these variables and of the cell-specific unemployment rates to alleviate concerns that reverse causality runs from the dependent variable (the wage gap) to the explanatory variables. Due to data availability, labour market variables are defined by country, year of survey and 5-year age groups.

Table 2.5. Educational attainment in Germany 2006: by area of study and gender. Individuals born between 1951 and 1970.

		Males	Females
Studied in formerFRG			
	Living in former FRG	12.53	12.43
Studied in formerGDR			
	Living in former FRG	12.64	12.72
	Living in former GDR	12.86	12.69

Source: our computations based on SOEP data

2.6 Data

We use both German and European data. For Germany, we compare the returns to education earned in the period 2000-2009 by the cohorts of individuals born between 1945 and 1970 who completed their schooling in Eastern Germany before the fall of the Berlin Wall with the returns earned during the same period by the same birth cohorts who completed their education in Western Germany.[12] Our data are drawn from the German Socio Economic Panel (SOEP), a longitudinal survey of private households and persons in the former FRG. The survey was started in 1984 and was expanded in 1990 to cover the territory of the former GDR. An advantage of these data is that we can identify where education was attained, independently of where the individual is currently residing and working.[13] Therefore, we can compute returns to education for individuals educated in the former GDR who are currently working either in the regions of the former GDR or in the regions of the former FRG.[14] Following Haisken De New and Frick, 2005, we compute years of schooling as the difference between the age when the highest level of education was attained and country specific information on the age when school typically starts. Real hourly earnings in Euro (CPI deflated) are obtained as the ratio of monthly gross earnings to the number of hours worked.[15] Table 2.5 shows that the average years of schooling attained by East Germans living in the Western Germany are very similar to the years attained by East Germans living in Eastern Germany. Independently of their place of residence, East Germans are slightly more educated than West Germans, mainly because of the broader diffusion of upper secondary education. Table 2.6 shows instead that average real hourly gross wages in 2006 were generally higher for Germans educated in the former FRG.[16] We extend our analysis to other Eastern European countries by using EU-SILC, a survey of living conditions in European countries, which covers both Western and Eastern European countries. The

[12]For each year in the sample we include individuals aged at most 55.

[13]This information is contained in the variables 'psbilo' and 'pbbilo' in the SOEP dataset. See-Haisken De New and Frick 2005 for details.

[14]We exclude Berlin (ex-West and ex-East) from our data because of the lack of information on some of the controls used in the regressions.

[15]We only consider individuals with at least 15 and at most 80 hours worked per week. For most countries we use data on gross personal income.

[16]The relatively high average wage earned by West Germans employed in the former GDR is based on a small sample of individuals, who most likely are filling high paying managerial jobs.

EU-SILC is based on nationally representative samples, which collects comparable cross sectional and longitudinal micro data on income poverty and social exclusion and contains information on income, housing, material deprivation, labour, health, demography and education. We use data from three waves (2006-2007-2008) and 23 countries: Bulgaria, The Czech Republic, Estonia, Hungary, Poland, Romania, Slovenia and Slovakia in Eastern Europe and Austria, Belgium, Denmark, Cyprus, Ireland, Spain, France, Finland, Greece, Italy, Netherlands, Norway, Portugal, Sweden and the United Kingdom in Western Europe. For each country in the sample, we only consider the sub-sample of individuals born between 1951 and 1964 who reside in the country of birth (more than 94 percent of the total in 2008), because EU-SILC data do not report the country of origin. By doing so, we minimize the risk of assigning the wrong education system to individuals. This would happen if we were to assume that a person living in France but born and educated in Poland completed her education in the former country rather than in the latter. By selecting the birth cohorts from 1951 to 1964, we exclude individuals who have attained part of their education before and part after the fall of the Berlin Wall. Further details on the data used in this paper are in Appendix B2.

Table 2.6. Real hourly gross wage of Germans born between 1951 and 1970, by area of education and employment. Germany 2006.

	Educated in the GDR	Educated in the FRG
Employed in former GDR	12.34	22.24
Employed in former FRG	16.08	18.42

Notes: See Table 5.

2.7 Results

We organize the presentation of our results in three sub-sections. The first sub-section looks at the returns to education in Germany, the second sub-section considers these returns in Western and Eastern Europe (Germany excluded), and the final sub-section presents returns in Europe by level of schooling.

2.7.1 Returns to education in Germany

German re-unification offers a unique opportunity to test the appropriateness of the education acquired under communism in a capitalistic labour market. Two previously separated economic and educational systems, radically different in many respects, merged in 1990 and the institutions of the capitalistic former FRG were imposed to Eastern Germany. As remarked by Orlowski and Riphahn, 2009, East and West Germany have shared since unification similar labour market institutions in addition to history and language. Nonetheless, and especially in the early 1990s (Rainer and Siedler, 2009), the higher wages and the better economic conditions of the West on the one hand, and the minor cultural and geographical barriers on the other hand, induced a large number of Eastern Germans to seek a job in the West. According to our data, as many as 15 percent of those born between 1952 and 1970 who completed their education in the former GDR were residing and working in the West in 2007.[17] Due to this massive migration, an appraisal of the returns to education earned by individuals educated in the former GDR needs to consider the returns earned both by those who moved to the West and by those who remained in the Eastern Landers. Letting $1-q$ and q be the probability of being employed in the East and the West respectively, and and W_e the wages earned in the West and the East, the average wage earned by the individuals educated under communism is

$$W = qW_o + (1-q)W_e = W_e + q(W_o - W_e) \tag{2.12}$$

and the marginal return to one year of education under communism is

[17] During the same period, less than 1 percent of Western educated Germans moved to the previous East. These data do not include the city of Berlin.

$$\frac{\partial lnW}{\partial E} = \frac{\partial lnW_E}{\partial E}[1 - s(1-q)]\left(\frac{\partial lnW_o}{\partial E} - \frac{\partial lnW_E}{\partial E}\right) + \frac{\partial lnq}{\partial E}(1-s) \qquad (2.13)$$

where $s = \left(\frac{W_e}{W}\right)$

Using the identification strategy described in sub-section 2.4.2, we evaluate this return by estimating separately each component on the right hand side of (2.13). Table 2.7 shows the estimated returns to education for the Germans who work in the group of regions where they were educated[18], and Table 8 presents the estimated returns earned by the Germans educated in the former GDR, who migrated after the reunification and were employed in the regions of the former FRG during the period 2000-2009. We group data by year of birth, year of the survey, area of education (former GDR or former FRG) and gender, and run weighted regressions, using as weight $s = \left(\frac{1}{N_M} + \frac{1}{N_F}\right)^{-1}$, where N_M and N_F are the number of males and females in each cell, as suggested by Card and Rothstein, 2007. We also cluster standard errors by year of birth to take into account the possibility that the error term is correlated across the ten available waves 2000 to 2009 - because of common year of birth effects.

[18]The full estimates are reported in Table B.3 in Appendix B4.

Table 2.7. Estimated returns to education. Germans working in the same areas of education (previous East and previous West). Years 2000 to 2009. Weighted regressions, without and with between-gender differences (BGD). Dependent variable: log hourly wage.

Years of schooling	Females educated in GDR	Females educated in FRG	Males educated in GDR	Males educated in FRG
without BGD	0.060	0.038	0.052	0.092
	(0.040)	(0.022)	(0.029)*	(0.029)***
with BGD	0.029	0.047	0.083	0.107
	(0.047)	(0.021)**	(0.033)**	(0.042)**
Test of difference between East and West (p-value)	0.60		0.87	
Observations	215	215	215	215

Notes: each regression includes the lagged relative population, the activity rate, parental age, education and social status, year and birth cohort dummies and age by year effects. Standard errors clustered by year of birth within parentheses. One, two and three stars for statistical significance at the 10, 5 and 1 percent.

Table 2.8. Estimated returns to education. Germans educated in previous East Germany and working in the Landers of the previous Federal Republic. Years 2000 to 2009. Weighted regressions, without and with between-gender differences (BGD). Dependent variable: log hourly wage.

Years of schooling	Females	Males
without BGD	0.115	0.099
	(0.023)***	(0.019)***
with BGD	0.146	0.115
	(0.027)***	(0.023)***
Test of difference between East and West (p-value)	0.00	0.87
Observations	215	215

Notes: see Table 7.

In the first rows of Tables 2.7 and 2.8 we report - separately for males and females - the estimated marginal returns to an additional year of schooling before removing common unobserved effects by gender differencing. In the second and third rows of either table we report instead the estimated returns to schooling after removing common effects by gender differencing and the p-values of the tests on the equality of returns across areas. We find that differencing the data by gender increases the returns to schooling in all cases but one (females educated in the former GDR and working there). We believe that two effects are at work here. On the one hand, gender differencing[19] eliminates common un-observables that affect earnings, which include common ability traits and other common environmental effects. Since common unobserved ability is positively correlated with education, its removal should reduce estimated returns. On the other hand, differencing removes also common measurement errors affecting years of education, thereby reducing attenuation bias. Our interpretation of the results in Tables 2.7 and 2.8 is that the second effect prevails in most cases. In the rest of this section, we shall focus on the estimates obtained using between gender differences. When we consider individuals working in the same area where they were educated, we find that having studied in the former GDR yields positive but lower returns than having studied in the former FRG. The estimated difference, however, is not statistically significant. When we consider instead the individuals who have been educated in the former GDR but work and reside in the former FRG, we find that they earn higher returns (14.6 percent for females and 11.5 percent for males) than those who have studied in the previous FRG (4.7 percent for females and 10.7 percent for males). For females, this premium is statistically significant. We believe that the latter estimates are not affected in a significant way by self-selection. On the one hand, the use of between gender differences removes common un-observables which are likely to affect the decision to migrate, such as health conditions, the presence of relatives or friends already residing in West Germany[20] and the distance from the previous Western border. On the other hand, we control for residual gender specific heterogeneity

[19]Since we estimate separate regressions by country of education within the same country, we cannot use country dummies, country by time dummies and country by age effects. Moreover, since age dummies are included in the model, log GDP at age 10 is omitted

[20]Rainer and Siedler (2009) find that East German immigrants are more likely to be employed and to hold high wage jobs when they are socially connected to the West prior to emigrating.

in the cells defined by time and year of birth with parental background variables, time dummies, age dummies and age trends. To evaluate returns in (2.13) we need also to estimate the effect of education on the probability of residing and working in the regions of the former FRG. For this purpose, we pool the available individual observations for the period 2000 to 2009 and regress the probability of migrating from the East to the West for the relevant age cohorts on a battery of observable individual characteristics, including age, education and family background. We present our results in Table 2.10.[21] Consistently with previous literature in this area - see Rainer and Siedler, 2009, Hunt 2006 and Fuchs-Schundeln and Schundeln, 2009 -we find that migration declines with age, is lower for those married and with parents in the household, but is not affected by educational attainment. We can use the estimates in Tables 2.7, 2.8 and 2.9 to compute the marginal effect of education on the average wage, as shown in eq. (2.13). If we do so, we find that German females educated in the East earn a 5.3 percent marginal return [2.9%+(1-0.94*(1-0.155))*(14.6%-2.9%)-0.002/0.155*(1-0.94)], very similar to the return earned by females educated in the previous FRG (4.7 percent). On the other hand, German males educated in the East earn on average a 9.6 percent marginal return [8.3%+(1-0.71*(1-0.156))*(11.5%-8.3%)+0.005/0.156*(1-0.71)][22] , slightly less than the 10.7 percent return earned by Western educated Germans. This evidence suggests that including the returns to education earned by migrants who work in different labour markets does not alter the overall result: education attained either under communism or in the West before the fall of the Berlin Wall yields in the late 2000s similar payoffs to employed workers, in spite of the presumption that Western education was probably designed to better fit the needs of free market economies.[23] A comprehensive evaluation of returns to education in a comparative perspective requires, however, that we take explicitly into account the fact that German males educated under communism have a lower employment probability, in both Eastern and Western Landers, than German males educated in the West: in 2007 the percentage of males in the relevant age group who were educated in the East and the West but did not work was equal to 15.7

[21] Fuchs-Schundeln and Schundeln (2009) adopt a similar estimation technique. They include in their regressions the characteristics of the regions of origin but do not control for family background.

[22] For females, s=0.94, q=0.155; for males, s=0.71 and q=0.156.

[23] Our results confirm findings on an earlier sample by Smolny and Kirbach 2011 .

and 6.7 percent, respectively.[24] In the case of females, the percentage of those not working was higher (20 percent) but almost equal across the two groups. Because the replacement rate of pensions and unemployment benefits typically declines with earnings (see OECD 2010a,OECD 2011), the monetary returns to schooling for those not working are smaller than the returns for those working.[25] Therefore, when we consider also those out of work, the difference between average returns to education in the former FRG and in the former GDR is bigger than when we restrict our attention only to the employed.

[24]Since we explicitly control for selection into employment, our estimates show the expected return to schooling accruing to a randomly drawn individual from the population. However, the probability of being drawn into employment is significantly different across areas.

[25]To illustrate, assume that the marginal return to education for a wage earner is β and let unemployment and/or pension benefits be given by $B = \alpha + \gamma W$. In this case the replacement rate B/W declines with earnings and the marginal return to education for the unemployed and retired is $\beta \frac{\gamma W}{\alpha + \gamma W} < \beta$.

Table 2.9. Probability of migrating from East to West Germany. East Germans born between 1945 and 1970. Dependent variable: dummy equal to 1 in the case of migration.

	(1) Females who moved after 1989	(2) Males who moved after 1989
age	-0.006***	-0.007***
	(0.002)	(0.001)
years of education	-0.002	0.005
	(0.004)	(0.003)
children under 16 in household	0.011	0.026*
	(0.016)	(0.014)
married	-0.069***	-0.025
	(0.021)	(0.019)
father lives in household	-0.042	-0.069***
	(0.043)	(0.019)
mother lives in household	-0.095***	-0.061***
	(0.019)	(0.019)
father education (more than secondary)	-0.008	0.025
	(0.037)	(0.035)
mother education (more than secondary)	-0.019	-0.034
	(0.048)	(0.028)
Observations	8109	7375

Notes: Marginal effects of a pooled probit model. The sample consists of Eastern Germans born between 1945 and 1970 who were included in at least one wave of SOEP between 2000 and 2009. All specifications include year dummies. Standard errors – within parentheses – are clustered at the individual level.

2.7.2 Returns to Education in Western and Eastern Europe

Do the results for Germany hold also for other Eastern European countries which moved from communism to a free market economy after the fall of the Berlin Wall? We address this question by extending our analysis to the 23 European countries included in the EU-SILC dataset. However, since the country where education took place is identified in these data exclusively for the individuals who reside in their country of birth, we can only estimate the first element on the right hand side

of equation (2.13). As shown in Table 2.10, we find that the estimated returns to schooling earned by Eastern Europeans employed in Eastern Europe who have studied under communism are similar to the returns earned by equally aged Western Europeans who have studied and are employed in the West: one additional year of schooling is expected to raise wages by 6.6 percent in the East and by 5.8 percent in the West in the case of females, and by 5.8 percent and 5.4 percent in the case of males. These differences, however, are not statistically significant, as in the case of Germany (see Table 2.7). As in Germany, Eastern Europeans in the relevant age groups were more likely to be out of work in the late 2000s than Western Europeans: during this period, only 65.4 percent of Eastern females and 75.9 percent of Eastern males in the relevant age groups were earning a positive wage, compared to 69.1 percent and 87.1 percent of Western females and males. These gaps in employment rates have an impact on our results, because the replacement rate of pensions and unemployment benefits typically declines with earnings and informal earnings are lower than wages in the regular economy.[26] Had we considered the entire working age population rather than the formally employed, the comparison of returns to education would have been less favourable for the individuals educated under communism, especially for males, with non - negligible differences across countries.[27] An important question is whether the higher unemployment and inactivity rates in Eastern Europe are due to education under communism or are the result of the large frictions due to the transition or whether they depend on the interaction between these two causes. Table 2.1 shows that unemployment and activity rates between 2006 and 2008 for the cohorts born between 1951 and 1964 are higher than in Western Europe for any level of education, mainly reflecting the relatively low

[26] In both Eastern and Western Europe, a relevant proportion of those reporting to be unemployed or inactive was employed in the informal or shadow economy. The dimension of this economy varies across countries and depends on a number of factors, including the average tax burden, the quality of economic institutions and the generosity of the unemployment and early retirement schemes. During the transition, especially in countries such as Romania and Bulgaria, which lacked a proper system of social benefits, informal employment was a safety net against poverty (Parlevliet and Xenogiani, 2008). According to (Schneider, 2011), informal employment in 1998 was estimated to be as large as 63, 43 and 31 percent of the official labour force in Bulgaria, Romania and Slovenia, respectively. These figures compare with 16 percent in Austria, 19 to 23 percent in Germany and 30 to 48 percent in Italy.

[27] Note that unemployment and inactivity rates are lower in Table B.1 than in Table 2.1 because the former does not include inactivity due to disability and because the cohorts we are looking at in Table B.1 are about ten years younger than in Table 2.1.

job creation capacity of CEE economies more than fifteen years after the onset of the transition. It also shows that having acquired a college degree under communism exposes individuals born between 1951 and 1964 to a risk of unemployment or inactivity close to 20 percent of the risk incurred by those with less than high school education. This relative risk is about half as big as the risk incurred in Western Europe. Qualitatively similar results, especially for males, are reported in Table B.1, where the same cohorts are observed between 1994 and 1998, i.e. well within the transition period, for a narrower subsample of countries . We conclude from this that college education acquired during communism has been a relative insurance device for the lucky few who had access to higher education. Such device, however, did not operate for the vast majority with a high school degree. To see this, consider Table 2.11, which shows the 1998 to 2008 percent changes in the employment share of jobs with different education content in Eastern and Western European countries for the cohort of individuals born between 1954 and 1963. We define elementary occupations and production labour as 'low education', white collar, service jobs and skilled blue collar jobs as 'medium education' occupations and professionals, technicians and managers as 'high education' jobs.[28] In both areas, we notice a similar pattern of polarization of jobs at the bottom and top of the distribution, and a consistent decline of 'medium education' jobs, which require high school education. Needless to say, this evolution of labour demand is better matched to the structure of education in the West than to the one prevailing in the East during communism, with its strong emphasis on upper secondary (vocational) education. Notice also that the persistence of high unemployment and inactivity rates emerging from the comparison of Table 2.A.1 and Table 2.1 cannot be explained exclusively by hypothetical labour market frictions associated to the transition but unrelated to education. If so, these frictions should have been reabsorbed by the end of the Nineties. We therefore conclude that the structure of education and the interaction between this structure and the transition process are the likely causes of the high unemployment and inactivity rates that we still observe in recent years among Eastern Europeans educated under the communism. We have tested the reliability of our results with a set of robustness checks: first, we have restricted the sample of

[28]This classification is admittedly gross and is based on the average education content of occupations.

countries to those passing a pooling test; second, we have excluded from the sample the oldest cohorts, who could have benefitted from early retirement schemes; third, we have added controls for labour market experience and last, we have removed individuals employed in agriculture. Further details on these estimates are reported in Appendix B4. In all cases, our key results remain qualitatively the same.

