CERGE
Center for Economic Research and Graduate Education
Charles University Prague

Skill Intensity of Occupations, Labor Market Polarization, and Occupational Allocation of College Graduates

Barbara Pertold-Gębicka

Dissertation

Prague, April 2011

Barbara Pertold-Gębicka

Skill Intensity of Occupations, Labor Market Polarization, and Occupational Allocation of College Graduates

Dissertation

Prague, April 2011

Dissertation committee:

Štěpán Jurajda (chair, CERGE-EI)

Alena Bičáková (CERGE-EI)

Randall Filer (Hunter College, City University of New York)

Byeongju Jeong (CERGE-EI)

Referees:

Anna Rute Cardoso (Institute for Economic Analysis, Spain)

Andrea Weber (University of Mannheim, Germany)

Acknowledgements

I would like to thank my advisor for his patient guidance,
my husband for support and motivation,
and my parents for showing me the way.

Table of Contents

Abstract

The first chapter is motivated by the rapid expansion of higher education systems in Central European countries, where universities are largely state-funded and provision of higher education is a public policy decision. In this paper, I investigate an indicator of college skills usage – the fraction of college graduates employed in "college" occupations. Gottschalk and Hansen (2003) propose to identify "college" occupations based on within-occupation college wage premia; I build on their strategy to study the local-labor-market relationship between the share of college graduates in the population and the use of college skills. Empirical results based on worker-level data from Czech NUTS-4 districts suggest a positive relationship, thus supporting the presence of an endogenous influence of the number of skilled workers on the demand for them. Thus, the findings of this paper suggest that, in the long run, districts should be able to positively stimulate their labor markets by providing higher education to a larger fraction of their population.

In the second chapter, I propose a model-based measure of occupational skill intensity – a measure allowing me to consistently track technological progress on occupational level or to derive the demand for educated labor within different groups of occupations. I use the March CPS data from 1983 to 2002 to estimate such a measure corresponding to occupation-specific relative productivities of college and high-school educated. With imperfect substitution across skill types, the measurement of relative productivities requires estimation of substitution elasticities, and I propose a simple strategy to obtain these. The resulting measure is used to shed light on labor market polarization as documented by Goos and Manning (2007) and Autor et al. (2006). I show that in the 1980s technological progress was equally distributed across occupations from the whole spectrum of earnings distribution, but high earners sorted to more skill-intensive occupations and low earners sorted to less skill-intensive occupations. In the 1990s, there was no further reallocation and the least paying occupations experienced greater technological progress.

The last chapter provides further analysis of labor market polarization. I note that although much attention has been given to job polarization on national labor markets, there is little evidence on cross-country differences in the shape of employment changes distribution, which is used to depict polarization. This paper analyzes job polarization in 12 European countries using the skill-intensity measure developed in the second essay, which is independent of current labor supply conditions. I show that extensive north-south differences in the extent and skewness of polarization correspond to cross-country dissimilarities in employment protection policies.

Abstrakt

První část disertace je motivovaná výraznou expanzí vyššího vzdělávání ve Střední Evropě, kde university jsou převážně financované z veřejných zdrojů a poskytovaní vyššího vzdělání je tak součástí rozhodování ve veřejné politice. V této části disertace zkoumám indikátor použití vysokoškolsky vzdělané pracovní síly – podíl absolventů vysokých škol ve „vysokoškolských" pozicích. Gottschalk a Hansen (2003) navrhují identifikovat „vysokoškolské" pozice na základě mzdové prémie, kterou dostanou na této pozici vysokoškolsky vzdělaní lidé. Navazuji na jejich práci a studuji lokální vztah mezi podílem vysokoškolsky vzdělaných lidí v populaci a využitím vysokoškolsky vzdělané pracovní síly. Empirické výsledky založené na analýze individuálních dat z českých regionů na úrovni NUTS-4 naznačují existenci pozitivního vztahu, tudíž potvrzují přítomnost endogenního vlivu počtu vzdělané pracovní síly na poptávku po ní. Z těchto závěru vyplývá, že v dlouhém období regiony by měly být schopny pozitivně stimulovat lokální trh práce pomocí poskytování vyššího vzdělání většímu podílu populace.

V druhé části navrhuji jak odhadnout na základě modelu míru kvalifikační náročnosti zaměstnání. Tato míra dovoluje konzistentně zachytit technologický pokrok na úrovni zaměstnání. V této části používím americké March CPS data z let 1983 až 2002 pro odhad míry kvalifikační náročnosti zaměstnání, která odpovídá relativní produktivitě středoškolsky a vysokoškolsky vzdělaných pracovníků na jednotlivých pozicích. Za předpokladu, že substituce mezi pracovníky s různou kvalifikací je nedokonalá, měření relativní produktivity vyžaduje odhad elasticity substituce. V této části navrhuji jednoduchou strategii jak tuto elasticitu odhadnout. Výsledek je poté využit pro analýzu polarizace trhu práce jak je dokumentovaná v Goos a Manning (2007) a Autor er al. (2006). Ukazuji, že technologický pokrok v osmdesátých letech byl rovnoměrný mezí zaměstnání z celého spektra rozdělené mezd, ale ti s vyššími mzdami se začleňují do zaměstnání s vyšší kvalifikační náročností, kdežto ti s nižšími mzdami do zaměstnání s nižší kvalifikační náročností. V devadesátých letech již nebylo toto rozčleňování a zaměstnání s nižšími mzdami byly více ovlivněny technologickým pokrokem.

Poslední část poskytuje podrobnou analýzy polarizace trhu práce ve 12 evropských zemí. Pro tuto analýzu využívám míru profesní náročnosti zaměstnání v takové podobě, jak byla odhadnuta v druhé části. Ukazuji, že velké rozdíly mezi severními a jižními evropskými zeměmi v rozsahu a rozdělení polarizace koresponduje s rozdíly v rozsahu institucionální ochrany zaměstnanců. Tento článek tak přispívá do současné literatury, která se zatím koncentrovala pouze na polarizaci trhu práce na národní úrovni.

General Introduction

This dissertation consists of three chapters investigating technological progress and its impact on the labor market at the occupational level. The occupation-focused research is motivated by the inability to consistently explain recent trends in the wage and employment structure in the developed world by the traditional framework with high- and low-skilled workers supplying labor to a homogeneous labor market with factor-augmenting technology.

The first chapter focuses on the allocation of college graduates across occupations with special attention given to college skills utilization within individual occupations. This topic is especially interesting in the context of countries where the majority of higher education institutions are public. As both the over- and undersupply of college seats could result in efficiency losses for society, there is a need to understand the forces shaping the demand for skilled labor to inform policy decisions concerning the provision of higher education.

The starting point for my analysis is the study of Gottschalk and Hansen (2003) who propose a methodology for classifying occupations into "college" and "noncollege" based on a rigorous, though simple, model. They assume that "noncollege" occupations do not value college-gained skills and thus pay none or a very small wage premium to college graduates, while "college" occupations pay a significant college wage premium. This property allows me to order occupations according to their estimated returns to college and to classify as "college" those occupations that fall above a certain threshold. Several studies follow this approach to measure the fraction of college graduates employed in "noncollege" occupations in the U.S. (Gottschalk and Hansen, 2003), Portugal (Cardoso, 2007), and the U.K. (Grazier, 2008). These papers analyze the time trend of the overskilling measure at the aggregate (country) level.

To better understand the mechanisms behind college graduates' allocation across occupations, I extend the framework proposed by Gottschalk and Hansen (2003) to explicitly model the relationship between the number of college graduates available in the labor market and the fraction of them working in the so called "noncollege" occupations. Using this framework as the baseline, I estimate this relationship using the within-country cross-regional variation in the fraction of college graduates working in "noncollege" occupations. This approach not only allows me to use more data points than the time trend studies but also makes it easier to break the simultaneity between the number of college graduates in the market and their occupational allocation. Nevertheless, I also present an analogous analysis on a panel of regions within the country. Interestingly, the relationship of interest is found to be negative in the cross-regional analysis and positive in the over-time analysis. These results suggest that the long-run equilibrium is shaped by the endogenous influence of the number of skilled workers on the demand for them. In the short run, however, the endogenous effect is not strong enough to compensate for movements along the demand curve.

In the second chapter, I employ a measure of the skill intensity of occupations to investigate the differences between wage structure trends in the U.S. in the 1980s and 1990s. The occupation-focused literature, represented by the works of Autor et al. (2003, 2006) and Goos and Manning (2007), proposes a modified version of the skill-biased technological change (SBTC) hypothesis to explain the varying extent of skill-biased technological progress or the recently documented earnings growth polarization. They argue that new technologies have a heterogeneous impact on workers. In particular, technologies complement workers performing non-routine cognitive tasks and substitute for workers performing routine tasks. Thus, to the extent that occupations capture the task content of work, the occupation-level analysis offers a key towards understanding the impact of technological change on wage structures. In the context of technological progress and the related demand for skilled labor, it is helpful to link occupations to their skill intensity. The latter would translate the occupation-specific task mix into the demand for skills defined by an occupation-specific production function.

I propose to use the within-occupation relative productivity of college and high school graduates, where college graduates represent highly-skilled labor and high school graduates represent less skilled-labor as a proxy for occupation-specific skill intensity. The second chapter develops a strategy for estimating the skill-intensity measure and applies it to analyze the recent polarization of earnings growth in the U.S. Polarization has been defined in the literature as employment or earnings growth in low- and high-skilled occupations at the cost of middle-skilled occupations. Interestingly, it is only observed since the 1990s. In the earlier decades, earnings at the low end of the distribution were falling and those at the top end were increasing. To shed more light on these differences, I use the measure of skill intensity of occupations to show that in the 1980s technological progress was equally distributed across occupations from all the earnings distribution, but high earners sorted to more skill-intensive occupations and low earners sorted to less skill-intensive occupations. In the 1990s, however, there was no further reallocation and the least-paying occupations experienced greater technological progress.

The last chapter focuses on labor market polarization in Europe. I use the European Union Labor Force Survey to report differences in the extent of job polarization (job polarization is defined as growth of employment in high- and low-skilled occupations with simultaneous decrease, or stagnation, of employment in middle-skilled occupations) across European countries, adopting the measure of the skill requirements of occupations developed in the second chapter. This is a preferable measure to document polarization across countries, as it is independent of supply conditions in local labor markets. The discussion and examples provided in the current study confirm this statement. With the use of the skill requirements measure, I provide extensive evidence on cross-country differences in the extent of polarization. Specifically, one can observe that polarization is strongest in the Southern European countries and Ireland, while it is weaker in Northern Europe. As a potential explanation for this observation, I suggest differences in the economic growth and educational attainment of their populations. The remaining cross-country variation in the extent of polarization might be partially driven by dissimilarities in labor market institutions.

CHAPTER 1

College Degree Supply and Occupational Allocation of Graduates - the Case of the Czech Republic

Barbara Pertold-Gębicka*

CERGE-EI†

Abstract

Public funding drives much of the recent growth of college degree supply in Europe, but few indicators are available to asseses its optimal level. In this paper, I investigate an indicator of college skills usage – the fraction of college graduates employed in "college" occupations. Gottschalk and Hansen (2003) propose to identify "college" occupations based on within-occupation college wage premia; I build on their strategy to study the local-labor-market relationship between the share of college graduates in the population and the use of college skills. Empirical results based on worker-level data from Czech NUTS-4 districts suggest a positive relationship, thus supporting the presence of an endogenous influence of the number of skilled workers on the demand for them.

---

*I am grateful to Štěpán Jurajda for his guidance and advice. I also thank Alena Bičáková, Daniel Münich, Randall Filer and CERGE-EI Brown Bag Seminar participants for helpful hints and comments. All remaining errors are mine. This research has been supported by the World Bank Research Fellowship grant. E-mail: barbara.gebicka@cerge-ei.cz

†A joint workplace of the Center for Economic Research and Graduate Education, Charles University, Prague, and the Economics Institute of the Academy of Sciences of the Czech Republic. Address: CERGE-EI, Politickych veznu 7, Prague 11121, Czech Republic.

1 Introduction

While primary and some form of secondary education is available for the vast majority of citizens in the developed countries, higher education is only accessible to a limited number of people. These limits are partially driven by public funds devoted to higher education, which is especially binding in countries where the majority of higher education institutions are public. As both the over- and undersupply of college seats could result in efficiency losses for society, there is a need to understand the forces shaping the demand for skilled labor to inform policy decisions concerning the provision of higher education.

Recent economic literature has approached the topic of optimal level of college degree supply by analyzing different indicators of college skills utilization. The most straightforward, analyzing social returns[1] to higher education (Acemoglu and Angrist, 2000; Moretti, 2004), directly captures the benefits of educating people, but is difficult to measure. An alternative is offered by the overskilling literature (see McGuinness, 2006 for a review), which investigates employment of college graduates in the so-called "noncollege" occupations (Pryor and Schaffer, 1997; McGuinness and Bennett, 2007) in order to quantify the oversupply of college skills. This line of research offers an easy to measure indicator of college skills usage which is not, however, supported by an economic model. Only recently, Gottschalk and Hansen (2003) proposed a methodology for classifying occupations into "college" and "noncollege" based on a rigorous, though simple, model. This equips us with a more reliable tool to measure the fraction of college graduates employed in "noncollege" occupations – an indicator useful in assessing whether changes in the supply of skilled labor meet changes in the demand for them. In this paper, I use the measure of college graduates employed in "noncollege" occupations, as proposed by Gottschalk and Hansen (2003), to find out whether an increased number of college graduates attracts firms

[1] There is also a vast stream of literature on private returns to higher education, known as the college wage premium, and their connection to the relative supply and demand for skilled labor (Bound and Johnson, 1992; Katz and Autor, 1999; Fortin, 2006). As the college wage premium is a relative measure of returns to higher education, it is not informative of the absolute demand for college graduates.

using advanced technologies and thus triggers a shift in the demand for skilled labor.