Table 2.10. Estimated returns to education for individuals educated in the country of residence and employment. 23 European countries (Germany excluded). Years: 2006, 2007 and 2008. Weighted regressions, with between - gender differences (BGD). Dependent variable: log hourly wage

	Females Eastern Europe	Females Western Europe	Males Eastern Europe	Males Western Europe
Years of schooling (with BGD)	0.066 (0.021)***	0.058 (0.008)***	0.061 (0.018)***	0.054 (0.009)***
Test of difference between East and West (p-value)	0.739		0.719	
Observations	258	567	258	567

Notes: each regression includes the lagged relative population, the unemployment and the activity rate, parental education, country and birth cohort dummies, country by year dummies and country specific age effects. Clustered country by year of birth standard errors within parentheses. One, two and three stars for statistical significance at the 10, 5 and 1 percent. The countries in Eastern and Western Europe are: Bulgaria, The Czech Republic, Estonia, Hungary, Poland, Romania, Slovenia, Slovakia, Austria, Belgium, Denmark, Cyprus, Ireland, Spain, France, Finland, Greece, Italy, Netherlands, Norway, Portugal, Sweden and the United Kingdom.

2.7.3 Returns by level of schooling

The estimates in Tables 2.7, 2.8 and 2.10 rely on the assumption that the relationship between log hourly earnings and years of schooling is linear. This is equivalent to assuming that, at the individual level, the returns to an additional year of post-secondary education are equal to the returns to primary and secondary education.[29] We relax this assumption by replacing Eq. (2.11) with

$$\Delta lnw_{act} = \bar{\beta}_f \Delta Y_{act} + (\bar{\beta}_m - \bar{\beta}_f) Y_{mact} + \gamma_f^H \Delta S_{act}^H + (\gamma_m^H - \gamma_f^H) S_{mact}^H + \gamma_f^C \Delta S_{act}^C + (\gamma_m^C - \gamma_f^C) S_{mact}^C \Delta \hat{\theta}_{ct} + \Delta \zeta_{act} \tag{2.14}$$

[29] In this paper, primary and secondary education corresponds to ISCED levels 0 to 3 and post-secondary education to ISCED levels 4 to 6.

where S^H and S^C are average years of schooling in primary and secondary education and post-secondary education respectively, and estimate (15) on the same sample of 23 European countries used in Table 2.10. We compute individual values of S^H and S^C as follows: for the individuals who have completed at most upper secondary education (ISCED=3), S^H is equal to the number of years of schooling required to attain the highest degree and S^C is equal to zero. For the individuals who instead have completed post-secondary education, S^H is the country specific modal number of years of schooling required to attain upper secondary education (ISCED 3; 12 years for most countries in the sample), and post-secondary schooling is the difference between total years of schooling and this modal number. As already explained before, cell averages are obtained from individual values by averaging over gender, country, wave and year of birth. Table 2.12 reports our results. We find that the gap between the returns to post-secondary schooling and primary or secondary education is positive and highest among those educated under communism: in Eastern Europe, the returns earned by males (females) are equal to 11.4 percent (9.2 percent) per year of post-secondary education and to -2.2 percent (3.9 percent) per year of primary and secondary schooling.[30] This compares to 5.9 percent (7.1 percent) and 4.6 percent (4.3 percent) for Western educated individuals. While Eastern males with primary and secondary education earn significantly less than their Western counterparts, Eastern males with post-secondary education earn a substantially higher return than Western Europeans.[31] These differences are less precisely estimated for Eastern and Western females.

[30]The coefficients associated to primary and secondary education are not statistically different from zero.

[31]Estimates of (2.14) based on German data yield imprecise results, due to the smaller sample size and the lack of cross-country variability. Yet, results -available from the authors upon request - are qualitatively comparable to those reported here.

Table 2.11. Percent changes (1998 to 2008) in employment shares by type of occupation. Western and Eastern Europe. Cohort born between 1954 and 1963.

		Males	Females
Western Europe			
	Low education jobs	+1.4	+1.9
	Medium education jobs	-6.6	-3.9
	College jobs	+5.2	+2.1
Eastern Europe			
	Low education jobs	+2.6	+3.1
	Medium education jobs	-3.2	-7.9
	College jobs	+0.5	+4.9

Noters: Low education jobs: ISCO-88 codes 800 and 900 (production labour and elementary occupations). Medium education jobs: ISCO-88 codes 400, 500 and 700 (white collars, service workers and skilled blue collars). College jobs: ISCO-88 codes 100, 200 and 300 (managers, professionals and technicians). Source: our computations on European Labour Force Survey data. The countries in Eastern and Western Europe are: The Czech Republic, Estonia, Hungary, Poland, Romania, Slovenia, Slovakia, Austria, Belgium, Denmark, Ireland, Spain, France, Finland, Greece, Italy, Netherlands, Norway, Portugal, Sweden and the United Kingdom.

We conclude that the similarity of returns to education in Eastern and Western Europe found in sub-section 6.2 is partly driven by the assumption of linearity. Once this assumption is abandoned, the estimates reveal that the less educated - and especially males - are relatively penalized from having studied under communism. As remarked in the previous sections, one should add to these differences those associated with the probability of having a paid job. Since the probability of being out of work is particularly high among the low educated who have completed their education under communism, the average return to schooling earned by this group is likely to be even lower in a comparative perspective than shown in Table 2.12.

Table 2.12. Estimated returns to years of schooling S^H (until upper secondary) and S^C (post-secondary) for individuals educated in the country of residence and employment. 23 European countries (Germany excluded). Years 2006, 2007 and 2008. Weighted regressions, with between - gender differences. Dependent variable: log hourly wage

	Females Eastern Europe	Females Western Europe	Males Eastern Europe	Males Western Europe
Years of schooling S^H	0.039	0.043	-0.022	0.046
	(0.031)	(0.012)***	(0.029)	(0.013)***
Test of difference between East and West: S^H (p-value)	0.902		0.044	
Years of schooling S^C	0.092	0.071	0.114	0.059
	(0.023)***	(0.013)***	(0.020)***	(0.012)***
Test of difference between East and West: S^C (p-value)	0.439		0.020	
Test of difference between S^H and S^C (p-value)	0.104	0.191	0.000	0.459
Observations	258	567	258	567

Notes: see Table 10

2.8 Conclusions

Is education acquired under communism still valuable in the market economies of Eastern Europe more than 10 years after the fall of the Berlin Wall? To answer this question, we have estimated the returns to education earned by Eastern Europeans - including previous East Germans - who have completed their schooling before 1989 and compared these returns with those earned by the benchmark group of coetaneous Western Europeans. In contrast with most of the empirical literature in the area, we have proposed an estimation method that explicitly addresses the endogeneity of education in earnings regressions. We have found evidence that the returns to education earned by individuals who completed their education under

communism are similar to the returns obtained by the individuals belonging to the benchmark group. It is tempting to interpret these results as suggestive that education under communism is as valuable in the first decade of the new millennium as the education acquired in Western Europe by the same cohorts of individuals. We have argued that such conclusion is unwarranted for two reasons. First, fewer individuals educated under communism are still at work, compared to their Western European counterparts. Since the unemployed, disabled and retired typically earn lower returns, average returns to education which include these groups are likely to be lower, not higher, for those who have studied under communism than for the benchmark comparison group. Second, we have assumed that each year of schooling yields the same marginal return, independently of the level of education. Once this assumption is removed, we have found on the one hand that senior males who have attained only primary or secondary education under communism earn significantly lower returns in the post-transition Eastern European labour markets than equally educated Western Europeans employed in the West, and on the other hand that senior females with less than college education earn similar returns in the East and the West. We have also presented evidence that males and females who have completed post-secondary education under communism enjoy in these markets higher payoffs from this education than similarly educated Western Europeans who are employed in the West (albeit the difference is statistically significant only for males). One reason for the relatively poor performance of less educated Eastern males is that the radical transformation of the economy in CEE countries after the fall of the Berlin Wall affected mainly the industrial sector, where male employment with less than college education was heavily concentrated.[32] Senior male employees with primary and secondary education, who entered the labour market or were already in the market when the transition to market economies began in the East, took the brunt of the recession either in terms of unemployment or in terms of lower wages. On the one hand, this negative cohort effect has been quite persistent and lasted until the end of the transition. On the other hand, the industrial skills developed before the end

[32]If we consider employees in the East and the West born between 1951 and 1964, we find that in the East the percentage of male and female employees working in industry in 2008 was 55% and 29% respectively. In the West, the share of male and female employment in industry has been much lower, at 38% and 12% respectively.

of communism have become increasingly less suitable to the new market economies. Female labour has been spared because it was mainly employed in the expanding service economies of the East. The relatively low appropriateness of primary and secondary education attained under communism to modern market economies does not extend to college education. Quite the contrary, this type of education yields higher returns than the education obtained in the West. To interpret these findings, we look at demand and supply by education. Starting with the latter, Table 2.2 shows that the percentage of individuals who have obtained a college degree under communism is much lower than the percentage of equally aged individuals with a college degree in the West. In effective terms, this gap is even larger if we consider that the percentage of retired, unemployed and disabled college graduates belonging to the same cohorts is higher in Eastern Europe (Table 2.1). Turning to the demand side, we have mentioned in Section 2.3 that the service sector expanded significantly in CEE countries between 1989 and 1998 both in terms of employment and of its share of GDP. This process continued afterwards, although at a slower pace. According to the World Bank (WDI indicators), the value added generated by services in 2008 was between 58 and 66 percent of GDP in the major CEE economies, with an important expansion of the financial sector and other skill-intensive services, following progressive liberalization. Most likely, the increased demand for college graduates in the area - generated by the expansion of skill-intensive services - exceeded the growth in the stock of graduates, which could expand mainly via higher education of the younger cohorts, thereby contributing to the relatively high returns to post-secondary education for senior employees, independently of gender.

Chapter 3

Mental Health and Education Decisions.

3.1 Introduction

Poor mental health in childhood is strongly linked to poor mental health later in life and has been shown to have a serious impact on life chances (Richard and Abbott, 2009). Mental health problems may impact on human capital accumulation by reducing both the amount of schooling and the productivity level, which may in turn have lifelong consequences for employment, income and other outcomes (Eisenberg, Golberstein, and Hunt, 2009). Although the link between education and poor mental health has long been established, it has not often been examined in large-scale longitudinal studies. In this paper, we look at this issue in the context of a very recent and large scale study of adolescents in England. England is a particularly interesting country for analysing this issue because of a notably bad performance both on measures of child wellbeing and early drop-out from full-time education. For example, the UK made headlines in the last couple of years for ranking 24th out of 29 European countries on a league table of child wellbeing (Bradshaw and Richardson, 2009). The 'long tail" in the educational distribution has long been known to be a feature of the UK labour force and remains the case for younger cohorts. A relatively high proportion of young people end up classified as "not in education, employment or training" (NEET). The 2007 figures from the

OECD suggests that the UK ranks 21st out of 25 OECD countries in this respect (OECD, 2010b). Specifically, 11 per cent of 15-18 year olds are not in education, employment or training. This is similar to Italy and Spain but very different from countries such as Germany, France and the US where the relevant statistics are 4.2%, 5.8% and 6.3% respectively.

To what extent is poor mental health and low educational attainment/ drop-out linked? Clearly the association can operate in both directions. From a policy perspective, one would like to know the causal influence of poor mental health on these outcomes. This is notoriously difficult to establish and most research addresses the association rather than the causal impact. The latter can only be established by experiments (which can be difficult to generalise from) or from techniques that allow one to use "exogenous variation" in mental health to predict its causal impact on later outcomes. Recent work by Ding and Lehrer (2007) makes some progress in this direction by using genetic markers. However, such data are hard to come by and not uncontroversial since genes may impact on behaviour through more than one channel. In general, it is difficult to argue that indicators of mental health are exogenous because they are likely to be influenced by life events that are not fully measured in surveys. Nonetheless, it is still useful to know about the association between poor mental health and educational outcomes as this gives some information about the likely importance of mental health compared to other contributing factors (e.g. school or family characteristics). It is of interest to see whether such indicators continue to have an influence after controlling for many other factors that might explain educational outcomes. Moreover, it is interesting to see to what extent a simple screening device (like the 12 item General Health Questionnaire, used in this paper) is useful for predicting negative outcomes even after controlling for many observable characteristics. Such indicators might be useful for practitioners at school as well as for researchers, particularly since a large amount of mental health problems are thought to go unrecognised and untreated (Richard and Abbott, 2009). Also, early-onset mental disorders tend to co-occur in a complex and poorly understood patterns of comorbidity (Kandel et al. 1999).

The General Health Questionnaire (GHQ) is a screening instrument designed for use in general populations to detect the presence of symptoms of mental ill-

health and depression in particular (Goldberg, 1972). It has been extensively used in the psychological literature and is regarded as one of most reliable indicators of psychological distress or disutility (Argyle, 1989). The 12 item version of the GHQ (GHQ-12) is based on the questions that provided the best discrimination among the original criterion groups. Although most studies use the overall GHQ score as an indicator of mental health, it can be useful to separate the indicator into different factors as they may not all work in the same direction. For example, at lower levels anxiety can actually be productive (Sadock and Sadock., 2000). Graetz (1991) found years of education to be positively correlated with anxiety but negatively correlated with loss of confidence.

One of the contributions of this study is to look not only at the impact of an overall measure of mental health, but also to look at how the different components of the GHQ measure relate to educational attainment and the probability of moving into inactivity at an early age. We find strong patterns of association with respect to the overall measure, particularly for girls. However, we also find that different components are not equally important and that the effects of 'anxiety' and the other factors are indeed associated with outcomes in opposite directions. Secondly, we contribute by saying something about potential mechanisms through which poor mental health may impact on outcomes. For example, poor mental health may impact on later outcomes by intermediary choices such as insufficient investment in effort (e.g. playing truant) and self-medication (e.g. substance abuse). We attempt to say something about the likely importance of these factors. Finally, we perform our analysis using a very recent cohort of young people where there is longitudinal data - and in a country where both poor mental health and early drop-out are known to be very big problems by international standards. It is rare to have data for such a recent cohort (aged 14/15 in 2004) and this might be important because adolescent emotional problems and conduct disorder are known to have become more prevalent in recent decades (Collishaw et al. 2004).

3.2 Literature review

The relationship between mental health and education has been explored in both the psychological literature and the economic literature.

There are many small-scale studies in the psychological literature looking at the relationship between indicators of mental health and educational outcomes. The first study to examine the educational consequences of mental disorders in a national sample for the US was by Berslau et al. (2008). They find strong associations between child-adolescent mood, anxiety, substance use and conduct disorders with termination of schooling prior to each of three educational milestones (high school graduation, college entry among school graduates and completion of four years of college among college entrants). A more recent study also finding large effects (though among a broader set of disorders) is by Berslau et al. (2008). They find that the proportion of school terminations attributable to mental disorders was largest for high school graduation (10.2%) but also meaningful for primary school graduation, college entry, and college graduation. A disadvantage of these studies is that they are cross-sectional and rely on retrospective questions of 'early onset' mental health indicators.Within the psychological literature, longitudinal studies are rare. An example is the study by Fergusson and Woodward (2002). They find that the relationship between adolescent depression and subsequent educational underachievement could be fully explained by a range of social, familial and personal factors. Johnson, Cohen, and Dohrenwend. (1999) come to a similar conclusion with regard to the association between depression/anxiety disorders and subsequent staying on decisions.

The economic literature has only fairly recently begun to consider the relationship between mental health and educational outcomes. A strength of the contribution made by economists is that typically studies are longitudinal and have big sample sizes. Currie and Stabile (2006) and Fletcher and Wolfe (2008) both focus on the relationship between ADHD[1]and subsequent educational attainment and find evidence of a strong negative association. This is important because ADHD is one of the most common chronic mental health problems among young children together with conduct disorder and anxiety. However, there are other mental health problems that become more prevalent in early adolescence such as depression. An interesting

[1]Attention deficit hyperactivity disorder.

observation is that the sex difference in mental health problems is reversed in childhood and in early adolescence. For example, depression (and other types of mental health problems) are more prevalent in males in childhood whereas the opposite is true among adolescents and adults (Peterson et al. 1993)[2]and research on this issue suggests that this is not related to factors such as response bias on questionnaires or greater openness to acknowledging psychological difficulties. Furthermore depressive symptoms increase (for boys and girls) through early adolescence and the finding that girls suffer more than boys has been consistently documented in many countries (Seiffge-Krenke and Stemmler, 2003). This is true for both clinical levels of depression and subclinical levels such as depressive symptoms and depressive mood (Cicchetti and Toth., 1998). Theories about why this might be the case relate to the timing of puberty, different coping resources, and reaction to stressful life events
.

Using longitudinal data, Fletcher (2008) finds a robust negative relationship between depression in high school and subsequent educational attainment, even after controlling for a range of factors. In later work, he finds that the relationship is not very sensitive to the inclusion of sibling fixed effects (Fletcher, 2010). These studies pertain to a recent cohort (students in grades 7-12 in 1994-1995) and are for the US. The timeframe of the research could be important for what he finds because the prevalence of mental health problems has increased over time. In fact, there has been a rise internationally in the prevalence of depression (Cross-National Collaborative Group, 1992). Furthermore, work based on the British birth cohorts and the British Child and Adolescent Mental Health Survey suggests a rise in adolescent emotional problems and conduct disorder from the mid-1970s up to recent times (Callishaw et al. 2004). Fortunately, we are able to look at the relationship between poor mental health and educational outcomes for a very recent cohort of English students (aged 14/15 in 2004).

Other recent longitudinal studies that consider the relationship between adolescent mental health problems and educational attainment have much to say about depression in particular (Ding and Lehrer 2007; Eisenberg, Golberstein, and Hunt 2009; Fletcher 2008) and all suggest that this has a strong negative impact on ed-

[2]This finding has been commonly reported in the psychological literature for some time (e.g. Eme 1979; Gove and Herb 1974; Locksley and Douvan 1979).

ucational attainment. Ding and Lehrer (2007) and Fletcher (2008) look at this separately by gender and find that effects are only important for girls. The paper by Fletcher (2008) is closest to our paper in terms of the age group of students, outcomes and methodology (although he has a different measure of mental health, and the paper relates to a different time and country). He comments that it is not possible to provide evidence on the mechanism behind the association between depression and dropping out of high school because many of the choices that adolescents make before dropping out of school (e.g. skipping school) are not adequately captured in the data set. We are fortunate to be able to say something about these potential mechanisms because relevant questions are asked in the survey that we use.