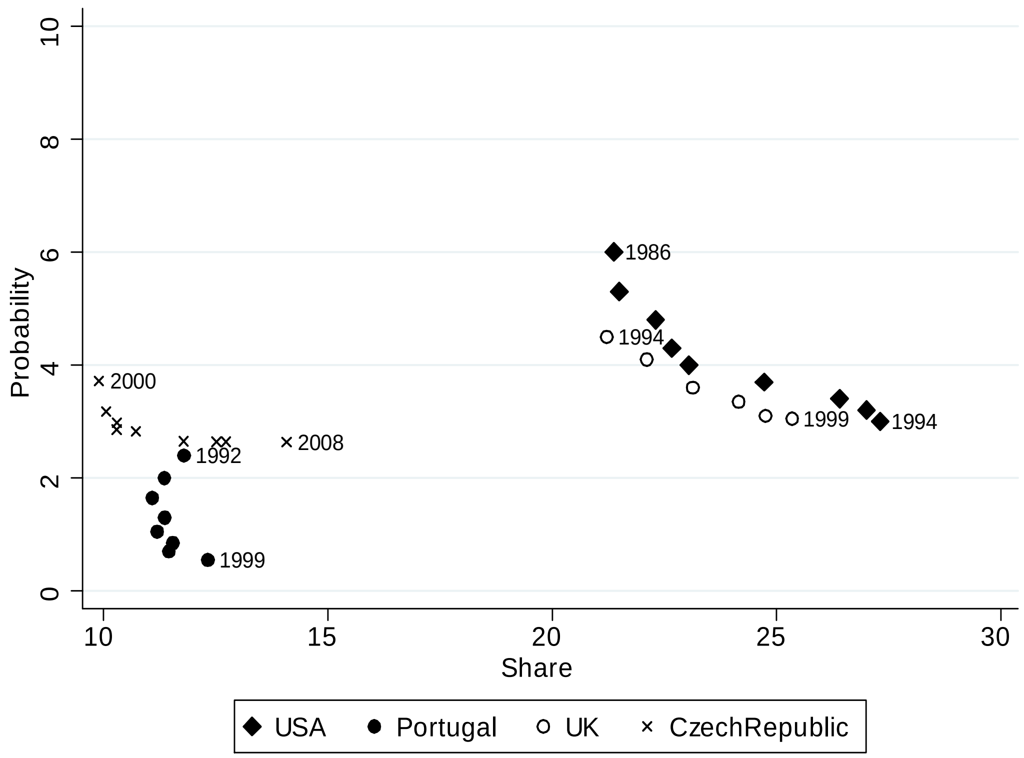
The model proposed by Gottschalk and Hansen (2003) assumes that "noncollege" occupations do not value college-gained skills and thus pay none or very little wage premium to college graduates, while "college" occupations pay a significant college wage premium. This property allows us to order occupations according to their estimated returns to college and to classify as "college" those occupations which fall above a certain threshold. Several studies follow this approach to measure the fraction of college graduates employed in "noncollege" occupations in the U.S. (Gottschalk and Hansen, 2003), Portugal (Cardoso, 2007), and the U.K. (Grazier et al., 2008). These papers only analyze the time trend of the overskilling measure at the aggregate level. It would be more informative, however, to see whether the extent of overskilling is correlated with the number of college graduates in the economy. This relationship is depicted in Figure 1, which plots the probability of a young college graduate to be employed in a "noncollege" occupation,[2] as reported by the authors of the above-mentioned articles, against the fraction of college graduates in the young population.[3] This figure also presents an analogous relationship for the Czech Republic, a country which is analyzed in more detail in this paper.

Two features stand out in Figure 1. First, within a country the probability of a college graduate to work in a "noncollege" occupation is negatively correlated with the fraction of college graduates in the population. Second, in countries with a higher proportion of highly educated people in the population, the likelihood of observing a college graduate work in a "noncollege" occupation is higher. The latter observation could be an artifact of the constant college wage premium threshold used in these studies to distinguish between "college" and "noncollege" occupations. It is generally understood that economies with a relatively low endowment of skilled labor report high college premia (Brunello et al., 2000; Card and Lemieux, 2001), which could be reflected in more occupations being classified as "college" in these countries.

[2] The probability of being employed in a "noncollege" occupation is a disaggregated measure of the fraction of college graduates employed in "noncollege" occupations.

[3] A young population is defined as 20-39 years of age.

Figure 1: Propensity of a college graduate to work in a "noncollege" occupation vs. the share of college graduates in the labor force across countries



Source: Own compilation using Gottschalk and Hansen (2003), Cardoso (2007), Grazier et al. (2008), Eurostat, and U.S. Census Bureau as well as the ISPV data.

More robust and more interesting is the negative within-country correlation between the fraction of college graduates in the population and the probability of a college graduate to work in a "noncollege" occupation. Following a simple supply-demand analysis, one would expect the opposite relationship.[4] Thus, it is tempting to interpret this feature as the positive influence of an increased number of skilled workers on the number of skill-intensive positions offered by firms (i.e., as a spillover effect). Yet, the observed correlation could be spurious and reflect just the simultaneous reaction of the demand and supply side of the labor market for college graduates to positive technological shocks.

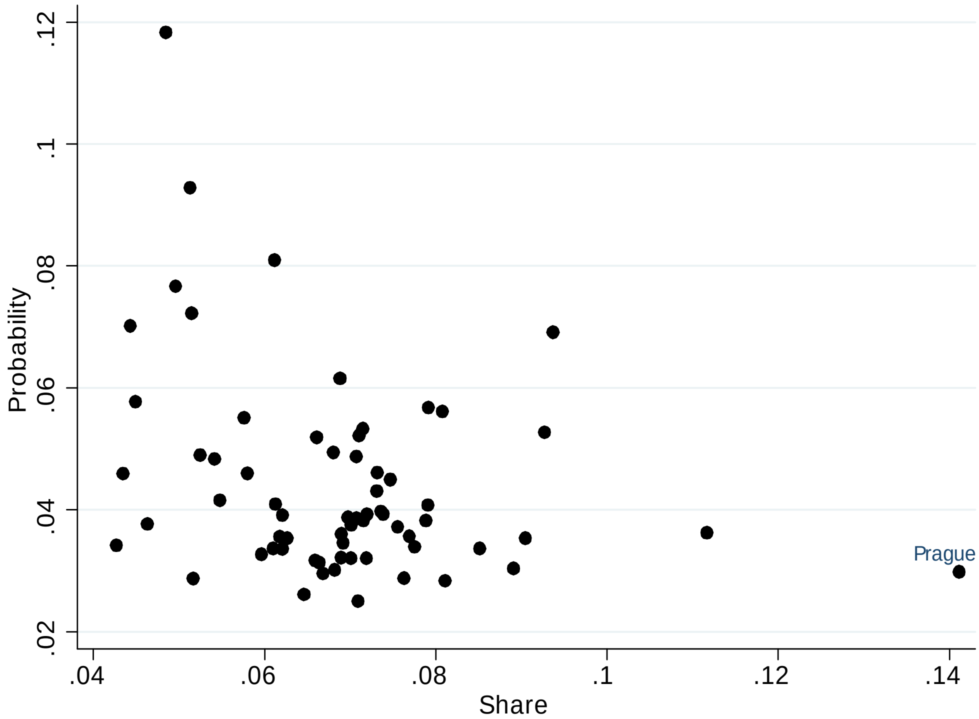
To better understand the patterns observed in Figure 1, I extend the Gottschalk

[4]This is a consequence of movement along a downward-sloping demand curve.

and Hansen (2003) setup to explicitly model the relationship between the number of college graduates available in the labor market and the fraction of them working in “noncollege” occupations. Instead of working with an aggregate time trend, I estimate this relationship using the within country cross-regional variation in the fraction of college graduates working in “noncollege” occupations. This approach not only allows me to use more data points but also makes it easier to break the simultaneity between the number of college graduates in the market and their occupational allocation. As presented in Figure 2, cross-regional patterns are similar to those observed within a country over time. For comparison, I also present an analysis on a panel of regions within the country. Interestingly, the relationship of interest is found to be negative in the cross-regional analysis and positive in the over-time analysis. These results suggest that the long-run equilibrium is shaped by the endogenous influence of the number of skilled workers on the demand for them. In the short run, however, the endogenous effect is not strong enough to compensate for movements along the demand curve. Thus, the patterns observed in Figure 1 might be driven by exogenous technological shocks.

The analysis presented in this paper concentrates on the Czech Republic. This Central European country is especially interesting because its higher education system has been expanding rapidly but unequally in recent years, resulting in significant between-year and across-region variation in the educational structure of the population. Moreover, as Central European countries are still lagging behind the Western economies in terms of technological development, there is a lot of opportunity for technological progress to happen and advanced capital to flow in. Finally, the choice of the Czech Republic adds policy relevance to this research. The higher education system in this country is largely state-funded and thus the provision of college education is a public policy decision. Awareness of the channels which affect the demand for college-educated labor would facilitate decision-making concerning the extent of higher education expansion. In the absence of the endogenous effect college enrolments should simply reflect the trend in technological progress of the economy, while the existence of this effect implies that increasing the educational attainment of the

Figure 2: Propensity of a college graduate to work in a "noncollege" occupation vs. the share of college graduates in the young population across Czech districts in 2001



Note: Young population consists of people below the age of 35. Source: Own calculations using 2001 Census and the ISPV data.

local population could be used to attract advanced technologies and to increase the skill bias of the economy.

The reminder of the paper is organized as follows. Section 2 places this study in the context of the existing literature and Section 3 describes higher education in the Czech Republic. The theoretical and empirical models of college and high school graduates' allocation across different occupations are described in Sections 4 and 5, respectively, followed by a definition of "college" and "noncollege" occupations in Section 6. Estimation of the causal relationship between the relative stock of college graduates and the fraction of them working in "noncollege" occupations is then discussed. Section 8 concludes.

2 Demand for College Graduates in the Literature

Several streams of literature are related to this paper. First, Acemoglu (2002, 2003) suggests that the extent of the skill bias of technology, and thus the demand for skills, can be shifted endogenously by intense international trade and by the presence of many skilled workers. Similar conclusions are reached by Moretti (2004), who shows that a high concentration of college-educated workers in a city's population has a positive effect on wages of all education groups in that city, including the college graduates. This implies the existence of positive productivity spillovers from the spatial concentration of skills and suggests that a large number of college graduates in a labor market can trigger a shift in the demand for them. Fortin's (2006) findings of a negative relationship between the production of college graduates and the college-high school wage gap across the U.S. states suggest that the positive effect of a high concentration of college graduates on local wages is stronger for high school-educated workers. These findings are challenged by Bound et al. (2004), who find that the production of college graduates in U.S. states does not correspond to their stock, because of a significant level of migration. If this is also true for the Czech Republic, the policy implications of the present study could be limited. Nevertheless, it is generally known that in Central Europe both the within-country and across-countries mobility of labor is much lower than in the U.S. (e.g., Fidrmuc, 2004) and enrolments in higher institutions translate into a future supply of college graduates to local labor markets in these countries. Thus, to identify potential endogenous shifts in the demand for labor, I follow Moretti (2004) and investigate the relationship between the presence of college-educated individuals in the economy and the demand for skilled labor. However, instead of analyzing college graduates' wages, I investigate their occupational allocation as the indicator of college skills usage.

Occupational allocation of college graduates is the central focus of another stream of literature related to this paper, widely known as the overeducation (overskilling) literature. Studies in this field measure the fraction of college graduates employed in occupations not requiring a college degree and estimate the wage effects of being employed in such an occupation. They find that the incidence of overeducation is increasing over time (Walker and Zhu, 2005, evidence for the U.K.) and that it is associated with a significant wage punishment (McGuinness, 2006, a metastudy) which, however, is largely reduced if individual heterogeneity is taken into account (Bauer, 2002, evidence for Germany). This literature typically classifies individuals as being overskilled if they work in an occupation which has the median (average) year of schooling lower than that of the individual, has an official schooling requirement, as defined in the job description, lower than that of the individual, or is assessed by the individual to require lower skills than she has. While this line of research studies a phenomenon directly reflecting the demand for college graduates, it suffers from the lack of an economic model supporting the measures of overskilling. This gap is filled by Gottschalk and Hansen (2003), who develop a simple supply-demand framework which models the allocation of college graduates between "college" and "noncollege" occupations. I depart from their model when investigating the occupational allocation of college graduates.

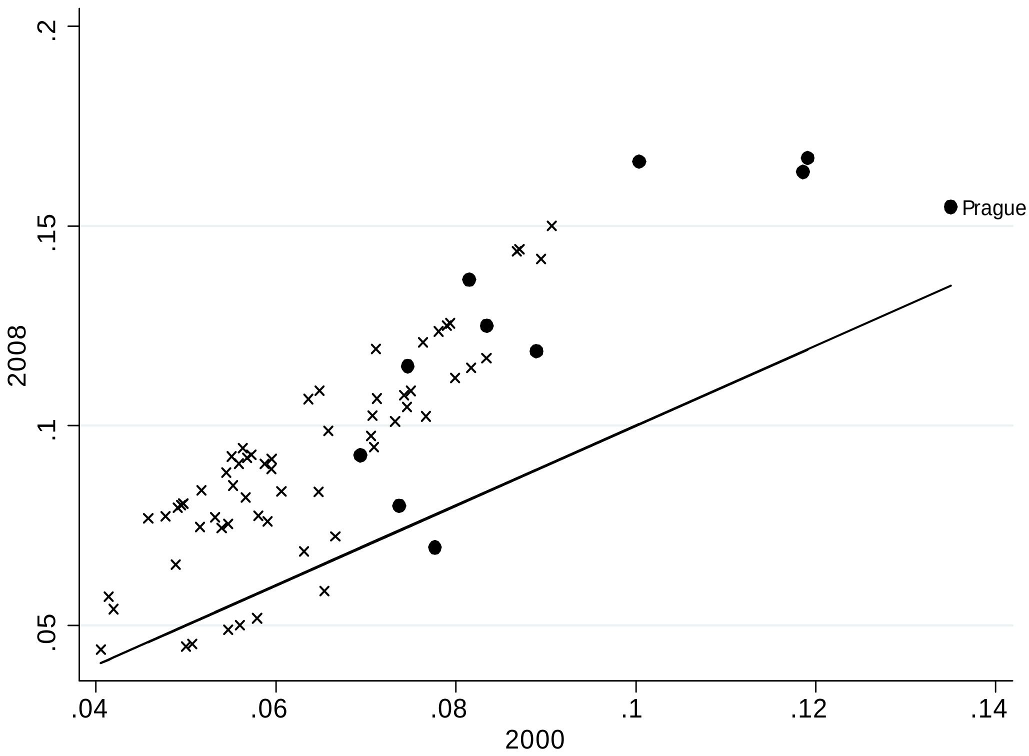
Research on the demand for college-educated workers has not been that extensive in the context of Central Europe. The only comparative study by Flabbi et al. (2008) shows that the returns to education were increasing or stayed constant in several Central and Eastern European countries throughout transition. Analyses concentrating on the Czech Republic in particular, e.g., Filer et al. (1999), Jurajda (2005), Munich et al. (2005), also confirm this finding. Another study by Jurajda (2004) shows that college graduates' wages are insensitive to their concentration across Czech districts. In a related work, Jurajda and Terrell (2007) find that significant differences in unemployment rates across regions of post-communist economies can be to a large extent explained by variations in local human capital endowment. Additionally, they show that FDI flows to regions are characterized by higher human capital endowment, which is in line with Acemoglu's hypothesis of endogenous technological progress. My study falls into this line of research, as it investigates the relationship between the educational structure of the local population and the labor market situation of college graduates.

3 The Czech Republic

The analysis presented in this chapter focuses on the employment of college graduates in the Czech Republic. This country is particularly interesting for its organization of tertiary education. The majority of Czech public universities were established under communism and underwent restructuring only in the 1990s. Yet, the mass expansion of college enrolments happened much later, with the most significant increase happening in the last decade (CSO, 2009). The growth in college enrolment and the resulting increased inflow of graduates is changing the educational profile of the Czech population. The fraction of college graduates in prime-age population (25 - 54 years of age) is growing – from 11% in 2000 to 14% in 2008 (Eurostat, 2009). This growth is even more visible in the young population (up to 35 years of age) – between 2000 and 2008 the share of college graduates in the young population increased from 8% to 19% (CSO, 2009). Despite these changes, the fraction of the prime-age population with higher education is still very low in the Czech Republic as compared to other countries. The OECD average fraction of college graduates among the prime-age population was 27% in 2006 (OECD, 2009) with the U.S. having the highest number (39%). International comparison suggests that the Czech Republic will experience a further increase in its proportion of the highly educated in the years to come in order to catch up with other countries. Thus, it is important to know how these changes shape the labor market. As an illustration, Figure 3 presents district-level[5] changes in the shares of highly educated in the young population between 2000 and 2008.