3.3 Data

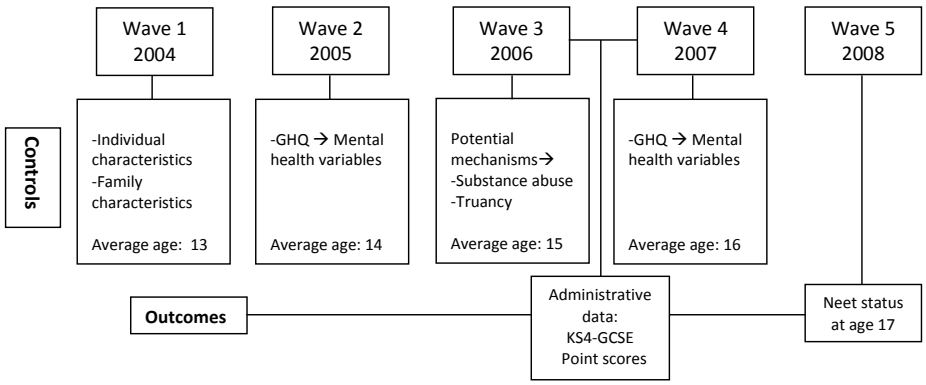
We use data from the Longitudinal Survey of Young People in England (LSYPE). This is a longitudinal data set which surveyed children aged between 13 and 14, beginning in 2004, for a total of around 14,000 young people. Parents are also surveyed and that data has been linked with administrative data on pupil test scores (including prior performance) and school-level information. Pupils (and parents) are surveyed each year up to age 18/19 (so far). The data set contains a very rich set of information about each young person. For example, it provides information on educational attainment, school information, family background as well as attitudes and behaviour. Young people respond to the 12-item General Health Questionnaire (GHQ) on two occasions - when the they are aged 14/15 (i.e. Wave 2) and again when they are aged 16/17 (i.e. Wave 4). We restrict our sample to people who answer all the GHQ questions in both waves. About 75% of young people answered all the GHQ questions in each Wave. 60% of young people answered all the questions in both waves and this reduces the sample to 8,122.[3]

The GHQ measure will be further described in the next section (a detailed description is provided in Appendix C.1 and C.2). We only retain observations for which we have valid test scores. The sample size is then 7,832. Descriptive statistics

[3]We have replicated our analysis when including people who answered 11 out of the 12 questions. Our results are not sensitive to this increase of our sample.

for the variables used in our analysis are shown in Appendix C.1 (Table C1 and C2). The sample used is similar to the full sample in many respects (such as the proportion 'not in education, employment or training' at age 17/18; parental qualifications and work status; family structure). For the most part, differences between the samples are quite small - although the sample used is a little better performing than the full sample in terms of exam results and in terms of socio-economic status (income and parental education). The samples are compared in Table C1.2. Our outcome variables are the (standardized) test score at age 16 and whether the person is classified as 'not in education, employment or training' (i.e. NEET) in Wave 5 (i.e. at age 17/18). The age 16 test score comes from the GCSE exam (General Certificate of Secondary Education) which all students in the UK undertake before leaving the compulsory phase of education at age 16. The National Curriculum is organized into different Key Stages. The GCSE exam marks the end of Key Stage 4. In many of our specifications, we control for test scores taken in national tests at the end of primary school (the end of Key Stage 2). The examination scores are all taken from administrative data that have been merged to survey data. Figure 3.1 summarizes the main variables used in the analysis.

Figure 3.1. LSYPE Dataset. Measures of Mental Health and Educational Attainment



3.3.1 The GHQ

The 12-item General Health Questionnaire (GHQ-12) is a self-reported measure of psychological morbidity intended to detect "psychiatric disorders among respondents in community settings and non-psychiatric clinical settings" (Goldberg and Williams., 1988). It is a measure of state which focuses mainly on the inability to carry out normal functions and the emergence of distressing symptoms. The GHQ-12 is a shorter version of a longer health questionnaire (originally 60-items) assessed by the World Health Organization and is used in studies about psychological disorders in primary health care. Due to its brevity and its capacity to retain many desirable psychometric properties, the GHQ-12 is widely used in clinical practice, epidemiological research and psychological research (Goldberg et al. 1997; Graetz 1991; Thomas, Benzeval, and Stansfeld 2005; Sweeting et al. 2009). It is also a very commonly used measure of individual well-being by economists in the UK literature (e.g. Clark and Oswald 1994; McCulloch 2001; Wigging et al. 2004; Gardner and Oswald 2007). [4] The questionnaire consists of 12 statements about aspects of wellbeing relating to worry, tension or sleeplessness. The respondent is asked to report his/her status over the past four weeks compared to what he/she considers "usual". There are six items that are positive descriptions of mood states (e.g "felt able to overcome difficulties"), and six that are negative descriptions of mood states (e.g. "felt like a worthless person"). The respondent states whether he/she is experiencing the symptom "much less than usual", "less than usual", "the same as usual" or "more than usual"(see Appendices C.2 and C.3). The most common scoring methods are as follows:

1) a Likert score, which assigns each response a value from zero to three, with zero indicating the highest level of well-being and three indicating the lowest. The answers are then summed to form the overall GHQ measure of psychiatric illness or

[4]Many of these studies use data from the British Household Panel Survey (BHPS) since it is one of the most detailed panel surveys which contains GHQ data. McCulloch (2001) uses the GHQ12 as an outcome of individual adversity associated with a census-based indicator of deprivation. Clark, Georgelli, and Sanfey (2001) use the GHQ to show that the unemployed have lower levels of mental well-being compared to working people. Similarly, Thomas, Benzeval, and Stansfeld (2005) use the GHQ as an outcome variable to measure the impact of different kinds of employment transitions (into various forms of non-employment) on psychological wellbeing. The GHQ has also been used widely in the literature on job satisfaction (Gardner and Oswald 2007, Callan et al. 2001).

mental well-being (total range 0-36).

2) a binary score system which assigns binary values to the responses from each question (where 1 indicates a low level of psychological well-being). The total score (over all items) varies between 0 and 12.

In both cases, the scoring is done such that high numbers indicate decreased levels of psychological well-being. Psychologists refer to being over a given threshold (beyond which the respondent is deemed to have mental health problems) as "caseness". When the binary score system is used, thresholds commonly applied in the literature are two, three and four positive items. We apply the most stringent threshold to indicate mental health problems or "risky cases" (i.e. 0-3: no ill-health; 4-12: high probability of common mental disorders). Many studies have analysed the dimensionality of the GHQ, assessing psychological morbidity in two or three dimensions rather than as a unidimensional index. The most common factorization is the one by Graetz (1991). He has proposed a three-dimensional model of the GHQ where questions can be used to create three distinct factors: Factor 1: "Anxiety and depression"- related to excessive worrying and difficulty controlling this worrying, Factor 2: "Anhedonia and social dysfunction"- related to reduced interest or pleasure in usual activities, and Factor 3: "Loss of confidence or self-esteem". This is a useful distinction since different aspects of GHQ-12 may be associated with behaviour in different ways (potentially in opposite directions). In our analysis we consider both the overall measure of mental health (both over a certain threshold and measured continuously), and these different components. In the survey, the GHQ questions are asked directly to the young person in Waves 2 and 4. In Table 3.1 we show summary statistics for key variables in our analysis.

Table 3.1. Main variables

Variable	Description	Boys	Girls	Boys at risk	Girls at risk
Panel A: Mental Health variables*					
GHQ at risk	GHQ 12 score >= 4. Risky threshold for GHQ12 "caseness".	0.110 (0.31)	0.248 (0.43)	1 (0)	1 (0)
GHQ Likert (0-1)	GHQ expressed in a continuous range [0-1]. The 12-GHQ questions are measured with the Likert scoring method (1-2-3-4) and then divided by 36.	0.236 (0.13)	0.312 (0.17)	0.500 (0.13)	0.543 (0.14)
Anxiety and Depression **	Continuous values ranging from 0-1. Includes four 'negative' items related to anxiety and depression.	0.235 (0.21)	0.338 (0.25)	0.615 (0.18)	0.670 (0.17)
Loss of confidence**	Continuous values ranging from 0-1 Includes two 'negative' items related to self confidence.	0.139 (0.21)	0.240 (0.28)	0.501 (0.29)	0.572 (0.29)
Anhedonia and Social dysfunction. **	Continuous values ranging from 0-1. Include six "positive" items testing the ability to perform daily activities and to cope with everyday problems.	0.269 (0.12)	0.318 (0.13)	0.423 (0.16)	0.449 (0.16)
Panel B: Output variables					
Standardized point scores	GCSE Standardized point scores Key Stage 4. (i.e. 16/17 years.)	-0.107 (1.02)	0.115 (0.96)	-0.097 (1.09)	0.060 (1.01)
Neet in W5	Not in education, employment or training at age 17/18.	0.106	0.076	0.124	0.098

*= Mental health variables collected in wave 2 (i.e. when young person is 14/15 years old). These variables are available also in W4, see appendix for detailed descriptive statistics.
**= See appendix for the construction of these indexes.

Panel A shows that a fairly high percentage of boys and girls are classified as "at risk" by the binary measure (at Wave 2) - 11 per cent of boys and almost 25 per cent of girls. Girls have a higher probability of mental health problems in each of the three dimensions of the GHQ (anxiety and depression, loss of confidence, anhedonia and social disfunction). If we look at the outcome variables (panel B), we see that both boys and girls "at risk" have lower outcomes with regard to the test score at age 16 and the probability of being "not in education, employment or training" than people not at risk. However, on average girls fare better than boys (even within the subpopulation of people "at risk"). In Table 3.2 we show the proportion of boys and girls who scored positively (i.e. indicating worry/stress) with respect to each component of the GHQ at age 14/15 (Wave 2; columns 1 and 4), and at age 16/17 (Wave 4; columns 2 and 5). Panel A reports the proportion of adolescents who could be defined as "at risk" according to the stringent threshold (i.e. where worry/anxiety is indicated in the response to at least 4 out of 12 questions). We also report results for a lower threshold - at least 2 out of 12 questions . For comparison, we show

the same data for 15 year olds from a recent survey of Scottish children (Sweeting, Young, and West, 2009). The comparable data are shown in columns 3 and 6. It is interesting to observe how similar the English and the Scottish studies are in terms of the overall incidence of poor mental health as well as for each separate indicator. [5] Other insights from this Table are that girls report a higher level of stress or worry than boys according to all indicators. Also, the incidence of poor mental health increases with age.

[5]Our GHQ scores are in line also with a study in the Netherlands about young people aged 18-24 (Hoeymans, Garssen, Westert, and Verhaak, 2004). They find a 'GHQ caseness' of 25% for young people aged 18-24, as well as higher rates for females.

Table 3.2. Comparison of GHQ scores in the LSYPE with Scottish data.

VARIABLES	(1) Boys Wave 2 Age 14/15	(2) Boys Wave 4 Age 16/17	(3) Boys in Scotland 2006 Age 15	(4) Girls Wave 2 Age 14/15	(5) Girls Wave 4 Age 16/17	(6) Girls in Scotland 2006 Age 15
Panel A: GHQ at risk						
GHQ at Risk: standard (>=2)	15.5	22.8	21.5	32.2	38.7	44.1
GHQ at Risk: stringent (>=4)	11	16.1	10.2	24.7	29.8	26.7
Panel B: Factor 1- Anxiety and Depression						
Felt constantly under strain	19.6	29.1	21.9	33.2	41.4	36.5
Being feeling unhappy or depressed	14.7	18.9	18.5	29.6	32.5	37.2
Lost much sleep over worry	11.9	17.7	16.4	24.2	31.8	29.5
Felt you could not overcome your difficulties	13.4	16.3	14.9	23.6	26.3	26.2
Panel C: Factor 2- Loss of confidence						
Been losing confidence in yourself	10.9	12.4	12.8	22.2	24.2	26.4
Been thinking of yourself as a worthless person	6.1	7.2	6.8	14.7	14.9	16.1
Panel D: Factor 3- Anhedonia and Social Dysfunction						
Been feeling reasonably happy, all things considered (disagree)	6.5	9.1	9.7	13.6	17.1	21.7
Felt you were playing a useful part in things (disagree)	5.8	8.9	8.1	9.6	11.8	15.1
Felt capable of making decisions about things (disagree)	3.2	4.1	5.2	6.3	9.6	13.3
Been able to face up to your problems (disagree)	4.7	5.9	6.9	10.8	13.3	16.6
Been able to enjoy your day-to-day activities (disagree)	7.5	12.8	11.1	12.6	18.6	15.2
Been able to concentrate on whatever you are doing (disagree)	9.7	11.7	15.4	17.6	19.4	32.0
Observations	4067	4067	1505	3765	3765	1539

3.3.2 Predicting poor mental health

Although not the main focus of our work, it is of interest to investigate how poor mental health, as measured by the GHQ, relates to pupil characteristics. A table showing summary statistics for variables used in our analysis for the whole

sample and according to whether young people are "at risk" is shown in Appendix C.1, Table C1. We estimate a Probit model where the dependent variable is the threshold beyond which someone might be thought of as "at risk". The results are reported in Appendix C.1, Table C3. We have run separate regressions for boys and girls. The first specification includes only basic controls (family income, ethnicity, parental education). In a second specification, we include a broad range of controls - many personal and family characteristics as well as school level characteristics. Results are qualitatively similar when using the continuous mental health measure. One of the most striking facts is how poorly the variables collectively explain poor mental health (no matter how we measure it). This suggests either that the GHQ does not have much informational content or that it simply does not correlate well with the usual indicators found in surveys, even though the information set is fairly rich. The main part of our analysis (and much of the literature) rejects the first explanation - it seems that the GHQ-12 does indeed have informational content. However, poor mental health is not well predicated by the usual indicators available to researchers and to schools (e.g. knowledge of test scores, family circumstances, socio-economic status and school characteristics). Relatively few variables are significantly different from zero, and this is more often the case for girls. For girls, among the variables that significantly effect the probability of having "poor mental health" (i.e. above the critical threshold) are family income (negative), whether the young person has a disability (positive), whether English is the main language of the household (positive), whether the parent is in good health (negative), the age 11 test score in English (positive), whether the young person goes to an independent school (negative). For boys, significant variables include whether the mother works full-time (negative), age 11 test score in Science (negative), age 11 test score in English (positive), and some school-level variables. Although the data set used here is very rich, it is nonetheless true that variables highlighted in the psychological literature are probably not well captured by the included variables. For instance, many psychological studies emphasize 'deficient active coping capacity' as a relevant variable (Andrews et al. 1978, Seiffge-Krenke 1995 and 2000). Various other studies point to a strong association between negative self-related cognitions and attribution styles including low self-esteem, low self-consciousness

and helplessness in depressive adolescents (Harter and Jackson, 1993). Moreover, parental rejection, lack of parental warmth and support, and disturbed parent-child relationships have also been frequently identified as strong correlates of adolescent depression.[6]Vernberg (1990) highlights the importance of low peer contact and peer rejection. Steinhausen and Metzke (2000) correlate depression with "a strongly controlling, highly competitive, less participation-oriented and low accepting school environment" (Zurich Adolescence Psychiatry and Psychopathology Study). Unfortunately these concepts are difficult to measure in survey data.

3.4 Conceptual framework

In order to investigate the relationship between mental health and educational attainment, we use a simple model of human capital accumulation. We follow the model proposed by Rosen (1977).[7] The relationship between earnings, y, and years of schooling, s, is assumed to be deterministic, and individuals, who differ in ability, A, maximize the present value of lifetime earnings and compare benefits with costs in deciding how much schooling to acquire.

$$y = f(s; A)$$

The discounted value of schooling net of foregone earnings, depends on the price of the skills acquired at school, the interest (discount) rate, and the ability of the individual. The benefit of schooling is increasing in both ability and price of skills acquired and decreasing in the interest rate. A worker characterised by a certain level of ability will decide to continue studying if the benefit exceeds

the cost. Fletcher (2008) was the first to include mental health in this framework.[8] He interprets ability as a function of mental health (d), and identifies two ways in which mental illness can influence education. First, assuming that mental illness decreases concentration during schooling (i.e. $A'(d) < 0$), mental illness lowers the returns to education because it affects the "individual's capacity or ability

[6]Barrera and Garrison-Jones (1992); Stark (1900); Steinhausen and Metzke (2000).
[7]This relies on Becker's fundamental contribution (Becker, 1962).
[8]Also Eisenberg, Golberstein, and Hunt (2009) use the same conceptual framework.

to learn".[9] Furthermore, Fletcher argues that mental illness can negatively affect the entire length of life or the duration of employment and therefore reduce the expected labour market benefits of education. This could lead individuals to invest less in schooling. Within this framework we investigate the relationship between mental health and education decisions using a reduced-form approach. We assume schooling to be a function of individual, family and school-level characteristics.

$$S^* = s(C, F, Sc) \tag{3.1}$$

where S^* is both the optimal schooling level and schooling performance; C represents individual characteristics (including mental health status and ability); F family characteristics; and Sc school level characteristics. We are mainly interested in C, particularly mental health status. Our analysis is conducted separately for boys and girls. Our main objective is to investigate the importance of mental health on schooling, where for schooling we mean both examination performance (test scores in the national exam before the end of compulsory education - GCSE) and schooling decisions (dropping out or NEET, "Not in Education, Employment or Training "). Thus, our outcome variables are the GCSE standardized test score and whether an individual is NEET at age 17/18. We separately consider three different measures of mental health: "GHQ caseness" (i.e. an indicator variable denoting whether the individual is "at risk" of poor mental health according to the highest threshold used by pscyhologists with regard to the GHQ); a continuous measure ranging from 0 to 1, GHQ Likert; and the three components of the continuous measure (i.e. the Graetz factors). Our basic OLS specification includes the mental health variable(s) and socio-economic and demographic controls (income, ethnicity and parental education). We later include a wider range of other potentially confounding variables (personal and family characteristics, and school level controls).

Let $MH_{i,t}$be the mental health status of an individual measured in wave 2; $edu_{i,s,t+n}$ represents our outcome variables: the GCSE point scores ($st.ptsc_{i,s,t+2}$) and NEET status ($neet_{i,s,t+3}$) of individual i in school s at time t.

We consider the following main (OLS) specification:

[9]For a summary of the empirical evidence on the link between schooling and mental health, seeRoeser, Eccles, and Strobel (1998).

$$edu_{i,s,t+n} = \alpha_1 + \alpha_2 MH_{it} + \alpha_3 X_i + \alpha_4 Z_{i,s} + \varepsilon_i \tag{3.2}$$

where X_i is a vector of personal and family characteristics, Z_{is} a vector of school characteristics, and ε_i the error component.