[5]Districts are NUTS-4 (Nomenclature of Statistical Territorial Units of the European Union) regions with populations of fewer than 150,000 individuals.

Figure 3: Changes in the fraction of college graduates in Czech NUTS-4 districts' young population between 2000 and 2008 together with a 45-degree line



Note: Full circles denote districts which had a college by the end of communism, while crosses denote districts which did not have a college at that time. Growth rates are aggregated at region-level (NUTS-3) due to representative data availability. Young population consists of people below the age of 35.
Source: Own calculations using 2001 Census and the 2000-2008 Czech Labor Force Survey.

Two major forces might be driving the changes observed in Figure 3 – differential provision of higher education, and cross-district migration of college graduates. As tertiary education in the Czech Republic is largely state funded (OECD, 2006), the supply of places in tuition-free colleges (which is significantly lower than the demand for them) is determined by the funds allocated by policy makers. Public universities are financed on a by-student basis, but they are restricted to increasing enrolments by no more than a specified percentage as compared to the previous academic year. This results in a very diversified educational structure of Czech districts' populations.

Differences in the fraction of the adult population with tertiary education and the rates of growth in this measure are strongly determined by the initial (i.e., before the transition) distribution of colleges across the country. It stands out in Figure 3 that districts which had a college established by the end of communism are characterized by significantly larger shares of a highly-educated population. This is used as an exclusion restriction when identifying the influence of the relative supply of college graduates on their fraction working in "noncollege" occupations, which is discussed in detail in Section 4.

Table 1: Public colleges and universities in the Czech Republic in the 1969-2008 period

Year	Number of public colleges and universities	Total number of students	Newly admitted	Graduates
1969	24	84 784	15 563	8 814
1974	23	85 608	17 467	11 040
1979	23	116 141	22 254	14 657
1983	23	114 529	21 636	18 828
1989	23	112 980	22 944	16 069
1993	23	127 137	30 964	14 896
1995	23	148 235	36 769	16 603
1996	23	166 135	39 782	18 398
1998	23	186 730	41 981	23 043
1999	23	198 961	40 794	23 582
2001	24	219 514	42 604	29 156
2003	24	256 408	49 873	31 503
2005	25	271 940	55 246	39 376
2007	26	303 731	60 766	53 635
2008	26	319 615	62 952	60 183

Source: Czech Statistical office (CSO), 2010

During the communist regime the number of higher education institutions in the Czech Republic was constant, as presented in Table 1. Shortly after the change of regimes, expansion of higher education took the form of increased enrolments at existing universities. New public universities were established as late as in the 21st century: Tomas Bata University in Zlin (2001), Technical University in Jihlava (2005), and Technical and Economic University in Ceske Budejovice. Simultaneously, first private colleges and universities were established. In 2000 there were 8 private colleges in the Czech Republic: four in Prague and one in Kunovice, Karlovy Vary, Mlada Boleslav, and Ostrava, respectively. The first 447 graduates of private colleges entered the labor market in 2002. In 2008 there were already 45 private colleges and universities and 9451 of their graduates entered the market (CSO, 2010). As presented in Table 8 in the appendix, the majority of private higher education institutions are located in Prague.

District-specific production of college graduates is almost directly translated into the number of skilled workers in local labor markets because of low cross-district migration of graduates.[6] While young Czechs move across districts to obtain a college education,[7] they are much less likely to move after graduation. Low migration within the Czech Republic has already been documented by Fidrmuc (2004). This author, however, did not distinguish education-specific migration. To fill this gap, I compare district-specific numbers of college graduates in two 5-year age cohorts (30-34 and 35-40) as recorded by the 1991 Census, with the same cohorts ten years later (i.e., 40-44 and 45-50 years-old).[8] This comparison, presented in Table 9 in the appendix as percentage changes over the 10-year period, suggests that cross-district migration of college graduates in the Czech Republic is very low. The total change in the number of college graduates in the Czech Republic for both age cohorts over the analyzed 10-year period reaches 12%. This growth could be mainly accounted for by people completing their education while working. Only a few districts experience percentage changes in the number of college graduates much different from the country average, which does not allow us to treat these districts as separate labor markets. To see how this fact influences the results, the final analysis is conducted with and without the high migration districts.

---

[6] Bound et. al (2004) show that the relationship between the production and stock of college graduates in U.S. states is weak, and thus state-specific educational policies might not have the desired effect on the labor market. This, however, appears not to be the case in the Czech Republic.

[7] There is also a significant number of students commuting to study.

[8] The districts of Prague (capital of the Czech Republic and the main city of Bohemia) and Brno (the main city of Moravia), the outliers in the number of college graduates, have been removed from this analysis.

Focusing on a country with significant district-level differences in the educational profile of the population driven by public policy decisions enables me to investigate how these decisions influence the situation of graduates in the labor market. It is especially interesting to see if, in those districts with higher skill endowment and/or where higher education is expanding more rapidly, it is easier or more difficult for college graduates to find employment that takes advantage of their skills. The preliminary analysis, presented in Figure 3, shows that indeed districts with a higher share of college graduates in their populations tend to offer more "college" workplaces. This analysis is of particular policy interest because it reveals whether in this setting the expansion of higher education can improve employment possibilities of college graduates (and thus their skill usage) by attracting advanced technologies.

4 Theoretical Framework of Workers' Allocation Across Occupations

In this paper I analyze the influence of variations in the relative number of college graduates in the population on their allocation between "college" and "noncollege" occupations. The first question to be answered before proceeding to the empirical analysis is why we would observe some college graduates working in "noncollege" occupations, and how to recognize which occupations are "college" and which are "noncollege". A model dealing with these issues has been proposed by Gottschalk and Hansen (2003). I modify it to directly model the influence of supply and demand conditions on the equilibrium allocation of college graduates. Later on, I also allow for endogenous influence of the number of college graduates in the labor market on their productivity in "college" occupations. This leads to an ambiguous prediction of the sign of the relationship between the relative number of college graduates in the population and their occupational allocation.

The model proposed by Gottschalk and Hansen (2003) assumes that there are two sectors in the economy: a "college" sector and a "noncollege" sector. Competitive firms in both sectors produce the same uniform good.[9] They have the following production functions:

$$Q_1 = F_1(\alpha_{C1}L_{C1} + \alpha_{N1}L_{N1}) \quad (1)$$

$$Q_2 = F_2(\alpha_{C2}L_{C2} + \alpha_{N2}L_{N2}), \quad (2)$$

where Q_j measures the output of sector j, L_{Cj} and L_{Nj} are the amounts of college- and high school-educated labor in sector j, α_{ij} are productivities of labor type i in sector j, and F_js are twice-differentiable functions with $F_j'(\cdot) > 0$ and $F_j''(\cdot) < 0$. It is assumed that in sector 1 college-educated labor is relatively more productive than high school-educated labor as compared to sector 2 ($\frac{\alpha_{C1}}{\alpha_{N1}} > \frac{\alpha_{C2}}{\alpha_{N2}}$). That is why sector 1 is called the “college” sector.

Firms’ profit maximization under the price of output normalized to unity and labor input prices being w_{C1}, w_{C2}, $w_{N1}$and w_{N2}, respectively, gives the following condition:

$$\frac{w_{C1}}{w_{N1}} = \frac{\alpha_{C1}}{\alpha_{N1}} > \frac{\alpha_{C2}}{\alpha_{N2}} = \frac{w_{C2}}{w_{N2}}, \quad (3)$$

i.e., the wages of college graduates relative to high school graduates are higher in sector 1, the “college” sector. This property will be further used to distinguish between “college” and “noncollege” occupations.

To complete the model, I modify the supply functions of different labor types to both sectors proposed by Gottschalk and Hansen (2003). Like these authors, I assume that workers in a pool of all college and high school graduates decide to work in either sector “*based on their heterogenous preferences and the relative wages available to them across sectors*” (p. 5). On top of that, however, I specify the relationship between the total number of college and high school graduates in the labor market and the sector-specific supply functions, which is not explicitly shown in the original model.[10] The authors do not need to model this because they do not

[9] Allowing the two sectors to produce different goods does not influence the inference of this model. This assumption is kept for the purpose of clarity.

[10] The supply functions of college and high school graduates to the “college” sector used by Gottschalk and Hansen (2003) are the following: $L_{C1}^S = \lambda_C + \beta_C \frac{w_{C1}}{w_{C2}}$ and $L_{N1}^S = \lambda_N + \beta_N \frac{w_{N1}}{w_{N2}}$. Note that they do not explicitly account for the total amount of college- and high school-educated labor in the economy.

analyze the relationship between the structure of the labor force and the allocation of workers across occupations. In my version of the model it is assumed that the total supply of a given labor type to a given sector is a proportion of all workers of this type in the population. This allows for direct analysis of the influence of changes in the structure of the labor force on the market equilibrium. The assumed supply functions are the following:

$$\frac{L_{C1}^S}{L_C} = \lambda_C + \beta_C \ln\left(\frac{w_{C1}}{w_{C2}}\right) \tag{4}$$

$$L_{C2}^S = L_C - L_{C1}^S \tag{5}$$

$$\frac{L_{N1}^S}{L_N} = \lambda_N + \beta_N \ln\left(\frac{w_{N1}}{w_{N2}}\right) \tag{6}$$

$$L_{N2}^S = L_N - L_{N1}^S, \tag{7}$$

where L_C and L_N are the total numbers of college and high school graduates in the labor market, and β_i and λ_i are the aggregate preference parameters of workers of type i.

Together, equations (3)[11] and (4) - (7) define the equilibrium allocation and wages of college and high school graduates among the two sectors. An important property of this model is that in equilibrium there are some college-educated workers employed in both sectors. This study concentrates on the fraction of college graduates working in the "noncollege" sector, which is defined as

$$\pi_C \equiv \frac{L_{C2}}{L_C}. \tag{8}$$

The main advantage of the proposed model is that it directly captures the influence of the supply conditions (the total amount of each labor type in the economy, L_i) and demand conditions (labor productivities, α_{ij}) on the equilibrium fraction of college graduates working in the "noncollege" sector (π_C^*).

$$\pi_C^* \equiv 1 - \frac{L_{C1}^*}{L_C} = f\left(L_C, L_N, \alpha_{C1}, \alpha_{N1}, \alpha_{C2}, \alpha_{N2}\right). \tag{9}$$

[11] Equation (3) actually consists of 4 equations: $w_{C1} = \alpha_{C1}F_1'(L_1)$, $w_{N1} = \alpha_{N1}F_1'(L_1)$, $w_{C2} = \alpha_{C2}F_2'(L_2)$, and $w_{N2} = \alpha_{N2}F_2'(L_2)$, where $L_1 = \alpha_{C1}L_{C1} + \alpha_{N1}L_{N1}$ is the total labor aggregate used in sector 1 and $L_2 = \alpha_{C2}L_{C2} + \alpha_{N2}L_{N2}$ is the total labor aggregate used in sector 2.

To understand the forces influencing the occupational allocation of college graduates, let me analyze how the equilibrium fraction of college graduates working in the "noncollege" sector reacts to the shifts in supply- and demand-characterizing variables, i.e., the structure of the labor market ($\frac{L_C}{L_N+L_C}$) and the extent of the skill bias of technology ($\frac{\alpha_{C1}}{\alpha_{N1}}$).

First, I analyze how the equilibrium allocation changes when the skill-biased technological change (SBTC) happens in the "college" sector, i.e., when $\frac{\alpha_{C1}}{\alpha_{N1}}$ grows and all other variables are kept unchanged. This change should increase wages offered by firms in the "college" sector to college graduates (demand for college graduates in sector 1 shifts up). Higher wages attract more college graduates to the "college" sector, as described by equation (4). This, in turn, lowers a bit their wages in sector 1 and increases their wages in sector 2. Finally, wages adjust in such a way that no more workers want to change jobs. The new equilibrium is characterized by higher wages for college graduates in both sectors, but wages in sector 1 increase more as compared to the initial level. This makes the new $\frac{w^*_{C1}}{w^*_{C2}}$ higher than the initial one and thus the new π^*_C lower than the initial one. To sum up,

$$\frac{\partial \pi^*_C}{\partial \left(\alpha_{C1}/\alpha_{N1}\right)} < 0. \tag{10}$$

Next, let me analyze what happens when the relative stock of college graduates in the labor market ($\frac{L_C}{L_N+L_C}$) increases, which is a result of growth in L_C and a related fall in L_N. This change results in an upward shift in the supply of college graduates and a downward shift in the supply of high school graduates to both sectors, as shown by equations (4) and (6). As a result, wages of all labor types in the "college" sector fall. In the "noncollege" sector wages fall as well, but less dramatically, as long as $\frac{\alpha_{C2}}{\alpha_{N2}} \geqslant 1$. If $\frac{\alpha_{C2}}{\alpha_{N2}} < 1$, wages in sector 2 may actually rise. In any case, the ratio $\frac{w_{C1}}{w_{C2}}$ falls and some workers reallocate from the "college" to the "noncollege" sector. This, in turn, lowers a bit wages in sector 2 and increases them in sector 1 (but not above the initial level) so that ultimately nobody wants to change jobs. The new equilibrium is characterized by lower wages for college graduates in both sectors, but wages in sector 1 decrease more as compared to the initial level. This makes the new

$\frac{w^*_{C1}}{w^*_{C2}}$ lower than the initial one and thus the new π^*_C higher than the initial one. To sum up,

$$\frac{\partial \pi^*_C}{\partial \left(\frac{L_C}{L_N+L_C}\right)} > 0. \tag{11}$$

The above analysis leads to the following formulation of the relationship between the relative supply of college graduates to the labor market and the fraction of them working in "noncollege" occupations:

$$\pi^*_C = f\left(\underset{+}{\frac{L_C}{L_N+L_C}}, \underset{-}{\frac{\alpha_{C1}}{\alpha_{N1}}}, \textit{other factors}\right). \tag{12}$$

Assuming that the relationship is approximately linear[12] and other factors vary randomly, it can be written it in the following form:

$$\pi^*_C = \gamma_0 + \gamma_1 \frac{L_C}{L_N+L_C} + \gamma_2 \frac{\alpha_{C1}}{\alpha_{N1}} + \varepsilon, \tag{13}$$

where $\gamma_1 > 0$ and $\gamma_2 < 0$, as derived.