We then attempt to control for unobserved heterogeneity by including school fixed effects. Our preferred specification is:

$$edu_{i,s,t+n} = \alpha_1 + \alpha_2 MH_{i,t} + \alpha_3 X_i + u_s + \varepsilon_i \tag{3.3}$$

where u_s is the secondary school fixed effect. When we consider "NEET" as an outcome variable we include a measure of mental health in wave 4 in some specifications (i.e. GHQ) in addition to the measure taken at wave 2. In some specifications we include the examination score at age 16 as a control variable (i.e. the GCSE standardized point score).

The most detailed specification for "NEET" as an outcome variable is thus:

$$neet_{i,s,t+3} = \beta_1 + \beta_2 MH_{i,t} + \beta_3 MH_{i,t+2} + \beta_4 X_i + \beta_5 st.ptsc_{i,s,t+2} + u_s + \varepsilon_i \tag{3.4}$$

In the last part of the paper we consider schooling as a function of both mental health and risky behaviours. We hypothesize that the individual may respond to poor mental health by engaging in "risky behaviours". We are interested to investigate the extent to which the effect of mental health on outcomes might be "explained" through a behavioural response. We measure "risky behaviour" (RB) as consumption of cigarettes, alcohol and cannabis; and whether the individual says that he/she skips classes (i.e. truancy).

We estimate the following model:

$$st.ptscu_{i,s,t+2} = \gamma_1 + \gamma_2 MH_{i,t} + \gamma_3 X_i + \gamma_4 RB_{i,t+1} + u_s + \varepsilon_i \tag{3.5}$$

for standardized test scores as an outcome and

$$neet_{i,s,t+3} = \beta_1 + \beta_2 MH_{i,t} + \beta_3 MH_{i,t+2} + \beta_4 X_i + \beta_5 RB_{i,t+1} + \beta_6 st.ptsc_{i,s,t+2} + u_s + \varepsilon_i \tag{3.6}$$

for the "NEET" outcome.

In these models, mental health has a potential indirect effect on outcomes via risky behaviour (substance abuse and truancy). There might also be a direct effect of "risky behaviour" on outcomes. The timing is the following: mental health status, $MH_{i,t}$, is measured in wave 2, risky behaviors (RB) are collected in wave 3, and outcome variables (exam score and NEET status) are collected in waves 4 and 5 respectively. One potential problem is that the indicators of risky behaviour and mental health are likely to be serially correlated with (their own) past measures. Thus, past "risky behaviour" might potentially cause the onset of mental health problems (rather than the other way round). Although this generates an additional problem of interpretation with regard to equations (2.5) and (2.6), we still think this is an interesting exercise that will at least give some suggestive results on the interrelationship between mental health, "risky behaviour" and outcome variables. A more general problem is omitted variable bias. Mental health and outcome variables may both be influenced by a third unobserved variable. This problem is particularly intractable with regard to the issue at hand because it is difficult to think of variables that influence mental health while having no direct influence on educational outcomes. As referred to earlier, recent work on genetic markers (Ding and Lehrer, 2007) has made some progress in this direction. In our analysis, we have no such instrument. However, we have an extremely rich longitudinal data set which allows us to deal with this problem (at least partially) by controlling for a very large number of individual, family and school characteristics.

3.5 Results

The outcome variables considered are as follows: the "standardized points score" measured in a national examination at age 16 (GCSE) - from administrative data linked to the Wave 4 survey, and whether the individual is classified as "not in edu-

cation, employment or training" (NEET) measured in Wave 5 (age 17/18). Mental health variables are recorded in Waves 2 and 4. In all regressions we cluster standard errors at the school-level. In this section we present OLS results with a set of "basic controls" (ethnicity, parental income and education), with "additional controls" (very detailed controls for individuals, families and schools), and then we show the results including school fixed effects. Summary statistics for the full set of controls are reported in Appendix C.1, Table C1 .

3.5.1 Mental health and examination performance at age 16

Table 3.3 presents the results when we consider the standardized point score as outcome variable. The table is structured in two panels: the first refers to boys and the second to girls. Column 1 shows the results when only basic controls are included. Then we progressively introduce more controls in columns 2 and 3. In column 3 we also control for secondary school fixed effects. Coefficients are shown for the variable of interest - whether the individual is deemed to be at risk of mental illness because he/she scores positive on at least 4 of the 12 items of the General Health Questionnaire (GHQ) in Wave 2 (i.e when he/she was 14/15). In the simplest specification (with only basic controls), a negative relationship between the mental health indicator and exam performance at age 16 is shown only for girls. Poor mental health is associated with a reduction in exam scores of 0.086 standard deviations for girls. The inclusion of additional controls strengthens the relationship for both boys and girls (with the inclusion of school fixed effects being particularly important for boys). The most detailed specification (column 3) suggests that poor mental health is associated with lower exam performance of 0.083 and 0.158 standard deviations for boys and girls respectively. These are large coefficients and indicate that poor mental health may be a serious problem (for educational outcomes) if these associations reflect causality. Furthermore, these results suggest that the GHQ measure has strong predictive power even after controlling for a rich set of variables.[10]

In Table 3.4 we replicate the regressions presented in Table 3 using a continuous measure of mental health (i.e. the GHQ Likert). Results show a similar pattern as in Table 3.3, except that in the regression with only basic controls (column 1), the

[10]See Appendix Table C1.4 for the full set of controls.

association between the mental health indicator and exam performance is positive for boys. We explore this counter-intuitive result by breaking down the mental health indicator into its components (below). However with regard to the overall measure, the positive coefficient turns negative as soon as additional controls are included (column 2). In Table 3.5 we break down the continous measure of mental health to its constituent parts (as described in Section 3.3.1). Panels A and B show results for boys and for girls respectively. Column 1 shows that a positive association with the first factor ("anxiety and depression") is set against a negative association with the second factor ("loss of confidence"). This makes intuitive sense in the context of the literature as lower levels of anxiety may be productive Sadock and Sadock. (2000). This result is also consistent with results reported by Graetz (1991) who found that, for young people, anxiety is associated with more schooling. However, as we include more controls, the association between "anxiety and depression" and exam performance becomes smaller and statistically insignificant both for girls and boys. This suggests that any positive effect of anxiety on exam performance is captured by past educational attainment at age 7, family characteristics, and student sorting to secondary schools. When we focus on our preferred specification - the most detailed specification including secondary school fixed effects (column 3) - we see that "loss of confidence" remains important for boths boys and girls. "Anhedonia and social dysfunction" is also important (and the dominant factor) for girls.

Table 3.3. Standardized point score as outcome. GHQ at risk as mental health measure

Panel A: Boys

VARIABLES	(1) Basic controls	(2) Additional controls	(3) (2) + school Fixed effects
GHQ at risk	-0.001	-0.055	-0.083*
	[0.048]	[0.037]	[0.036]
Observations	3923	3923	3923
Adjusted R-squared	0.134	0.478	0.569

Panel B: Girls

GHQ at risk	-0.086**	-0.148**	-0.158**
	[0.034]	[0.027]	[0.026]
Observations	3644	3644	3644
Adjusted R-squared	0.137	0.479	0.574

Notes: Robust standard errors in brackets. ** p<0.05, * p<0.10. Each regression includes terciles of income dummies (Baseline is lowest tercile), ethnicity dummies (Mixed, Indian, Pakistani, Bangladeshi, Black Caribbean, Black African, other. Baseline is white ethnicity) and parental education dummies for both father and mother (University qualification, A-Level qualification, GCSE qualification, other qualification. Baseline is no qualification). The "Additional Controls" specification includes both personal and family characteristics and school level controls such as: whether young person has a disability, English as the main language of household, whether is a step family, dummies for family type (Married couple, lone father, lone mother, no parents in the household. Baseline is cohabiting couple), whether mother and father are working full time or part time, number of siblings, birth weight, whether born on time, if single parent family at birth, whether parents are in good health, total score in science, maths and english at KS2. School level controls are: average key stage 2 score of the primary school the pupil attended, school size, % of students with statements of special educational needs, % of students eligible to receive free school meals, % of students who do not speak English as a first language, School type dummies (Independent school; semi-autonomous school; special school. Baseline is other state school), whether grammar school, % achieving 5 or more grades at A-C in GCSE, 2004.

Table 3.4. Standardized point score as outcome. GHQ Likert (0-1) as mental health measure

Panel A: Boys

VARIABLES	(1) Basic controls	(2) Additional controls	(3) (2) + school Fixed effects
GHQ Likert (0-1)	0.278**	-0.143	-0.158*
	[0.116]	[0.093]	[0.085]
Observations	3923	3923	3923
Adjusted R-squared	0.133	0.478	0.569

Panel B: Girls

GHQ Likert (0-1)	-0.191**	-0.407**	-0.420**
	[0.093]	[0.075]	[0.066]
Observations	3644	3644	3644
Adjusted R-squared	0.137	0.479	0.574

Notes: Robust standard errors in brackets. ** p<0.05, * p<0.10. Each regression includes terciles of income dummies, ethnicity dummies and parental education dummies for both father and mother. The "Additional Controls" specification includes both personal and family characteristics and school level controls.

Table 3.5. Standardized point score as outcome. Graetz factors as mental health measure

Panel A: Boys

VARIABLES	(1) Basic controls	(2) Additional controls	(3) (2) + school Fixed effects
Anxiety and Depression	0.488**	0.142**	0.075
	[0.087]	[0.069]	[0.072]
Loss of confidence	-0.479**	-0.252**	-0.218**
	[0.087]	[0.065]	[0.069]
Anhedonia and Social Dysfunction	0.301**	-0.030	0.017
	[0.134]	[0.106]	[0.103]
Observations	3923	3923	3923
Adjusted R-squared	0.143	0.480	0.569

Panel B: Girls

Anxiety and Depression	0.432**	0.021	-0.051
	[0.085]	[0.068]	[0.065]
Loss of confidence	-0.435**	-0.161**	-0.143*
	[0.075]	[0.059]	[0.057]
Anhedonia and Social Dysfunction	-0.280**	-0.332**	-0.238**
	[0.142]	[0.109]	[0.102]
Observations	3644	3644	3644
Adjusted R-squared	0.147	0.480	0.574

Notes: Robust standard errors in brackets. ** p<0.05, * p<0.10. Each regression includes terciles of income dummies, ethnicity dummies and parental education dummies for both father and mother. The "Additional Controls" specification includes both personal and family characteristics and school level controls.

3.5.2 Mental health and the probability of being "Not in Education, Employment or Training" (NEET)

Our dataset allows us to measure whether the teenager effectively drops out of education and employment (known as "Not in Education, Training or Employment" or NEET) at age 17/18 (Wave 5). As shown in Table 3.1, the percentage of boys

and girls that are NEET at this age is 10.6% and 7.6% respectively . In Tables 3.6 and 3.7 we show estimates using the same specifications as described above for exam peformance, progressively introducing controls (columns 1-3). Then we include exam performance to discern whether there is an association between poor mental health and NEET over and above any association that works through exam performance (column 4). In column 5 we also include a measure of mental health at age 15/16 (as well the measure at age 13/14). This shows whether the association of early mental health problems with NEET remains after we control for later mental health and exam performance. Table 3.6 shows the results where mental health is measured by whether the individual is over the relevant threshold (4 out of 12 positive answers) and deemed to be "at risk". Table 3.7 presents the same specifications where we use the continuous measure of mental health (the Likert measure) instead of the threshold. In both tables, panels A and B show results for boys and girls respectively. In Table3. 6 the coefficients are similar for boys and girls - but more precisely estimated for girls. The results in columns 1-3 are not very sensitive to the inclusion of controls (although coefficients increase a little when more controls are included). Column 3 shows that poor mental health (measured at age 13/14) is positively associated with the probability of drop-out (or NEET) by 2.7 and 3.3 percentage points for boys and girls respectively - although this is only statistically significant for girls. This coefficient is very sizeable given the baseline figures for NEET. Surpisingly, the coefficient is only moderately reduced by including a control for exam performance (column 4) - suggesting that the association between poor mental health and NEET does not operate primarily through how the student does at school. Including a later measure of mental health - GHQ in wave 4 (column 5) - also moderately reduces the coefficient on the earlier measure. In the case of boys, the later measure of mental health shows a stronger association with NEET (and is statistically significant). When we replicate these regressions using the continuous measure of mental health (Table 7), the difference between boys and girls is larger. The early mental health measure only has explanatory power for girls. As before, the coefficient is little affected by including detailed controls, exam performance at age 16, and subsequent measures of mental health. However, when we use the later measure of mental health (measured at age 15/16), a positive coefficient is shown for

boys which is similar in magnitude to that shown for girls with regard to the earlier measure of mental health. In Table 3.8 we break down the continuous measure of mental health into its components. The structure of the table is the same as for table 3.7. We observe that "anhedonia and social dysfunction" seems to be driving the entire association for girls. For boys, this breakdown is largely uninformative - reflecting the overall lower association between early measures of mental health and NEET. However, when the analysis includes the components of the later mental health measure (column 5), "anhedonia and social dysfunction" shows up as an important factor for boys too. The magnitude of the association between this component of mental health and NEET is very similar to that found for girls when using the earlier measure of mental health. This analysis shows that there is a relationship between early indicators of mental health and the probability of drop-out. Although this is stronger for girls, we find similar patterns for boys using a later measure of mental health. The component "anhedonia and social dysfunction" seems to drive effects in both cases. These measures have a strong association with NEET over and above any association that might be influenced by examination performance.

Table 3.6. NEET in W5 as an outcome. GHQ at risk as mental health measure

Panel A: Boys

VARIABLES	(1) Basic controls	(2) Additional controls	(3) (2) + school Fixed effects	(4) (3)+GCSE pt.sc.	(5) (4) +MH at W4
GHQ at risk, W2	0.021	0.023	0.027	0.029	0.022
	[0.017]	[0.017]	[0.017]	[0.018]	[0.018]
GHQ at risk, W4					0.028*
					[0.016]
GCSE Standardized point score				-0.033*	-0.033*
				[0.009]	[0.009]
Observations	3655	3655	3655	3525	3525
Adjusted R-squared	0.010	0.019	0.045	0.054	0.055
Panel B: Girls					
GHQ at risk, W2	0.030**	0.031**	0.033**	0.027*	0.021*
	[0.011]	[0.011]	[0.011]	[0.011]	[0.012]
GHQ at risk, W4					0.017
					[0.011]
GCSE Standardized point score				-0.042**	-0.042**
				[0.008]	[0.008]
Observations	3468	3468	3468	3357	3357
Adjusted R-squared	0.014	0.038	0.053	0.061	0.062

Notes: Robust standard errors in brackets. ** p<0.05, * p<0.10. Each regression includes terciles of income dummies, ethnicity dummies and parental education dummies for both father and mother. The "Additional Controls" specification includes both personal and family characteristics and school level controls.

Table 3.7. NEET in W5 as an outcome. GHQ Likert as mental health measure

Panel A: Boys

VARIABLES	(1) Basic controls	(2) Additional controls	(3) (2) + school Fixed effects	(4) (3)+GCSE pt.sc.	(5) (4) +MH at W4
GHQ Likert in W2	0.012	0.032	0.026	0.026	-0.020
	[0.039]	[0.040]	[0.041]	[0.042]	[0.045]
GHQ Likert in W4					0.123**
					[0.043]
GCSE Standardized point score				-0.033**	-0.035**
				[0.009]	[0.009]
Observations	3655	3655	3655	3525	3525
Adjusted R-squared	0.010	0.020	0.046	0.055	0.057
Panel B: Girls					
GHQ Likert in W2	0.102**	0.108**	0.114**	0.087**	0.081**
	[0.031]	[0.031]	[0.029]	[0.029]	[0.032]
GHQ Likert in W4					0.015
					[0.031]
GCSE Standardized point score				-0.041**	-0.042**
				[0.008]	[0.008]
Observations	3468	3468	3468	3357	3357
Adjusted R-squared	0.016	0.040	0.055	0.062	0.062

Notes: Robust standard errors in brackets.** p<0.05, * p<0.10. Each regression includes terciles of income dummies, ethnicity dummies and parental education dummies for both father and mother. The "Additional Controls" specification includes both personal and family characteristics and school level controls.

Table 3.8. NEET in W5 as outcome. Graetz factors as mental health measure

Panel A: Boys

VARIABLES	(1) Basic controls	(2) Additional controls	(3) (2) + school Fixed effects	(4) (3)+GCSE pt.sc.	(5) (4) +MH at W4
Anxiety and Depression W2	-0.016	-0.001	0.005	-0.000	-0.015
	[0.031]	[0.031]	[0.034]	[0.034]	[0.036]
Loss of confidence W2	0.031	0.018	0.012	0.017	0.010
	[0.032]	[0.032]	[0.033]	[0.034]	[0.034]
Anhedonia and Social Dysfunction W2	-0.006	0.015	0.008	0.009	-0.026
	[0.050]	[0.051]	[0.050]	[0.051]	[0.054]
Anxiety and Depression W4					0.026
					[0.034]
Loss of confidence W4					0.018
					[0.033]
Anhedonia and Social Dysfunction W4					0.099*
					[0.058]
GCSE Standardized point score				-0.033**	-0.034**
				[0.009]	[0.009]
Observations	3655	3655	3655	3525	3525
Adjusted R-squared	0.009	0.019	0.045	0.054	0.056
Panel B: Girls					
Anxiety and Depression W2	0.027	0.041*	0.045	0.043	0.039
	[0.024]	[0.024]	[0.028]	[0.028]	[0.030]
Loss of confidence W2	-0.003	-0.015	-0.019	-0.017	-0.016
	[0.023]	[0.023]	[0.025]	[0.025]	[0.026]
Anhedonia and Social Dysfunction W2	0.109*	0.115*	0.126**	0.084*	0.080*
	[0.046]	[0.045]	[0.045]	[0.045]	[0.048]
Anxiety and Depression W4					0.013
					[0.029]
Loss of confidence W4					-0.010
					[0.025]
Anhedonia and Social Dysfunction W4					0.015
					[0.050]
GCSE Standardized point score				-0.042**	-0.042**
				[0.008]	[0.008]
Observations	3468	3468	3468	3357	3357
Adjusted R-squared	0.016	0.041	0.056	0.062	0.061

Notes: Robust standard errors in brackets. ** p<0.05, * p<0.10. Each regression includes terciles of income dummies, ethnicity dummies and parental education dummies for both father and mother. The "Additional Controls" specification includes both personal and family characteristics and school level controls.

3.5.3 Potential mechanisms

Finally, we investigate to what extent the observed associations can be "explained" by including possible behavioural mechanisms. For example, individuals may respond to poor mental health by substance abuse and by skipping classes at school. In addition to any direct effect these behaviours might have on exam performance and NEET, they might also wipe out some of the observed effect of poor mental health on these outcomes. "Substance abuse" is proxied by whether the individual (at age 14/15) has ever consumed cigarettes, alcohol or cannabis. Whether they have missed school is measured by whether they self-report as ever having played truant at age 14/15. The indicator of mental health is measured in the previous wave and the outcome indicators are all measured in subsequent waves. However, as discussed earlier, endogeneity problems can affect this analysis because of serial correlation between current and past behaviour. While substance abuse may lead to mental health problems, it might also be a reaction to mental health problems. Nonetheless, we believe that this is still an interesting exercise. In Table 3.9 we show results where the dependent variable is exam performance at age 16 and mental health is measured on the continuous Likert scale. We start by showing the association between poor mental health and exam performance in the most detailed specification (i.e. replicating Table 3.4, column 3). We then show how the association changes when controlling for truancy (column 2), indicators of substance abuse (column 3), and risky behaviours altogether (column 4). Results for boys (Panel A) show that the association between poor mental health and exam performance is wiped away by either or both of these controls. For girls (Panel B), controlling for these potential mechanisms reduces the association by about half (i.e. the coefficient goes from -0.42 in column 1 to -0.20 in column 5). This is "explained" equally by truancy and by consumption of both cigarettes and cannabis.

In Table 3.10 we estimate similar regressions when the dependent variable is NEET at Wave 5. Our starting point is the most detailed specification when controlling for measures of mental health at age 14/15 and at age 16/17, as well as exam performance at age 16 (table 7, column 5). As discussed previously, the earlier measure of mental health is important for girls whereas the later measure is important for boys. The regressions show that truancy has no association with the outcome

variable for boys (any effect is entirely absorbed through the association between truancy and exam performance). Although there is a small association between consumption of cigarettes and the probability of being NEET, this does nothing to the association between mental health and NEET for boys. Analogous regressions for girls (Panel B) show that only truancy has an association with NEET over and above any association that operates via exam performance. The association between mental health and NEET is reduced, but only moderately. Thus there seems to be evidence that these mechanisms play an indirect role in explaining NEET throught their impact on the exam scores.[11]

[11] We use a more stringent definition of truancy. In the questionnaire people are asked "Have you ever played truancy in the past 12 months?". We have excluded people who answered positively to the truancy question but, when asked about the frequency, they answered to have played truancy "only the odd day/that class". We have replicated our estimations using the original variable, results are slightly different: for neet in W5 as an outcome truancy does not have an effect for girls.

Table 3.9. Potential mechanisms with the GCSE standardized point scores as an outcome.

Panel A: BOYS

VARIABLES	(1) secondary school FE	(2) + Truancy	(3) +Consumption of substances	(4) +Risky behaviors
GHQ Likert (0-1) W2	-0.158*	-0.081	-0.024	0.010
	[0.085]	[0.084]	[0.083]	[0.082]
Cigarette ever			-0.312**	-0.273**
			[0.032]	[0.033]
Alcohol ever			-0.024	-0.016
			[0.031]	[0.031]
Cannabis ever			-0.211**	-0.188**
			[0.029]	[0.029]
Any truant		-0.518**		-0.359**
		[0.046]		[0.046]
Observations	3923	3923	3923	3923
Adjusted R-squared	0.569	0.586	0.601	0.609

Panel B: GIRLS

GHQ Likert (0-1) W2	-0.420**	-0.319**	-0.244**	-0.198**
	[0.066]	[0.066]	[0.065]	[0.065]
Cigarette ever			-0.247**	-0.219**
			[0.030]	[0.030]
Alcohol ever			-0.024	-0.017
			[0.032]	[0.032]
Cannabis ever			-0.174**	-0.143**
			[0.032]	[0.031]
Any truant		-0.510**		-0.374**
		[0.048]		[0.048]
Observations	3644	3644	3644	3644
Adjusted R-squared	0.574	0.589	0.600	0.608

Notes: Robust standard errors in brackets. ** p<0.05, * p<0.10. Each regression includes terciles of income dummies, ethnicity dummies and parental education dummies for both father and mother, controls for personal and family characteristics and school fixed effects.

Table 3.10. Potential mechanisms with NEET as an outcome.

Panel A: BOYS

VARIABLES	(1) secondary school FE	(2) (1)+ Truancy	(3) (1) +consumption of substances	(4) (1)+Risky behaviors
GHQ Likert (0-1) W2	-0.020	-0.021	-0.029	-0.028
	[0.045]	[0.045]	[0.045]	[0.045]
GHQ Likert (0-1) W4	0.123**	0.122**	0.120**	0.121**
	[0.043]	[0.043]	[0.043]	[0.043]
Cigarette ever			0.050**	0.051**
			[0.017]	[0.017]
Alcohol ever			-0.005	-0.005
			[0.016]	[0.016]
Cannabis ever			-0.010	-0.009
			[0.015]	[0.015]
Any truant		-0.001		-0.013
		[0.024]		[0.025]
GCSE Standardized point score	-0.035**	-0.034**	-0.029**	-0.030**
	[0.009]	[0.009]	[0.009]	[0.009]
Observations	3525	3525	3525	3525
Adjusted R-squared	0.057	0.057	0.059	0.059

Panel B: GIRLS

GHQ Likert (0-1) W2	0.081**	0.070**	0.082**	0.073**
	[0.032]	[0.032]	[0.033]	[0.033]
GHQ Likert (0-1) W4	0.015	0.009	0.017	0.011
	[0.031]	[0.031]	[0.031]	[0.031]
Cigarette ever			0.008	0.003
			[0.014]	[0.014]
Alcohol ever			-0.005	-0.006
			[0.014]	[0.014]
Cannabis ever			-0.007	-0.014
			[0.014]	[0.015]
Any truant		0.091**		0.094**
		[0.022]		[0.022]
GCSE Standardized point score	-0.042**	-0.035**	-0.042**	-0.037**
	[0.008]	[0.008]	[0.008]	[0.008]
Observations	3357	3357	3357	3357
Adjusted R-squared	0.062	0.067	0.061	0.067

Notes: Robust standard errors in brackets. ** p<0.05, * p<0.10. Each regression includes terciles of income dummies, ethnicity dummies and parental education dummies for both father and mother, controls for personal and family characteristics and school fixed effects

3.6 Conclusions

A growing literature shows that mental health disorders in adolescence have potentially long term effects on adult mental health. In this paper we show how poor mental health in early adolescence has a strong negative association with subsequent examination performance and drop-out from the labour market and education. England is a particularly interesting country for considering these issues because of a low international ranking both on measures of child wellbeing and on early drop-out. Although there are many studies that look at the relationship between mental health and education, there are not many large-scale longitudinal studies where it has been possible to look at the relationship for a recent cohort while also controlling for a wide range of personal, family and school characteristics. In our study we show that whereas it is difficult to "explain" poor mental health by a large range of characteristics, measures of poor mental health have a strong association with subsequent educational outcomes and the probability of being "not in education, employment or training" (or NEET). The measure of poor mental health comes from the General Health Questionnaire (GHQ-12), which is a screening instrument designed for use in general populations to detect the presence of symptoms of mental health (depression in particular). An insight from this study is how useful such a simple indicator might be for predicting who is "at risk" from negative educational outcomes - even after allowing for all the usual observable characteristics available to researchers and school practitioners. We conduct our analysis separately for boys and girls. Our findings show stronger patterns of association for girls than for boys (which is consistent with previous literature). We also find that different components of mental health are not equally important with respect to the outcomes considered here. "Anhedonia and social dysfunction" seems to be most important after including detailed controls. "Loss of confidence" also shows a strong association with examination performance. Finally, we consider potential mechanisms through which poor mental health may impact on outcomes. Individuals may respond to poor mental health by engaging in risky behaviours (consumption of substances -cigarettes, alcohol and cannabis) and by missing school. In addition to any direct effect these behaviours might have on exam performance and NEET, they may reduce some of the observed effect of poor mental heath on these outcomes. Our results suggest that this is a

reasonable hypothesis with regard to examination results but not with regard to NEET where they do not make any difference. The overall picture presented in this paper suggests that mental distress is strongly associated with poor educational outcomes and early drop-out. This research helps illustrate the potential importance of programmes aimed at improving the mental health of adolescents.

Bibliography

Acemoglu, D. (2003): "Patterns of Skill Premia," *Review of Economic Studies*, 70(2), 199–230.

Acemoglu, D., and J. Angrist (2001): "How Large are Human-Capital Externalities? Evidence from Compulsory-Schooling Laws," in *NBER Macroeconomics Annual 2000, Volume 15*, NBER Chapters, pp. 9–74. National Bureau of Economic Research, Inc.

Acemoglu, D., and D. Autor (2010): "Skills, tasks and technologies: Implications for employment and earnings," Discussion Paper 16082, NBER Working Paper, Cambridge, MA.

Andren, D., J. Earle, and D. Sapatoru (2005): "The wage effects of schooling under socialism and in transition: Evidence from Romania, 1950-2000," *Journal of Comparative Economics*, 33(2), 300–323.

Andrews, G., C. Tennant, D. Hewson, and G. Vaillant (1978): "Life Event Stress, Social Support, Coping Style, and Risk of Psychological Impairment.," *The Journal of Nervous and Mental Disease*, 166.

Argyle, M. (1989): *The Psychology of Happiness*. Routledge.

Ashenfelter, O., and C. Rouse (2000): "Schooling, intelligence and income in America.," in *Meritocracy and Economic Inequality*. Princeton Univ. Press, Princeton.

Autor, D. H., L. F. Katz, and M. S. Kearney (2008): "Trends in U.S. Wage Inequality: Revising the Revisionists," *The Review of Economics and Statistics*, 90(2), 300–323.

Bal, L. (2001): "Education to facilitate the way out of poverty," *Beyond Transition: The Newsletter about Reforming Economies*, 13(3), 31–32.

Bank, T. W. (2002): "Transition. The first ten years. Analysis and lessons for Eastern Europe and the former Soviet Union," Discussion paper, The World Bank, Washington DC.

Barrera, M. J., and C. Garrison-Jones (1992): "Family and peer social support as specific correlates of adolescent depressive symptoms.," *Journal of Abnormal Child Psychology*, 20, 1–16.

Barth, E., and C. Lucifora (2006): "Wage Dispersion, Markets and Institutions: The Effects of the Boom in Education on the Wage Structure," IZA Discussion Papers 2181, Institute for the Study of Labor (IZA).

Becker, G. S. (1962): "Investment in Human Capital: A Theoretical Analysis," *Journal of Political Economy*, 70.

Berend, I., and T. Berend (2001): *Decades of crisis: Central and Eastern Europe before World War II*. California University Press.

Berryman, S. (2000): *Hidden Challenges to the Education Systems in Transition Economies*. The World Bank.

Boeri, T., and J. F. Jimeno (2005): "The effects of employment protection: Learning from variable enforcement," *European Economic Review*, 49(8), 2057–2077.

Boeri, T., and K. Terrell (2002): "Institutional determinants of labor reallocation in transition," *Journal of Economic Perspectives*, 16(1), 51–76.

Bradshaw, J., and D. Richardson (2009): "An Index of Child Well-Being in Europe," *Child Indicators Research*, 2, 319–351.

Braga, M., D. Checchi, and E. Meschi (2011): "Institutional Reforms and Educational Attainment in Europe: A Long Run Perspective," IZA Discussion Papers 6190, Institute for the Study of Labor (IZA).

Breslau, J., M. Lane, N. Sampson, and R. C. Kessler (2008): "Mental disorders and subsequent educational attainment in a US national sample," *Journal of Psychiatric Research*, 42(9), 708 – 716.

Brunello, G., S. Comi, and C. Lucifora (2000): "The College Wage Gap in 10 European Countries: Evidence from Two Cohorts," IZA Discussion Papers 228, Institute for the Study of Labor (IZA).

Callan, M., D. Wainwright, M. Forsythe, B. Wall, and S. Almond (2001): "Mental health and stress in the workplace: the case of general practice in the UK," *Social Science and Medicine*, 52(4), 499–507.

Campos, N., and A. Jolliffe (2007): "Earnings, Schooling and Economic Reform: Econometric Evidence from Hungary (1986-2004)," Discussion Paper 2678, IZA Discussion Papers.

Card, D. (1999): *The Causal Effect of Education on Earnings*vol. Handbook of Labor Economics, chap. 3A, pp. 1802–1859. Amsterdam, North Holland.

Card, D., and J. E. DiNardo (2002): "Skill-Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles," *Journal of Labor Economics*, 20(4), 733–783.

Card, D., and T. Lemieux (2001): "Can Falling Supply Explain The Rising Return To College For Younger Men? A Cohort-Based Analysis," *The Quarterly Journal of Economics*, 116(2), 705–746.

Card, D., and J. Rothstein (2007): "Racial segregation and the black-white test score gap," *Journal of Public Economics*, 91(11-12), 2158–2184.

Charnoz, P., E. Coudin, and M. Gaini (2011): "Decreasing Wage Inequality in France 1976-2004: Another French Exception?," Discussion Paper 6, INSEE working paper.

Chase, R. S. (1998): "Markets for communist human capital: returns to education and experience in the Czech Republic and Slovakia," *Industrial and Labor Relations Review*, 51(3), 401–423.

Chay, K. Y., and D. S. Lee (2000): "Changes in relative wages in the 1980s Returns to observed and unobserved skills and black-white wage differentials," *Journal of Econometrics*, 99(1), 1–38.

Chevalier, A., and J. Lindley (2009): "Overeducation and the skills of UK graduates," *Journal Of The Royal Statistical Society Series A*, 172(2), 307–337.

Cicchetti, D., and S. L. Toth. (1998): "The development of depression in children and adolescents.," *American Psychologist*, 53, 221–241.

Clark, A., Y. Georgelli, and P. Sanfey (2001): "Scarring: The Psychological Impact of Past Unemployment," *Economica*, 68(270), 221–241.

Clark, A., and A. J. Oswald (1994): "Unhappiness and Unemployment," *The Economic Journal*, 104(424), 648–659.

Collishaw, S., B. Maughan, R. Goodman, and A. Pickles (2004): "Time trends in adolescent mental health.," *Journal of Child Psychology and Psychiatry*, 45, 1350–1362.

Currie, J., and M. Stabile (2006): "Child mental health and human capital accumulation: The case of ADHD," *Journal of Health Economics*, 25(6), 1094–1118.

Denny, K., and O. Doyle (2010): "Returns to basic skills in central and eastern Europe," *Economics of Transition*, 18(1), 183–208.

Deschenes, O. (2006): "Unobserved ability, comparative advantage, and the rising return to college in the United States, 1979-2002," Discussion paper, Department of Economics, University of California, Santa Barbara.

Dickens, R., S. Machin, and A. Manning (1999): "The Effects of Minimum Wages on Employment: Theory and Evidence from Britain," *Journal of Labor Economics*, 17(1), 1–22.

DiNardo, J., N. M. Fortin, and T. Lemieux (1996): "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach," *Econometrica*, 64(5), 1001–44.

Ding, W., and S. F. Lehrer (2007): "Do Peers Affect Student Achievement in China's Secondary Schools?," *The Review of Economics and Statistics*, 89(2), 300–312.

Edin, P., and B. Holmlund (1995): "The Swedish Wage Structure: The Rise and Fall of Solidarity Wage Policy?," NBER Chapters, pp. 307–344. National Bureau of Economic Research, Inc.

Eisenberg, D., E. Golberstein, and J. B. Hunt (2009): "Mental Health and Academic Success in College," *The B.E. Journal of Economic Analysis and Policy*, 9.

Eme, R. (1979): "Sex differences in childhood psychopathology: a review," *Psychological Bulletin*, 86, 574–95.

Fergusson, D. M., and L. J. Woodward (2002): "Mental Health, Educational, and Social Role Outcomes of Adolescents With Depression," *Arch Gen Psychiatry*, 59(3), 225–231.

Filer, R., and D. Munich (2003): *Public support for private schools in post-communist Central Europe: Czech and Hungarian experiences'*vol. Choosing Choice: School Choice in International Perspective, pp. 122–196. New York: Teachers College Press.

Flabbi, L., S. Paternostro., and E. Tiongson (2008): "Returns to education in the economic transition: A systematic assessment using comparable data," *Economics of Education Review*, 27(6), 724–740.

Flanagan, R. (1998): "Were communists good human capitalists? The case of the Czech Republic," *Labour Economics*, 5(3), 295 –312.

Fleisher, B., K. Sabirianova, and W. Xiaojun (2005): "Returns to skills and the speed of reforms: Evidence from central and Eastern Europe, China, and Russia," *Journal of Comparative Economics*, 33(2), 351–370.

Fletcher, J., and B. Wolfe (2008): "Child mental health and human capital accumulation: The case of ADHD revisited," *Journal of Health Economics*, 27(3), 794–800.

Fletcher, J. M. (2008): "Adolescent depression: diagnosis, treatment, and educational attainment," *Health Economics*, 17(11), 1215–1235.

——— (2010): "Adolescent depression and educational attainment: Results Using Sibling Fixed Effects.," *Health Economics*, 19, 855–871.

Fortin, N. M., and T. Lemieux (1997): "Institutional Changes and Rising Wage Inequality: Is There a Linkage?," *Journal of Economic Perspectives*, 11(2), 75–96.

Fuchs-Schundeln, N., and M. Schundeln (2009): "Who stays, who goes, who returns? East-West migration within Germany since reunification," *Economics of Transition*, 17(4), 703–738.

Gardner, J., and A. J. Oswald (2007): "Money and mental wellbeing: A longitudinal study of medium-sized lottery wins," *Journal of Health Economics*, 26(1), 49–60.

Goldberg, D., R. Gater, N. Sartorius, T. B. Ustun, M. Piccinelli, O. Gureje, and C. L. H. Rutter (1997): "The validity of two versions of the GHQ in the WHO study of mental illness in general health care.," *Psycholgical Medicine*, 27, 191–197.

Goldberg, D. P. (1972): *The detection of psychiatric illness by questionnaire.* London: Oxford University Press.

Goldberg, D. P., and P. Williams. (1988): *A user's guide to the General Health Questionnaire*. NFER-Nelson, Windsor.

Goldin, C., and L. F. Katz (2007): "The Race between Education and Technology: The Evolution of U.S. Educational Wage Differentials, 1890 to 2005," NBER Working Papers 12984, National Bureau of Economic Research, Inc.

Gove, W., and T. R. Herb (1974): "Stress and mental illness among the young: a comparison of the sexes," *Social Forces*, 53, 256–265.

Graetz, B. (1991): "Multidimensional properties of the General Health Questionnaire," *Social Psychiatry and Psychiatric Epidemiology*, 26, 132–138.