According to the model presented above, the relationship between $\frac{L_C}{L_N+L_C}$ and π^*_C is positive. However, this model does not take into account the endogenous influence of the labor force structure on college graduates' productivity in "college" occupations. Let me now introduce endogeneity (also known as productivity spillover) into the model to show that it can alter the relationship. A general representation of productivity spillovers commonly used in the literature is in the form of productivity being an increasing function of aggregate skills (e.g., Acemoglu and Angrist, 2000; Moretti, 2004). In this paper I use a simple linear relationship:

$$\frac{\alpha_{C1}}{\alpha_{N1}} = \alpha + \delta \frac{L_C}{L_N+L_C}, \tag{14}$$

where $\delta \geq 0$ ($\delta = 0$ implies no spillovers and $\delta > 0$ implies the existence of positive productivity spillovers). Incorporating this into equation (13), I get:

$$\pi^*_C = \gamma_0 + \underbrace{(\gamma_1 + \gamma_2 \delta)}_{\theta} \frac{L_C}{L_N+L_C} + \gamma_2 \cdot \alpha + \varepsilon. \tag{15}$$

[12] The model outlined in this section has no closed form solution. Therefore, I have to approximate its functional form.

When allowing for productivity spillovers from a high concentration of skills, the sign of the relationship between the relative supply of college graduates and the fraction of them working in "noncollege" occupations is not clearly predicted by the model. If the direct effect (γ_1) is stronger than the spillover effect ($\gamma_2\delta$), the overall relationship is negative; however, if the spillover effect is strong enough to compensate for the direct effect, the overall relationship is positive. The goal of this paper is to estimate the parameter $\theta_1 \equiv \gamma_1 + \gamma_2\delta$ to determine whether positive or negative effects prevail in the influence of the relative stock of college graduates on their allocation across occupations.

Before proceeding to the empirical analysis, let me discuss the assumptions behind the model and the limitations implied by them. First of all, it is important to acknowledge that the above model describes a single closed economy. One should be careful when applying it to compare districts within one country if workers and firms are mobile. In the context of the Czech Republic, however, mobility of labor is limited. As shown in Section 2, workers tend to stay in the district where they graduated. Additionally, there are other factors than labor availability influencing firms' decisions to locate in a given district, and thus firm mobility does not fully compensate cross-district differences in the labor force structure. This allows me to treat districts as separate labor markets and use equation (15) to analyze the cross-district relationship between the relative supply of labor and the fraction of college graduates working in "noncollege" occupations.

Second, the assumption of workers' heterogeneous preferences towards job attributes could be questioned. While this is the only approach used in this line of literature, one could come up with alternative explanations for why we observe college graduates in both "college" and "noncollege" occupations. Workers might have heterogeneous ability to use college-gained skills, and "college" firms employ only those with high enough ability. Discussion of this model is not within the scope of this paper. Let me note, however, that also the alternative explanation supports the prediction of the model used to classify occupations, i.e., that relative wages of college to high school graduates are higher in the "college" sector (see the appendix).

I base the analysis on the Gottschalk and Hansen (2003) model to be consistent with the literature.

5 Estimation Strategy

The theoretical model derived in the previous section serves as a baseline for analyzing the relationship between the relative stock of college graduates and the fraction of them working in "noncollege" occupations. Before formulating an econometric model based on these derivations, let me note that equation (15) accommodates an implicit assumption that the aggregate preference of workers, summarized by parameters β_C, β_N and λ_C, λ_N, are constant within and across districts. This is, however, a very unrealistic assumption. It can be argued that the composition of characteristics of individuals living in a given district influences their allocation across occupations through their preference parameters. If, for example, in a given district there are many females with a college education (who are, on average, less flexible in looking for employment), there might be a higher fraction of college graduates in "noncollege" occupations there. In order to account for such effects, I formulate an econometric model on the individual rather than on the aggregate level, i.e., I model the propensity of an individual college graduate to work in a "noncollege" occupation as a function of her characteristics and characteristics of the region where she lives, as shown in equation (16). This model can be thought of as a disaggregated version of equation (15).

$$\text{Prob}(nocollege_{ikt}) = \gamma_0 + \mathbf{X}'_{ikt}\boldsymbol{\theta}_0 + \theta_1 \left(\frac{L_C}{L_N + L_C} \right)_{kt} + \mathbf{Y}'_{kt}\boldsymbol{\theta}_2 + \varepsilon_{ikt}, \tag{16}$$

where $\text{Prob}(nocollege_{ikt})$ is an indicator whether a college graduate i in district k at time t is working in a "noncollege" occupation, $\mathbf{X}'_{ikt}$ is a vector of individual characteristics such as the worker's potential labor market experience (in years) and gender, $\left(\frac{L_C}{L_N+L_C} \right)_{kt}$ is the relative stock of college graduates in district k at time t, $\mathbf{Y}'_{kt}$ is a vector of other year-district specific characteristics, and ε_{ikt} represents the individual, time, and district specific unobservable determinants of college graduates' allocation across occupations. The parameter of main interest is θ_1; it describes the causal relationship between the relative number of college graduates in a district's population and their fraction working in "noncollege" occupations.[13]

The district specific characteristics in $\mathbf{Y}_{kt}$ include size measures such as the density of the district's population, and the logarithm of the district's labor force to account for assortative matching effects. It is generally accepted that in larger markets, workers and firms find each other more easily (Wheeler, 2001) and thus we could observe a lower fraction of college graduates working in "noncollege" occupations in large labor markets. I also control for the share of employment in the public sector because the individual level data used for estimations covers only employees from the commercial sector, while the public sector usually employs many college graduates, which can influence the district's equilibrium share of the highly educated.[14]

The source of identification used to estimate θ_1 is the variation in the fraction of highly-educated adults within and across Czech district populations and the simultaneous variation in the proportion of college graduates working in "noncollege" occupations in these districts. Because of the two-level structure of the variables,[15] the precision of $\widehat{\theta}_1$ might be significantly downward-biased if estimating the model (16) by standard methods. Simple clustering would not improve the situation because of a limited number of clusters (districts). As Donald and Lang (2007) show, standard errors of estimated parameters on variables that are constant within a group (here within a district in a given year) *"are asymptotically normally distributed only as the number of groups goes to infinity"* (p. 221). The same authors propose a two-step procedure to overcome this problem. I follow this procedure by first estimating

[13] Ideally, the above should be modeled as a choice between three alternatives: working in the "college" sector, working in the "noncollege" sector, and being unemployed. Unfortunately, the data set used in this paper does not contain information about the unemployed. Nevertheless, this is not an important issue in the case of the Czech Republic, where the unemployment rate of college graduates did not exceed 4.6% in any district over the 2000-2006 period.

[14] I have also experimented with using real GDP per capita as an additional explanatory variable, but it appears to have no power in explaining the variation in the fraction of college graduates working in "noncollege" occupations.

[15] The dependent variable is at the individual level, while the explanatory variable of interest is at the group (district) level.

the propensity of individual college graduates to work in noncollege occupations as a function of their individual characteristics and district-time dummies. In the second step I perform a weighted least squares (WLS) regression of the estimated parameters by district-time dummies on district-time characteristics, where the variance of the estimated parameters by district-time dummies is used as the weighting factor. This approach can be summarized in the following way:

$$1^{\text{st}} \text{ step:} \qquad \text{Prob}(nocollege_{ikt}) = \delta_0 + \mathbf{X}'_{ikt}\boldsymbol{\delta}_1 + \mathbf{TD}'_{kt}\mathbf{d} + \boldsymbol{\xi}_{ikt}, \tag{17a}$$

$$2^{\text{nd}} \text{ step:} \qquad \widehat{d}_{kt} = \gamma_0 + \theta_1 \cdot \left(\frac{L_C}{L_N + L_C}\right)_{kt} + \mathbf{Y}'_{kt}\boldsymbol{\theta}_2 + \varepsilon_{kt}, \tag{17b}$$

where $\mathbf{TD}'_{kt}$ is a vector of year-district dummies, $\boldsymbol{\xi}_{ikt}$ captures unobservable individual characteristics, and ε_{kt} represents the time and/or district specific unobservable determinants of college graduates' allocation across occupations.