Haisken De New, J., and J. Frick (2005): "The Desktop Companion to the German Socio-Economic Panel Study," Discussion paper, DIW Berlin, Germany.

Hanushek, E., and L. Woessmann (2009): "Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation," Discussion Paper 14633, NBER Working Papers, Cambridge, MA.

Harmon, C., H. Oosterbeek, and I. Walker (2003): "The Returns to Education: Microeconomics," *Journal of Economic Surveys*, 17(2), 115–156.

Hoeymans, N., A. Garssen, G. Westert, and P. Verhaak (2004): "Measuring mental health of the Dutch population: a comparison of the GHQ-12 and the MHI-5," *Health and Quality of Life Outcomes*, 2(1), 23.

Hunt, J. (2006): "Staunching Emigration from East Germany: Age and the Determinants of Migration," *Journal of the European Economic Association*, 4(5), 1014–1037.

Johnson, J., P. Cohen, and B. Dohrenwend. (1999): "A longitudinal investigation of social causation and social selection processes involved in the association between socioeconomic status and psychiatric disorder.," *Abnormal Psycholgy*, 108, 490–499.

Jurajda, S. (2005): "Czech Relative Wages and Returns to Schooling: Does the Short Supply of College Education Bite?," *Czech Journal of Economics and Finance*, 55(1-2), 83–95.

Kandel, D., J. Johnson, H. Bird, M. Weissman, S. Goodman, B. Lahey, D. Regier, and M. Schwab-stone (1999): "Psychiatric Comorbidity Among Adolescents With Substance Use Disorders: Findings From the MECA Study," *Journal of the American Academy of Child and Adolescent Psychiatry,*, 38, 693–699.

Katz, L. F., and D. H. Autor (1999): "Changes in the wage structure and earnings inequality," in *Handbook of Labor Economics*, vol. 3 of *Handbook of Labor Economics*, chap. 26, pp. 1463–1555. Elsevier.

Katz, L. F., and K. M. Murphy (1992): "Changes in Relative Wages, 1963-1987: Supply and Demand Factors," *The Quarterly Journal of Economics*, 107(1), 35–78.

Kessler, R. C., C. L. Foster, W. B. Saunders, and P. Stang (1995): "Social consequences of psychiatric disorders, I: Educational attainment," *American Journal of Psychiatry*, 157(7), 1026–1032.

Koeniger, W., M. Leonardi, and L. Nunziata (2007): "Labor Market Institutions and Wage Inequality," *Industrial and Labor Relations Review*, 60(3), 340–356.

Krueger, A. B., and M. Lindahl (2001): "Education for Growth: Why and for Whom?," *Journal of Economic Literature*, 39(4), 1101–1136.

Lamo, A., J. Messina, and E. Wasmer (2010): "Are specific skills an obstacle to labor market adjustment?," Discussion Paper 5250, IZA DP.

Lemieux, T. (2006): "Postsecondary Education and Increasing Wage Inequality," *American Economic Review*, 96(2), 195–199.

Leuven, E., H. Oosterbeek, and H. van Ophern (2004): "Explaining International Differences in Male Wage Inequality by differences in Demand and Supply of Skill," *Economic Journal*, (144), 478–498.

Levy, F., and R. J. Murnane (1992): "U.S. Earnings Levels and Earnings Inequality: A Review of Recent Trends and Proposed Explanations," *Journal of Economic Literature*, 30(3), 1333–1381.

Lochner, L., and E. Moretti (2004): "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports," *American Economic Review*, 94(1), 155–189.

Locksley, A., and F. Douvan (1979): *Problem behaviour in adolescence.* New York: Brunner-Mazel.

Machin, S. (1997): "The decline of labour market institutions and the rise in wage inequality in Britain," *European Economic Review*, 41(3-5), 647–657.

Machin, S., and S. McNally (2007): "Tertiary education systems and labour markets," Discussion paper, OECD, Paris.

Manacorda, M. (2004): "Can the Scala Mobile explain the fall and rise of earnings inequality in Italy? A semiparametric analysis, 1977-1993.," *Journal of Labor Economics*, 22(3), 585–613.

Martins, P. S., and P. T. Pereira (2004): "Does education reduce wage inequality? Quantile regression evidence from 16 countries," *Labour Economics*, 11(3), 355–371.

McCulloch, A. (2001): "Ward-level deprivation and individual social and economic outcomes in the British Household Panel Study.," *Environment and Planning*, 33, 667–684.

Melly, B. (2005): "Public-private sector wage differentials in Germany: Evidence from quantile regression," *Empirical Economics*, 30, 505–520.

Micklewright, J. (1999): "Education, inequality and transition," *Economics of Transition*, 7(2), 343–376.

Milligan, K., E. Moretti, and P. Oreopoulos (2004): "Does education improve citizenship? Evidence from the United States and the United Kingdom," *Journal of Public Economics*, 88(9), 1667–1695.

Mincer, J. (1974): *Schooling, Experience and Earnings,*. Columbia University Press, New York.

Munich, D., J. Svejnar, and K. Terrell (2005): "Returns to human capital under the communist wage grid and during the transition to a market economy," *The Review of Economics and Statistics*, 87(1), 100–123.

OECD (2000): "Literacy in the Information Age. Final report of the International Adult Literacy Survey," Discussion paper, OECD, Paris.

——— (2010a): "Benefits and Wages: OECD Indicators," Discussion paper, OECD, Paris.

——— (2010b): "Education at a Glance.," Discussion paper, OECD.

——— (2011): "Pensions at a Glance.," Discussion paper, OECD, Paris.

Orazem, P., and M.Vodopivec (1997): "Value of human capital in transition market: Evidence from Slovenia," *European Economic Review*, 41, 893–903.

Oreopoulos, P., T. Wachter, and A. Heisz (2008): "The short- and long-term career effects of graduating in a recession: Hysteresis and heterogeneity in the market for college graduates," Discussion Paper 3578, IZA Discussion Papers.

Orlowski, R., and R. Riphahn (2009): "The East German wage structure after transition," *Economics of Transition*, 17(4), 629–659.

Parlevliet, J., and T. Xenogiani (2008): "Report on informal employment in Romania," Discussion Paper 271, OECD Development Centre Working Paper, Paris.

Petersen, A. C., B. Compas, J. Brooks-Gunn, M. Stemmler, S. Ey, and K. E. Grant (1993): "Depression in adolescence.," *Amercan Psychology*, 48, 155–168.

Pfeifer, C. (2011): "Risk Aversion and Sorting into Public Sector Employment," *German Economic Review*, 12(1), 85–99.

Rainer, H., and T. Siedler (2009): "The role of social networks in determining migration and labour market outcomes'," *Economics of Transition*, 17(4), 739–767.

Richard, M., and R. Abbott (2009): "Childhood mental health and life chances in post-war Britain. Insights from three national birth cohort studies," Discussion paper, Center for Medical Health, London.

Riphahn, R., and P. Trubswetter (2010): "The intergenerational transmission of educational attainment in East and West Germany," Discussion paper, mimeo.

Roeser, R., J. Eccles, and K. Strobel (1998): "Linking the study of schooling and mental health: selected issues and empirical illustrations at the level of individual.," *Educational Psychologist*, 33(4), 58–66.

Rosen, S. (1977): "Human capital: a survey of empirical research.," *Research in Labor Economics*, 1, 3–40.

Sadock, B. J., and V. A. Sadock. (2000): *Kaplan and Sadock's comprehensive textbook of psychiatry.* Philadelphia, PA, US: Lippincott Williams and Wilkins Publishers.

Schneider, F. (2011): "The shadow economy and shadow economy labor force: what do we (not) know?," Discussion Paper 5769, IZA Discussion Papers.

Seiffge-Krenke, I., and M. Stemmler (2003): "Coping with everyday stress and links to medical and psychosocial adaptation in diabetic adolescents," *Journal of Adolescent Health*, 33(3), 180 – 188.

Smolny, W., and M. Kirbach (2011): "Wage Differentials between East and West Germany. Is it Related to the Location or to the People?," *Applied Economics Letters*, 18(9), 873–879.

Stark, K. D. (1900): *The treatment of depression during childhood: A school-based program.* New York: Guilford.

Steinhausen, H.-C., and C. W. Metzke (2000): "Adolescent Self-Rated Depressive Symptoms in a Swiss Epidemiological Study," *Journal of Youth and Adolescence*, 29, 427–440.

Sullivan, D. H., and T. M. Smeeding (1997): "Educational Attainment and Earnings Inequality in Eight Nations," *International Journal of Educational Research*, 27(6), 513–25.

Sweeting, H., R. Young, and P. West (2009): "GHQ increases among Scottish 15 year olds 1987-2006," *Social Psychiatry and Psychiatric Epidemiology*, 44(7), 579–586.

Taber, C. R. (2001): "The Rising College Premium in the Eighties: Return to College or Return to Unobserved Ability?," *Review of Economic Studies*, 68(3), 665–91.

Thomas, C., M. Benzeval, and S. A. Stansfeld (2005): "Employment transitions and mental health: an analysis from the British household panel survey," *Journal of Epidemiol Community Health*, 59, 243–249.

Vernberg, E. M. (1990): "Psychological adjustment and experiences with peers during early adolescence: Reciprocal, incidental, or unidirectional relationships?," *Journal of Abnormal Child Psychology*, 18, 187–198.

Walker, I., and Y. Zhu (2008): "The College Wage Premium and the Expansion of Higher Education in the UK," *Scandinavian Journal of Economics*, 110(4), 695–709.

Wiggins, R., P. Schofield, A. Sacker, J. Head, and M. Bartley (2004): "Social position and minor psychiatric morbidity over time in the British Household Panel Survey 1991-1998," *Journal of Epidemiology and Community Health*, 58(9), 779–787.

Appendices

A Appendix to Chapter 1

A.1 Data Appendix

Supply Index: This index is created from OECD data. It is a measure of relative supply and it is calculated for each gender in country, yearly, as the ratio of college graduates to non-college graduates (ISCED 5/ISCED 3). skilled workers.

Demand Index: This index is created from EU-KLEMS data. It is a measure of relative demand and it is calculated for each country, yearly, considering hours worked by high-skilled persons engaged (share in total hours) by industries relative to hours worked by middle skilled workers.

R&D intensity: Data ar drawn from the OECD-STAN database which provides information on imports, R&D and value added in the manufacturing sector from 1973-2009. Using these data I manage to build a proxy for technology using data on total manufacturing for R&D and value added for all countries.

Minimum Wage: This is the ratio of the statutory minimum wage to the median wage in each country. It is provided by the OECD. Germany, Denmark, Finlad and Italy have no statutory minimum wage.

Employment Protection Legislation (EPS): The employment protection legislation consist on a set of norms and procedures followed in case of dismissal of redundant workers. Act as deterrent: it protects workers with permanent contracts from the risk of early termination of their employment contract Decisions involve also third parties, the legitimacy of a layoff ultimately depends on court ruling. EPS is a strongly redistributive institution. It protects those who already have a job, notably a permanent contract in the formal sector. Unemployed individuals and workers with temporary contracts suffer in the presence of strict EPS for permanent contracts. The former experience longer unemployment spells, while the latter are caught in a secondary labor market of temporary contracts. The OECD indicators of employment protection are synthetic indicators of the strictness of regulation on dismissals and the use

of temporary contracts. These indicators are compiled from 21 items covering three different aspects of employment protection: Individual dismissal of workers with regular contracts, additional costs for collective dismissals and regulation of temporary contracts. Range {0, 6} increasing with strictness of employment protection.

Net Union Density: Union density expresses union membership as a proportion of the eligible workforce. Normally, union density rates are standardized by the calculation of union membership as a proportion of the wage and salary earners in the same year (preferably on the basis of some annual average year data). The data are updated from the ILO website.

Public Sector employment: Data are collected from the laborsta.ilo.org website (ILO). These are data covering all employment of general governmental sector plus employment of publicly owned enterprises and companies. It covers all persons employed directly by those institutions. Based on this data, I compute an index of "public sector employment" by calculating the percentage of public employees over total working population, yearly, by country.

To address any further concern regarding the presence of endogeneity, I then implement an IV strategy. The potentially endogenous relative supply variable is instrumented using the "tertiary education institutional set-up" variables. Data are taken from Braga, Checchi, and Meschi (2011) and contains information about student financing and univesity autonomy and selectivity. For details about the construction of the indicators and the sources of the information they use see the paper available at: http://ftp.iza.org/dp6190.pdf

A.2 Additional tables and figures

Table A1: Descriptives by cohorts

	MALES				FEMALES			
	High relative supply		Low relative supply		High relative supply		Low relative supply	
	ECHP	EUSILC	ECHP	EUSILC	ECHP	EUSILC	ECHP	EUSILC
Age<=28								
College	27.18%	23.27%	5.80%	14.36%	40.52%	40.66%	12.49%	30.37%
Secondary	40.76%	49.16%	40.60%	47.92%	40.84%	44.55%	47.89%	47.34%
Low	32.06%	27.56%	53.60%	37.72%	18.63%	14.79%	39.62%	22.29%
Years of edu.	12.12	12.76	11.02	11.99	12.79	13.85	11.83	13.04
Log wage	9.02	9.48	8.78	9.19	8.86	9.33	8.62	9.08
Age 29-34								
College	39.07%	43.88%	20.42%	29.00%	53.24%	58.47%	27.33%	42.62%
Secondary	35.12%	39.34%	39.95%	44.00%	31.57%	32.65%	42.58%	39.65%
Low	25.81%	16.78%	39.63%	26.99%	15.20%	8.88%	30.09%	17.73%
Years of edu.	13.09	14.71	12.90	13.48	13.86	15.67	13.51	14.31
Log wage	9.59	9.98	9.26	9.65	9.32	9.68	8.95	9.39
Age 35-49								
College	37.64%	39.99%	21.12%	28.36%	46.56%	49.58%	26.79%	36.41%
Secondary	33.49%	40.99%	39.42%	42.15%	33.64%	37.82%	40.22%	42.36%
Low	28.87%	19.02%	39.46%	29.49%	19.80%	12.60%	32.99%	21.23%
Years of edu.	12.94	14.37	12.65	13.30	13.53	15.08	12.86	13.68
Log wage	9.75	10.12	9.38	9.85	9.42	9.75	9.05	9.46
Age 40-45								
College	37.60%	34.12%	21.23%	27.48%	43.40%	42.11%	27.58%	34.43%
Secondary	32.51%	42.32%	38.94%	41.41%	31.12%	40.68%	36.52%	41.87%
Low	29.89%	23.56%	39.83%	31.12%	25.48%	17.20%	35.90%	23.70%
Years of edu.	12.84	13.71	12.35	13.11	13.04	14.32	12.32	13.27
Log wage	9.84	10.15	9.45	9.94	9.47	9.83	9.13	9.55
Age 45								
College	37.51%	31.65%	22.79%	24.80%	39.36%	37.91%	29.11%	30.88%
Secondary	29.37%	41.13%	35.66%	43.64%	29.91%	39.42%	32.57%	43.25%
Low	33.12%	27.22%	41.55%	31.56%	30.73%	22.67%	38.32%	25.86%
Years of edu.	12.74	13.43	9.51	9.98	12.57	13.76	12.28	12.91
Log wage	9.91	10.21	47.95	48.05	9.52	9.88	9.22	9.62

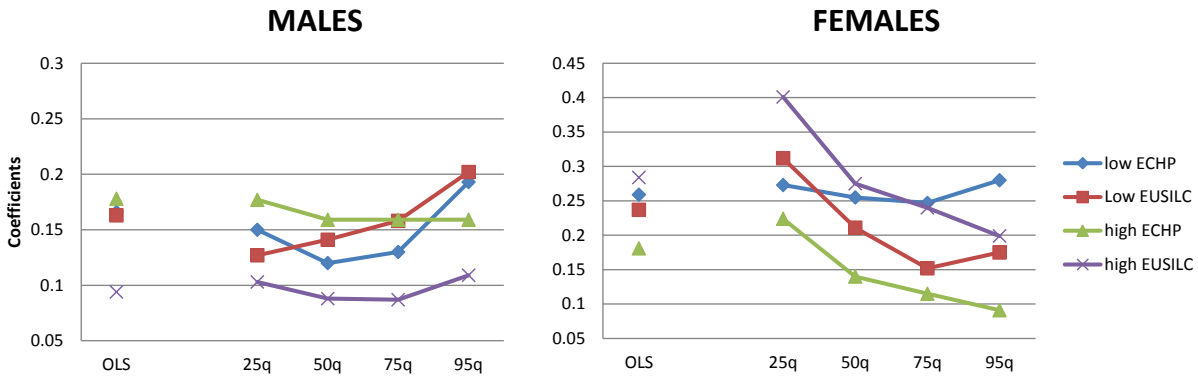
Table A2:Institutions by country

Country	Gini coefficient	Unemp. Rate (%)	Emp. Rate(%)	Relative supply	Relative demand	R&D Intensity	Emp. protection	Union density	Minimum wage	Wage compression	Pb.Emp(%)
Austria	0.25	4.16	69.14	0.08	0.19	5.63	2.09	34.52	0.00	3.16	13.12
Belgium	0.28	8.40	59.32	0.13	0.28	6.53	2.36	51.51	0.53	2.30	23.17
Germany	0.28	9.25	65.95	0.13	0.14	7.24	2.37	23.06	0.00	3.02	11.87
Denmark	0.22	5.34	74.95	0.20	0.12	7.60	1.55	72.35	0.00	2.48	33.66
Spain	0.29	15.60	56.67	0.21	0.68	2.30	2.98	15.74	0.43	3.34	15.96
Finland	0.25	10.68	66.56	0.15	0.77	8.43	2.05	73.63	0.00	2.31	27.49
France	0.28	10.10	62.27	0.20	0.22	9.27	3.02	7.94	0.57	2.90	24.23
Greece	0.32	9.63	58.01	0.15	0.53	0.89	3.16	25.75	0.48	3.24	21.83
Ireland	0.35	7.03	62.91	0.17	0.23	2.90	2.43	34.46	0.34	3.62	16.21
Italy	0.31	9.79	54.95	0.12	0.12	2.37	1.01	36.39	0.00	2.83	16.83
Portugal	0.36	6.53	67.13	0.09	0.77	0.95	3.56	21.61	0.51	4.98	12.89
United Kingdom	0.35	5.96	70.65	0.18	0.24	6.26	0.69	29.21	0.30	3.36	19.91

Table A3: Institutions by region

	Low relative supply		High relative supply	
	ECHP	EUSILC	ECHP	EUSILC
Gini coefficient	0.31	0.31	0.29	0.29
Unemployment rate (%)	8.29%	7.36%	11.68%	7.92%
Employment rate (%)	62.27%	64.41%	62.13%	67.48%
Relative supply	0.09	0.13	0.16	0.20
Relative demand	0.33	0.37	0.34	0.38
R&D intensity	2.47	3.92	5.24	7.03
Employnent protection	2.41	1.86	2.45	2.23
Union density	31.06	27.86	35.77	39.21
Miinimum wage	0.21	0.10	0.31	0.34
Wage compression	3.60	3.18	2.93	2.93
Public employment	15.74%	14.87%	23.80%	22.78%
Permanet contract	78.44%	88.63%	75.56%	87.82%

Figure A1: Within group inequality - secondary school degree. Quantile regressions



B Appendix to Chapter 2

B.1 Selection into Employment

Let the individual earnings function be given by $lnw_i = \gamma_0 S_i + \varepsilon_i$

Letting E_i be the probability of employment (and a positive wage), expected log wages conditional on employment are given by

$E\left[lnw_i | w_i > 0\right] = \gamma S_i + \varsigma\lambda(E_i)$

where λ is the inverse Mills ratio. Taking a first order Taylor approximation of λ and appending a zero mean error term υ to the earnings equation we obtain

$lnw_i = \gamma S_i + \delta E_i + \upsilon_i$ Aggregation by gender, cohort, country and year yields:

$lnw_{gact} = \gamma S_{gact} + \delta EP_{gact} + \upsilon_{gact}$

where EP is the employment - population ratio in the cell, which is closely related to the unemployment and activity rate in the cell.

B.2 Data Appendix

Our main data sources are:

SOEP (German Socio-Economic Panel): these data were kindly provided by DIW Berlin www.diw.de/soep/

EUSILC (European Survey on Living Standards): these data are available at the Department of Economics of the University of Padova upon formal approval by Eurostat. See for further details the website http://epp.eurostat.ec.europa.eu/portal/page/portal/microdata/eusilc

Data on country, age and gender specific unemployment rates and activity rates are from the Eurostat database http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/themes Since these data are only available by age group, we select three age groups - 45-49, 50-54 and 55-59 - and match these groups to the relevant cells defined by age, country and gender.

Data on parental background for the 23 European countries are taken from the ESS (European Social Survey). We have extracted the information for the relevant cohorts, countries and waves measuring parental background with the percentage of parents in the cell who have attained less than upper secondary education. Data are available online at: http://ess.nsd.uib.no/

B.3 Robustness checks

We have compared the returns to education for senior workers educated in Eastern and Western Europe by assuming that these returns are homogeneous among the countries within each area. Yet this restriction may not hold. We ask whether our qualitative findings still hold when we restrict the comparison to two groups of countries in the East and the West of Europe which share the same returns to education within each group. To identify these two groups, we estimate an augmented version of Eq.(2.11), which includes among the regressors the interactions of country dummies with the level and the between-genders difference in years of schooling, and select the sub-sample of countries for which the hypothesis of no joint statistical significance of these interactions cannot be rejected. It turns out that we need to exclude from the sample the following countries: Cyprus, Austria and Sweden in the West and Romania, Estonia and Slovakia in the East. When we do so, the qualitative results presented in Table 2.10 still hold .

In our definition of senior workers we have also included employed individuals who are aged between 51 and 55 at the time of the survey (2006-2008). A potential concern with this choice is that selection into retirement may affect average earnings and education in the oldest cells of the sample. To illustrate, the percentage of retired individuals is as low as 1 percent in the cells of individuals aged below 50, and equal to 7.9 and 4.1 percent respectively in the cells of Eastern and Western Europeans aged 51 to 55. In our estimates, we control for self-selection into employment and the labour force using the country by age by gender values of the unemployment rate and the activity rate. As a further control, we replicate our estimates of Eq. (2.11) by considering only cohorts born between 1956 and 1964, and find that we cannot reject the null hypothesis that the estimated returns to education earned by senior employees are not significantly different between Eastern and Western Europe. Therefore, we confirm the results presented in Table 2.10.

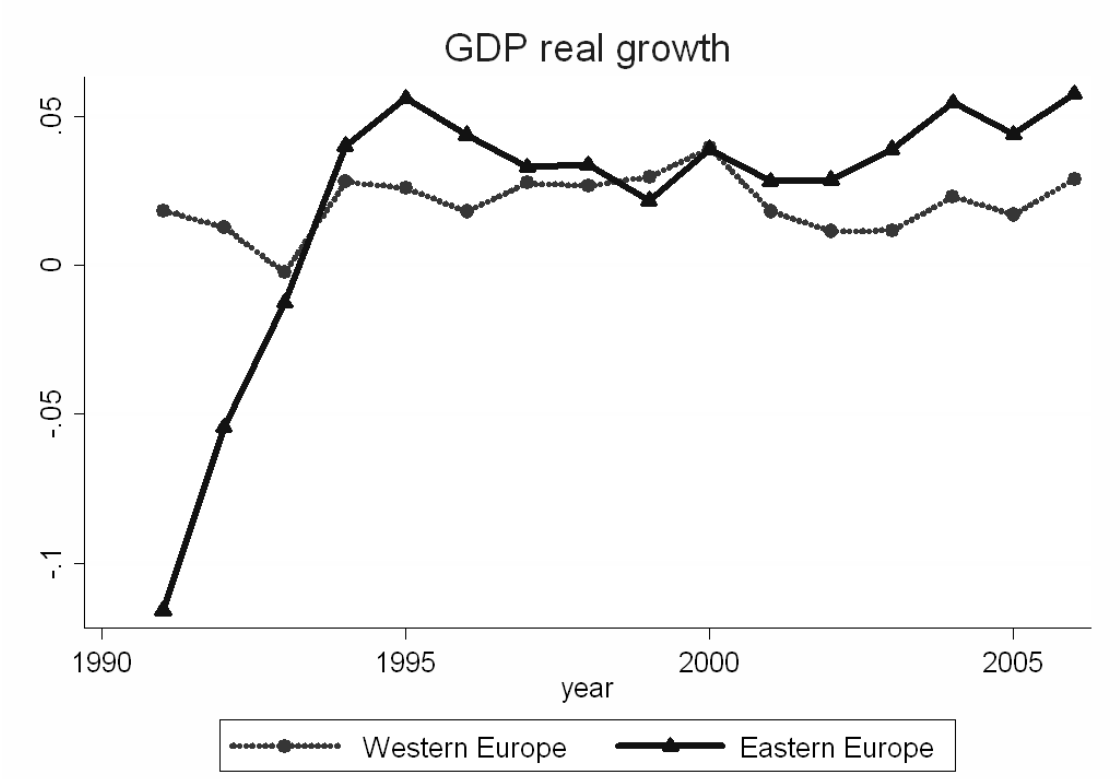
We also ask whether our estimates are affected by adding labour market experience, which in the EU-SILC data is defined as the number of years since the first regular job was started. In results available from the authors upon request, we show that adding the level and difference of experience to Eq. (2.11) does not alter our findings in a qualitative way. Moreover, when we test whether the level and difference of labour market experience is statistically significant, we cannot reject the null hypothesis of no significance.

Finally, we check the robustness of our finding to two additional variations: first, we remove agriculture from the data; second, we exclude those individuals who have completed

their schooling after 1987.[12]

B.4 Additional evidence.

Figure B1: Real GDP growth in Eastern and Western Europe. Source: Eurostat.



[12]In both cases, results available from the authors show no qualitative changes with respect to the baseline results presented above.

Table B1:Percentage retired and unemployed, by educational attainment. Cohorts born between 1951 and 1964, observed between 1994 and 1998. By gender, country of residence and level of education. Source: IALS

		Females				Males			
Region	Country	College degree	ISCED 3 academic	ISCED 3 vocational	Less than ISCED 3	College degree	ISCED 3 academic	ISCED 3 vocational	Less than ISCED 3
Western Europe	Ireland	1.5	2.8		18.7	10.5	17.0		19.3
	Netherlands	2.5	8.2	2.6	5.2	5.6	1.7	4.1	7.0
	Sweden	2.2	9.6		11.0	4.7	6.6		11.7
	Great Britain	8.1	9.3		11.4	1.8	6.7		9.3
	Belgium	2.9	7.2	19.3	35.7	2.4	3.0	10.5	5.4
	Italy	10.5	6.9	14.0	10.7	9.3	3.8	3.7	15.6
	Norway	0.6	2.9		6.6	0.2	2.4		0.0
	Denmark	2.4	15.2	9.8	27.2	4.3	8.5	2.6	14.7
	Finland	2.8	13.6	14.2	19.9	5.8	0.0	10.9	19.3
	West (average)	3.7	8.4	12.0	16.3	4.9	5.5	6.4	11.4
Eastern Europe	Poland	5.9	0.0	22.5	27.8	5.7	37.2	11.9	19.1
	Slovenia	5.0	13.1	10.7	21.7	0.0	10.0	9.5	18.9
	Czech Republic	5.4	14.7	4.8	11.2	2.2	0.0	6.2	12.3
	Hungary	6.2	10.2	16.5	34.2	4.6	10.8	24.3	38.4
	East (average)	5.6	9.5	13.7	23.7	3.1	14.5	13.0	22.2

Notes: for the comprehensive systems of Ireland, Sweden, Great Britain and Norway we report test scores in column *ISCED* 3 academic. Belgium includes only Flanders and Norway only the Bokmal region. In IALS, differently from EU-SILC, the category "disabled" is not included.

Table B2: Returns to schooling: selected OLS estimates (Marginal effects in earnings regressions).

Study	COUNTRY	PERIOD	CONTROLS	DATA	RESULTS
Andren et al. (2005)	Romania	1950-2000	Standard Mincerian controls (experience, experience squared).	Integrated Household Survey (IHS)	Men only Pre transition (1985-1989) 0.034 (0.003) Post transition (2000) 0.085 (0.001)
Bird, Schwarze and Wagner (1994)	Germany	1989-1991	Firm experience, work hours, civil service employment, firm size, white-collar employment, tenured civil servant status, and a dummy for married.	G-SOEP	Men only **East Germany** During communism 0.044 (0.003) During transition 0.041 (0.005) **West Germany** During transition 0.067 (0.003)
Campos and Jolliffe (2007)	Hungary	1986–2004	Firm size, industry dummies	WES	Both genders in 1989 0.073 (0.0004) Both genders in 2004 0.107 (0.0010)
Chase (1998)	Czech Republic and Slovakia	1984-1993	Standard Mincerian controls.	Social Stratification Survey	**Czech Republic** During communism Men 0.024 (0.0016) Women 0.042 (0.0018) During transition Men 0.052 (0.0031) Women 0.058 (0.0033) **Slovakia** During communism Men 0.028 (0.0021) Women 0.044 (0.0025) During transition Men 0.049 (0.0028) Women 0.054 (0.0029)
Denny and Doyle (2010)	Czech Republic, Hungary and Slovenia.	1994-1998	Age, dummy variables for gender, living in rural area, being a part-time worker and being self employed. Augmented specification including basic skills measure	IALS	**Both genders** **Czech Republic** 0.058 (0.005) with skills 0.049 (0.018) **Hungary** 0.058 (0.007) with skills 0.055 (0.008) **Slovenia** 0.078 (0.005) with skills 0.065 (0.006)

Flabbi et al. (2008)	Bulgaria Czech Republic Hungary Latvia Poland Russia Slovak Republic Slovenia	1991-2002	Dummies for living in urban areas and married, controls for current job (dummies for occupation, public employee, working full-time, member of a trade union), controls for current family (number of members, dummy for spouse working full-time).	ISSP	**Both genders** **Bulgaria** Early (1991-1996) 0.034 (0.001) Late (1997-2002) 0.044 (0.008) **Czech Republic** Early (1991-1996) 0.036 (0.006) Late (1997-2002) 0.042 (0.008) **Hungary** Pre (1986-1990) 0.049 (0.004) Early (1991-1996) 0.055 (0.002) Late (1997-2002) 0.069 (0.009) **Latvia** Early (1991-1996) 0.036 (0.007) Late (1997-2002) 0.039 (0.006) **Poland** Early (1991-1996) 0.054 (0.004) Late (1997-2002) 0.054 (0.010) **Russia** Early (1991-1996) 0.035 (0.007) Late (1997-2002) 0.051 (0.006) **Slovak republic** Late (1997-2002) 0.027 (0.001) **Slovenia** Early (1991-1996) 0.055 (0.006) Late (1997-2002) 0.051 (0.004)
Orlowski and Riphahn (2009)	Germany	2002-2006	Industry fixed effects, State fixed effects.	G-SOEP	**Men** **East Germany** Lower secondary: 0.005 (0.013) Upper secondary 0.04 (0.014) University 0,079 (0.009) **West Germany** Lower secondary: 0.042 (0.005) Upper secondary: 0.056 (0.003) University: 0.10 (0. 003)
Orazem and Vodopivec (1997)	Slovenia	1987-1992	Ethnicity , type of appointment (temporary or internship), industry fixed effects.	Administrative data base	**High school:** **Men** Pre transition 0.064 (0.0014) Post transition 0.081 (0.003) **Women** Pre transition 0.074 (0.0016) Post transition 0.090 (0.0036) **University:** Men pre transition 0.087 (0.012) **University:** Men pre transition 0.104 (0.0024) Men post transition 0.135 (0.005) Women pre transition 0.113 (0.002) Women post transition 0.137(0.0048)

Notes: Standard errors in parenthesis.

C Appendix to Chapter 3

C.1 Additional tables

Table C1: Descriptive statistics for boys and girls (restricted sample).

Variable	(1) Boys	(2) Boys At risk	(3) Girls	(4) Girls At risk
Name	Mean (sd)	Mean (sd)	Mean (sd)	Mean (sd)
Panel A: Outcome variables				
GCSE Point scores	402.98	404.33	433.72	426.08
	(142.35)	(151.86)	(133.31)	(140.72)
GCSE Standardized pt. scores	-0.056	-0.047	0.156	0.103
	(0.988)	(1.05)	(0.925)	(0.97)
Neet in W5	0.106	0.124	0.076	0.098
Panel B: Risky behaviour variables				
Ever smokes cigarettes.	0.177	0.234	0.251	0.360
Ever had alcoholic drink	0.694	0.731	0.690	0.758
Ever tried cannabis	0.255	0.303	0.211	0.323
Ever played truancy	0.065	0.102	0.057	0.108
Panel C: Mental Health variables				
GHQ at risk in w2	0.110	1	0.248	1
GHQ at risk in w4	0.161	0.412	0.298	0.541
GHQ likert in w2	8.496	18.0	11.218	19.56
	(4.822)	(4.70)	(6.13)	(5.10)
GHQ likert in w4	8.966	12.93	11.67	15.53
	(5.133)	(6.70)	(6.29)	(7.16)
GHQ likert in 0-1 w2	0.236	0.50	0.311	0.543
	(0.134)	(0.131)	(0.170)	(0.142)
GHQ likert in 0-1 w4	0.249	0.359	0.324	0.431
	(0.142)	(0.186)	(0.174)	(0.198)
Anxiety in w2	0.235	0.614	0.338	0.670
	(0.208)	(0.183)	(0.254)	(0.170)
Anxiety in w4	0.267	0.433	0.376	0.527
	(0.224)	(0.263)	(0.25)	(0.243)
Loss of confidence w2	0.139	0.5	0.239	0.572
	(0.212)	(0.29)	(0.282)	(0.29)
Loss of confidence w4	0.136	0.277	0.235	0.385
	(0.214)	(0.280)	(0.279)	(0.327)
Social dysfunction w2	0.269	0.423	0.317	0.449
	(0.124)	(0.157)	(0.135)	(0.158)
Social dysfunction w4	0.277	0.351	0.321	0.389
	(0.119)	(0.156)	(0.135)	(0.165)

continue

Panel D: Demographic controls				
White	0.736	0.723	0.709	0.707
Mixed eth.	0.047	0.064	0.048	0.060
Indian	0.068	0.047	0.056	0.044
Pakistani	0.047	0.044	0.050	0.044
Bangladeshi	0.029	0.033	0.043	0.044
Black Caribbean	0.026	0.020	0.031	0.032
Black African	0.026	0.040	0.029	0.030
Other	0.0182	0.267	0.029	0.036
1st lowest quantile of income	0.302	0.289	0.318	0.329
2nd lowest quantile of income	0.306	0.331	0.305	0.291
Highest quantile of income	0.392	0.378	0.376	0.379
Panel E: Key stage 2 variables				
Total score in science	56.86	56.79	56.23	57.47
	(17.68)	(19.58)	(17.69)	(17.51)
Total score in maths	64.305	65.19	61.30	63.09
	(24.65)	(25.703)	(24.09)	(24.11)
Total score in english	57.19	58.72	61.08	62.94
	(18.92)	(20.32)	(19.24)	(19.17)
Panel F: More detailed controls				
Mum with high qualification	0.255	0.285	0.230	0.265
Mum with gce a level	0.129	0.118	0.127	0.118
Mum Gcse grades a-c	0.262	0.251	0.263	0.250
Mum with Other qualification	0.088	0.08	0.091	0.088
Mum with No qualifications	0.186	0.178	0.204	0.187

continue

Dad with high qualification	0.200	0.216	0.197	0.210
Dad with gce a level	0.132	0.129	0.118	0.116
Dad Gcse grades a-c	0.166	0.146	0.166	0.161
Dad with Other qualification	0.048	0.046	0.051	0.045
Dad with No qualifications	0.149	0.124	0.142	0.128
Dad work full time	0.674	0.65	0.649	0.634
Dad work part time	0.030	0.031	0.30	0.033
Mum work full time	0.337	0.314	0.333	0.349
mum work part time	0.35	0.347	0.348	0.344
English spoken as main lang,	0.906	0.904	0.894	0.911
Married couple	0.709	0.665	0.705	0.68
Cohabiting couple	0.072	0.077	0.063	0.064
Lone parent	0.206	0.247	0.217	0.237
No parent	0.008	0.008	0.0082	0.005
Good_health (of parent in w1)	0.871	0.87	0.859	0.849
# of siblings	1.53	1.541	1.563	1.505
	(1.129)	(1.14)	(1.18)	(1.13)
Birth weight	3.149	3.082	3.014	3.02
	(1.056)	(1.13)	(0.975)	(0.969)
Single parent family at birth	0.178	0.207	0.168	0.168
Child with disability/illness	0.132	0.124	0.096	0.117
Child was born on time	0.424	0.432	0.458	0.454
Having at least one younger siblings	0.60	0.572	0.606	0.594
Having at least one elder siblings	0.58	0.628	0.586	0.597

Source: Authors' computation of LSYPE data.

Table C2: Descriptive Statistics. Sample comparison.

Variables	(1) Whole sample (13539)	(2) Completed MH Sample (8122)	(3) Restricted sample (7832)
Name	Mean (sd)	Mean (sd)	Mean (sd)
Panel A: Outcome variables			
GCSE Point scores	376.15	411.18	417.79
	(157.36)	(144.03)	(138.91)
GCSE Standardized pt. scores	0	0.222	0.264
	(1)	(0.91)	(0.88)
Neet in W5	0.105	0.093	0.091
Panel B: Risky behavior variables			
Ever smokes cigarettes.	0.210	0.213	0.212
Ever had alcoholic drink	0.59	0.68	0.69
Ever tried cannabis	0.217	0.233	0.233
Ever played truancy	0.067	0.062	0.061
Panel C: Key stage 2 variables			
Total score in science	52.05	55.21	56.55
	(20.29)	(19.25)	(17.68)
Total score in maths	56.56	61.20	62.86
	(26.84)	(25.93)	(24.43)
Total score in english	53.85	57.37	59.06
	(22.35)	(21.18)	(19.16)
No test taken in math	0.036	0.022	0
No test taken in science	0.030	0.019	0
No test taken in english	0.042	0.026	0
Panel D: More detailed controls			
Mum with high qualification	0.204	0.238	0.243
Mum with gce a level	0.114	0.126	0.128
Mum Gcse grades a-c	0.252	0.260	0.263
Mum with Other qualification	0.09	0.09	0.08
Mum with No qualifications	0.24	0.20	0.19

Continue

Dad with high qualification	0.163	0.196	0.199
Dad with gce a level	0.111	0.124	0.126
Dad Gcse grades a-c	0.160	0.166	0.166
Dad with Other qualification	0.052	0.05	0.05
Dad with No qualifications	0.166	0.148	0.146
Dad work full time	0.608	0.657	0.662
Dad work part time	0.033	0.03	0.03
Mum work full time	0.313	0.33	0.335
mum work part time	0.317	0.347	0.35
English spoken as main lang,	0.872	0.899	0.9
Married couple	0.667	0.704	0.707
Cohabiting couple	0.074	0.068	0.068
Lone parent	0.23	0.213	0.211
No parent	0.011	0.008	0.008
(of parent in w1)	0.841	0.863	0.866
# of siblings	1.62	1.56	1.54
	(1.23)	(1.17)	(1.15)
Birth weight	2.99	3.07	3.08
	(1.09)	(1.02)	(1.012
Single parent family at birth	0.192	0.176	0.173
Child with disability	0.128	0.120	0.115
Child was born on time	0.435	0.439	0.44
1st Lowest quantile of income	0.333	0.31	0.30
2nd lowest quantile of income	0.333	0.31	0.30
Highest quantile of income	0.333	0.387	0.39

Source: Authors' computation of LSYPE data.

Table C3: Mental health as outcome. Probit estimation- marginal effects.

	Boys		Girls	
	(1)	(2)	(4)	(5)
VARIABLES	Basic controls	Additional controls	Basic controls	Additional controls
Income: 2nd tercile	0.013	0.012	-0.045*	-0.050**
	(0.019)	(0.019)	(0.024)	(0.025)
Income: 3rd tercile (highest)	-0.004	-0.005	-0.039	-0.045*
	(0.018)	(0.018)	(0.025)	(0.026)
Ethnicity=mixed	0.038	0.035	0.054	0.039
	(0.027)	(0.027)	(0.036)	(0.036)
Ethnicity=Indian	-0.028*	-0.035*	-0.048	-0.048
	(0.017)	(0.018)	(0.030)	(0.033)
Ethnicity=Pakistani	0.001	-0.006	-0.018	0.009
	(0.027)	(0.030)	(0.034)	(0.045)
Ethnicity= Bangladeshi	0.031	0.029	0.024	0.051
	(0.035)	(0.042)	(0.049)	(0.061)
Ethnicity=Black Caribbean	-0.028	-0.022	-0.003	-0.010
	(0.027)	(0.029)	(0.042)	(0.043)
Ethnicity=Black African	0.051	0.035	-0.006	0.028
	(0.039)	(0.040)	(0.040)	(0.048)
Ethnicity=Other	0.059	0.056	0.059	0.066
	(0.049)	(0.048)	(0.043)	(0.047)
Mother: highest qualifc.=university	0.016	0.011	0.051**	0.013
	(0.018)	(0.018)	(0.025)	(0.026)
Mother: highest qualifcation=Alevel	-0.005	-0.005	0.002	-0.031
	(0.019)	(0.020)	(0.029)	(0.028)
Mother: highestQualification=GCSE	0.003	-0.001	0.005	-0.016
	(0.017)	(0.018)	(0.023)	(0.024)
Mother: highest qualifcation other	-0.005	-0.002	0.013	0.006
	(0.020)	(0.021)	(0.030)	(0.030)
Father: highest qualific=university	0.030	0.018	0.020	0.005
	(0.020)	(0.020)	(0.027)	(0.028)
Father: highest qualification=A-level	0.024	0.015	0.011	-0.005
	(0.023)	(0.021)	(0.030)	(0.030)
Father: highest qualification=GCSE	0.012	0.007	0.008	-0.001
	(0.020)	(0.020)	(0.027)	(0.027)
Father: highest qualification=other	0.018	0.006	-0.008	-0.012
	(0.030)	(0.028)	(0.038)	(0.038)
Step family		-0.009		0.021
		(0.019)		(0.031)
Married couple		-0.012		-0.025
		(0.024)		(0.034)
Lone father		0.112		0.061
		(0.163)		(0.138)
Lone mother		-0.062**		0.092
		(0.030)		(0.077)
No parents in the household		-0.039		-0.095
		(0.078)		(0.109)
Young person has a disability		-0.001		0.058**
		(0.014)		(0.025)
English is the main Language of household		-0.009		0.068**
		(0.025)		(0.031)

Continue

Father: works part-time		0.028		0.020
		(0.039)		(0.048)
Father: works full-time		0.026		-0.017
		(0.020)		(0.029)
Mother: works part-time		-0.014		0.026
		(0.013)		(0.021)
Mother: works full-time		-0.021		0.034
		(0.013)		(0.022)
Number of siblings		-0.001		-0.001
		(0.005)		(0.007)
Birth weight		-0.004		0.005
		(0.009)		(0.014)
Birth on time		0.008		0.002
		(0.010)		(0.016)
Single parent family(when child born)		0.017		-0.021
		(0.016)		(0.021)
Parent in good health		-0.009		-0.039*
		(0.016)		(0.022)
Total score in Science (KS2)		-0.002**		0.000
		(0.001)		(0.001)
Total score in Maths (KS2)		0.000		-0.000
		(0.000)		(0.001)
Total score in English (KS2)		0.002**		0.003**
		(0.001)		(0.001)
average point score (KS2)		0.004		0.004
		(0.004)		(0.005)
Nb. Of pupils on roll (KS3)		-0.000		-0.000
		(0.000)		(0.000)
% pupils with sen (KS3)		-0.004		0.009*
		(0.004)		(0.005)
% of 15 year old pupils achieving 5+ a*-c		-0.001**		0.000
		(0.000)		(0.001)
% pupils eligible for FSM		-0.001**		-0.001
		(0.001)		(0.001)
% pupils whose 1st language is not english		0.000		0.001
		(0.000)		(0.001)
Grammar school		0.013		-0.030
		(0.032)		(0.041)
Independent school		0.070		-0.247**
		(0.129)		(0.022)
Autonomous school		0.012		-0.014
		(0.012)		(0.018)
Special school		0.158		
		(0.262)		
Observations	4063	4047	3765	3765
Pseudo R2	0.00796	0.0304	0.00557	0.0262

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

Table C4: Standardized test scores as outcome. GHQ at risk as mental health measure. Full specification.

Panel A: BOYS

VARIABLES	(1) +parental education	(2) +KS2 attainment	(3) +secondary FE
GHQ at risk in W2	-0.001	-0.055	-0.083**
	[0.048]	[0.037]	[0.036]
Income: 2nd tercile	0.093	0.064	0.075*
	[0.058]	[0.048]	[0.045]
Income: 3rd tercile (highest)	0.279**	0.057	0.066
	[0.057]	[0.048]	[0.046]
Ethnicity=mixed	0.022	0.044	0.061
	[0.071]	[0.057]	[0.055]
Ethnicity=Indian	0.481**	0.416**	0.426**
	[0.071]	[0.064]	[0.060]
Ethnicity=Pakistani	0.157**	0.372**	0.412**
	[0.078]	[0.074]	[0.073]
Ethnicity= Bangladeshi	0.427**	0.440**	0.448**
	[0.110]	[0.105]	[0.098]
Ethnicity=Black Caribbean	-0.206**	0.075	-0.014
	[0.095]	[0.079]	[0.078]
Ethnicity=Black African	0.083	0.251**	-0.014
	[0.092]	[0.092]	[0.085]
Ethnicity=Other	0.112	0.256**	0.283**
	[0.128]	[0.108]	[0.092]
Mother: highest qualifc.=university	0.630**	0.210**	0.235**
	[0.061]	[0.045]	[0.043]
Mother: highest qualifcation=Alevel	0.464**	0.136**	0.158**
	[0.063]	[0.048]	[0.046]
Mother: highestQualification=GCSE	0.333**	0.105**	0.111**
	[0.057]	[0.043]	[0.040]
Mother: highest qualifcation other	0.075	0.021	0.010
	[0.071]	[0.056]	[0.049]
Father: highest qualific=university	0.424**	0.114**	0.115**
	[0.054]	[0.042]	[0.044]
Father: highest qualification=A-level	0.189**	0.070	0.054
	[0.058]	[0.046]	[0.046]
Father: highest qualification=GCSE	0.141**	0.063	0.050
	[0.053]	[0.042]	[0.043]
Father: highest qualification=other	-0.153*	-0.112*	-0.082
	[0.078]	[0.060]	[0.061]
Step family		-0.103**	-0.128**
		[0.052]	[0.047]
Married couple		0.000	0.041
		[0.055]	[0.052]
Lone father		0.035	-0.158
		[0.260]	[0.214]
Lone mother		-0.259**	-0.230*
		[0.115]	[0.118]
No parents in the household		0.152	-0.047
		[0.301]	[0.256]
Young person has a disability		-0.126**	-0.105**
		[0.038]	[0.034]
English is the main Language of household		-0.089	-0.087
		[0.067]	[0.057]

continue

Father: works part-time		0.115	0.149*
		[0.084]	[0.077]
Father: works full-time		0.076	0.079*
		[0.048]	[0.046]
Mother: works part-time		0.064*	0.035
		[0.035]	[0.033]
Mother: works full-time		0.035	0.031
		[0.038]	[0.034]
Number of siblings		-0.042**	-0.033**
		[0.012]	[0.012]
Birth weight		-0.003	-0.005
		[0.020]	[0.020]
Birth on time		-0.018	-0.020
		[0.024]	[0.024]
Single parent family(when child born)		-0.013	0.028
		[0.038]	[0.037]
Parent in good health		0.114**	0.104**
		[0.041]	[0.037]
Total score in Science (KS2)		0.006**	0.006**
		[0.002]	[0.002]
Total score in Math (KS2)		0.014**	0.013**
		[0.001]	[0.001]
Total score in English (KS2)		0.020**	0.020**
		[0.001]	[0.001]
average point score (KS2)		-0.051**	
		[0.009]	
Nb. Of pupils on roll (KS3)		0.000	
		[0.000]	
% pupils with sen (KS3)		0.015	
		[0.012]	
% of 15 year old pupils achieving 5+ a*-c		0.012**	
		[0.002]	
% pupils eligible for FSM		0.007**	
		[0.002]	
% pupils whose 1st language is not english		-0.003**	
		[0.001]	
Grammar school		-0.249**	
		[0.086]	
Independent school		-0.017	
		[0.531]	
Autonomous school		-0.573	
		[0.635]	
Special school		0.066	
		[0.041]	
Constant	-0.723**	-2.130**	-2.832**
	[0.074]	[0.286]	[0.135]
Observations	3923	3923	3923
Adjusted R-squared	0.131	0.478	0.569

Continue

Panel B: GIRLS

VARIABLES	(1) +parental education	(2) +KS2 attainment	(3) +secondary FE
GHQ at risk in W2	-0.086**	-0.148**	-0.158**
	[0.034]	[0.027]	[0.026]
Income: 2nd tercile	0.012	0.073*	0.057
	[0.054]	[0.043]	[0.042]
Income: 3rd tercile (highest)	0.241**	0.125**	0.121**
	[0.054]	[0.042]	[0.043]
Ethnicity=mixed	0.068	0.047	0.021
	[0.072]	[0.059]	[0.054]
Ethnicity=Indian	0.546**	0.406**	0.402**
	[0.075]	[0.065]	[0.061]
Ethnicity=Pakistani	0.289**	0.366**	0.400**
	[0.088]	[0.081]	[0.072]
Ethnicity= Bangladeshi	0.464**	0.390**	0.461**
	[0.091]	[0.089]	[0.088]
Ethnicity=Black Caribbean	-0.067	0.060	0.042
	[0.087]	[0.081]	[0.074]
Ethnicity=Black African	0.067	0.168**	0.260**
	[0.094]	[0.082]	[0.078]
Ethnicity=Other	0.422**	0.329**	0.336**
	[0.098]	[0.079]	[0.070]
Mother: highest qualifc.=university	0.593**	0.214**	0.217**
	[0.054]	[0.043]	[0.041]
Mother: highest qualifcation=Alevel	0.436**	0.139**	0.123**
	[0.057]	[0.045]	[0.045]
Mother: highestQualification=GCSE	0.320**	0.123**	0.106**
	[0.051]	[0.041]	[0.037]
Mother: highest qualifcation other	0.111*	0.053	0.038
	[0.060]	[0.047]	[0.047]
Father: highest qualific=university	0.391**	0.122**	0.136**
	[0.053]	[0.043]	[0.044]
Father: highest qualification=A-level	0.194**	0.029	0.072
	[0.056]	[0.043]	[0.046]
Father: highest qualification=GCSE	0.047	0.001	-0.010
	[0.051]	[0.041]	[0.043]
Father: highest qualification=other	0.045	0.099*	0.080
	[0.078]	[0.059]	[0.059]
Step family		-0.061	-0.025
		[0.045]	[0.047]
Married couple		0.057	0.087
		[0.057]	[0.054]
Lone father		0.284	0.243
		[0.219]	[0.196]
Lone mother		-0.192*	-0.203*
		[0.099]	[0.105]
No parents in the household		0.255	-0.073
		[0.221]	[0.224]
Young person has a disability		-0.095**	-0.115**
		[0.042]	[0.038]
English is the main Language of household		-0.149**	-0.124**
		[0.055]	[0.052]

continue

Father: works part-time		0.020	0.046
		[0.078]	[0.075]
Father: works full-time		0.050	0.029
		[0.046]	[0.045]
Mother: works part-time		0.068**	0.073**
		[0.032]	[0.033]
Mother: works full-time		0.055*	0.059*
		[0.032]	[0.034]
Number of siblings		-0.007	-0.011
		[0.012]	[0.011]
Birth weight		-0.013	0.007
		[0.020]	[0.021]
Birth on time		-0.051**	-0.043*
		[0.023]	[0.023]
Single parent family(when child born)		-0.010	-0.045
		[0.036]	[0.036]
Parent in good health		0.080**	0.053
		[0.040]	[0.035]
Total score in Science (KS2)		0.010**	0.009**
		[0.002]	[0.002]
Total score in Math (KS2)		0.013**	0.012**
		[0.001]	[0.001]
Total score in English (KS2)		0.017**	0.019**
		[0.001]	[0.001]
average point score (KS2)		-0.056**	
		[0.009]	
Nb. Of pupils on roll (KS3)		0.000	
		[0.000]	
% pupils with sen (KS3)		-0.002	
		[0.011]	
% of 15 year old pupils achieving 5+ a*-c		0.010**	
		[0.002]	
% pupils eligible for FSM		0.006**	
		[0.002]	
% pupils whose 1st language is not english		-0.003**	
		[0.001]	
Grammar school		-0.037	
		[0.161]	
Independent school		0.374	
		[0.244]	
Autonomous school		-0.011	
		[0.038]	
Special school		0.000	
		[0.000]	
Constant			
	-0.410**	-1.783**	-2.674**
	[0.073]	[0.265]	[0.132]
Observations	3644	3644	3644
Adjusted R-squared	0.135	0.478	0.573

Notes: Robust standard errors in brackets. ** p<0.05, * p<0.10.

C.2 GHQ Questions

Have you recently...

1. *Concen:* been able to concentrate on whatever you are doing?
2. *NoSleep:* Lost much sleep over worry?
3. *Useful:* felt you were playing a useful part in things
4. *Decide:* felt capable of making decisions about things
5. *Strain:* felt constantly under strain?
6. *Diffic:* feeling you couldn't overcome your difficulties?
7. *Activ:* been able to enjoy normal day to day activities?
8. *Probs:* been able to face up to your problems?
9. *Depress:* been feeling unhappy and depressed?
10. *NoConf:* been losing confidence in yourself?
11. *Wthless:* been thinking of yourself as a worthless person?
12. *Happy:* been feeling reasonably happy, all thing considered?

C.3 Graetz factors

The GHQ-12 has been extensively evaluated in terms of its validity and reliability as a uni-dimensional index of severity of psychological morbidity but the issue concerning the nature and the number of factors which are measured by the GHQ12 is still not completely clear. Many studies have assessed psychological morbidity in two or three dimensions. Several two- and three-dimensional models have been proposed, and to date no study examining the factor structure of the GHQ-12 has found it to be uni dimensional. A version of three-factor solution for the GHQ has been proposed by Graetz (1991). He shows that several advantages can be gained using multidimensional properties of GHQ as well as a single severity score. Using a large Australian sample of young people, and performing maximum likelihood factor analysis with oblique rotation , he reported the GHQ-12 measures three distinct constructs of "Anxiety", "Social dysfunction" and "Loss

of confidence". This three-dimensional model of Graetz has been identified the best fitting model by confirmatory factor analysis.

The three-dimensional model of Graetz is computed as follow:

- **Anxiety**= Obtained averaging the following GHQ questions:

Depress: been feeling unhappy and depressed?

Diffic: feeling you couldn't overcome your difficulties?

Strain: felt constantly under strain?

NoSleep: Lost much sleep over worry?

- **Loss of confidence**= Obtained averaging the following GHQ questions:

NoConf: been losing confidence in yourself?

Wthless: been thinking of yourself as a worthless person?

- **Anhedonia and Social Dysfunction**= Obtained averaging the following GHQ questions:

Decide: felt capable of making decisions about things

Useful: felt you were playing a useful part in things

Happy: been feeling reasonably happy, all things considered?

Probs: been able to face up to your problems?

Activ: been able to enjoy normal day to day activities?

Concen: been able to concentrate on whatever you are doing?