An omitted variable problem appears when estimating equation (17b) by WLS.[16] Some of the factors captured by the error term might bias the estimate of $\widehat{\theta}_1$ due to a correlation with the relative supply of college graduates. The major source of bias is the unobserved heterogeneity across districts, as well as over time, in the demand for labor. Both time and district specific productivity shocks might partially drive the variation in the stock of college graduates. For example, expansion of the hi-tech industry in one district may attract highly-educated workers to move there or the observation of country-wide SBTC could motivate more people to pursue higher education. This is why I expect $cov(\varepsilon_{kt}, \frac{L_C}{L_N+L_C}_{kt}) \neq 0$. The intuitive sign of this correlation is positive (i.e., positive productivity shocks induce a higher fraction of college graduates), thus the WLS estimates of the relationship from equation (17b) would be biased downwards.[17]

[16] An omitted variable bias might also appear when estimating equation (17a) if workers sort into cities according to their unobservable abilities. In this case, $\mathbf{TD}'_{kt}$ and $\boldsymbol{\epsilon}_{ikt}$ are correlated, which influences the estimate of d_{kt}. This could be addressed by controlling for workers' fixed effects. The data used in this study do have a repeated cross-section structure, which does not allow for this approach. Nevertheless, Moretti (2004) shows that omitted "*individual characteristics are not a major source of bias*" (p. 176).

[17] A positive demand shock in the "college" sector makes more graduates work there and thus decreases π^*_{Ckt}. At the same time, it triggers growth in $CollSh_{kt}$. What we observe is a growth in the relative supply of college graduates and a decline in the fraction of

Endogeneity of the fraction of the population with a college degree can be overcome in several ways. The first proposal is to use an instrument that predicts well the share of college graduates in a district's population but at the same time is uncorrelated with district-specific productivity shocks. In the search for an instrumental variable I draw from Moretti's (2004) approach towards estimating the social returns to education. He proposes that the historical presence of a college be used as an instrument for the relative supply of college graduates. Another proposal is to work with a panel of districts and use a fixed-effect estimation to difference out district-specific unobservable factors.

Moretti's (2004) idea to use the historical presence of a college as an exogenous predictor of the variation in the stock of highly-educated labor across districts can also be applied in the case of the Czech Republic (e.g., Jurajda, 2004). Because of limited cross-district labor mobility, as discussed in Section 2, the number of college graduates in the district population is to a large extent driven by the presence of a college in this district. Additionally, the majority of public colleges in the Czech Republic were established during communism, which makes their presence exogenous to current productivity shocks. Thus, the presence and/or size of a college[18] in a district as of the end of communism might be a good candidate for an instrument predicting the current stock of college graduates across districts. Although some colleges opened in the 1950s and 1960s were tied to local industries, which casts some doubt on the exogeneity of such instrumental variables, the industrial structure of districts changed during the period of transition[19] and the overall demand for labor has dropped during that time. That is why, while controlling for districts' industrial structure at the end of communism, I can safely use the chosen instruments.[20]

---

them employed in "noncollege" occupations, which creates the impression of a negative relationship between these two.

[18] Size of the district's college as of the end of communism is defined as the fraction of the district population holding a college degree in 1991.

[19] See Figure 4 in the appendix for a comparison of districts' industrial structure.

[20] Both presence of a college and size of a college in a district as of the end of communism are strong instruments (correlation with 2001 share of college graduates is 0.63 and 0.85, respectively). Additionally, Sargen's test of overidentifying restrictions suggests that,

The size and presence of a college in a district as of the end of communism can be used as instruments only in the case of cross-sectional analysis because these instruments do not vary over time. When applying the instrumental variable approach, I am left with a variation in the relative number of college graduates across districts that is due solely to the historical distribution of colleges and thus is uncorrelated with current district-specific productivity shocks. This should allow for identification of the unbiased cross-district relationship between the relative stock of college graduates and the fraction of them working in "noncollege" occupations.

Working with a panel of districts allows for identification of the influence of changes in the relative supply of college graduates on their allocation between "college" and "noncollege" occupations. It also allows me to use a fixed-effect estimation approach and difference out the time-constant district-specific demand shifters. In this way I eliminate the endogenous effect coming from the correlation of district-specific time-constant unobservables and the relative stock of college graduates in a district's population. Nevertheless, there still can be time-varying factors influencing the changes in the relative number of college graduates. Inclusion of a proxy for time-district specific demand factors – the Katz and Murphy (1992) demand shift index[21] – would remove some of the unobservable demand from the error term and minimize the bias of $\widehat{\theta}_1$.

6 Identifying "College" and "Noncollege" Occupations

In order to perform the estimations described above, I need to measure the fraction of college graduates employed in "noncollege" occupations. Thus, I need to classify all occupations of college graduates into "college" and "noncollege" ones. In doing so I follow Gottschalk and Hansen's (2003) approach based on the model presented in

given the presence of a college in 1991 is exogenous to the model, its size is exogenous as well (p-value = 0.512).

[21] Further details about the Katz and Murphy demand shift index can be found in Katz and Murphy (1992) and Moretti (2004).

Section 4. This approach exploits the property of the model described by inequality (3), i.e., that wages of college graduates relative to high school graduates are higher in sector 1, the "college" sector. This can be further extended to the situation in which there are many different occupations in each sector, but still it holds that in each "college" occupation, the relative productivity of college graduates is higher than in each "noncollege" occupation. Consequently, the relative wages of college graduates are also higher in occupations from the "college" sector than from the "noncollege" sector.

Based on this model, I can distinguish between "college" and "noncollege" occupations once knowing the wage premium paid to college-educated workers over high school-educated workers in each occupation employing both worker types. Gottschalk and Hansen, who perform an occupational classification for the U.S., use a 10% college wage premium as a threshold, i.e., they classify an occupation as "college" when it pays at least a 10% premium to highly-educated workers.[22] This value, as they justify it, is a bit higher than the lowest estimate of the overall college wage premium in the U.S. as estimated by Katz and Murphy (1992). Taking into account that the overall college wage premium in the Czech Republic is significantly higher than in the U.S., I also experiment with a higher threshold (15%). Nevertheless, as presented in Section 7.4, the qualitative results are insensitive to the chosen threshold.

Occupations in which one type of worker strongly prevails are classified automatically. Gottschalk and Hansen call occupations in which more than 90% of workers have a higher education as "college" ones. Due to the low fraction of college graduates in the Czech labor market, I also experiment with a 85% threshold. As in the case of college wage premium, final results are insensitive to the chosen threshold. Additionally, I classify occupations where more than 95% of workers have only a high school diploma as "noncollege" occupations.

The procedure for classifying occupations can be described as follows. For each 3-digit occupation where college graduates constitute between 5% and 90% of all

[22] The same threshold is used by Cardoso (2007) for analyzing the Portugese situation and by Grazier et al. (2008) for analyzing the British labor market.

employees, I estimate the following wage equation:

$$\log w_{ik} = \beta_{0k} + \beta_{1k} \cdot exp_i + \beta_{2k} \cdot exp_i^2 + \beta_{3k} \cdot female_i + \phi_k \cdot coll_i + \varepsilon_{ik}, \tag{18}$$

where $\log w_{ik}$ is the logarithm of hourly wage received by worker i in occupation k, exp_i and exp_i^2 are each worker's potential labor market experience (in years) and its square, $female_i$ is a dummy variable indicating a worker's gender and $coll_i$ is a dummy variable equal to 1 if a worker has a college degree and 0 otherwise.[23] This is a standard Mincerian regression used widely in the literature for identifying returns to different worker characteristics. The parameter used to classify occupations is ϕ_k, the college wage premium. Occupations for which the hypothesis that $\widehat{\phi_k} > threshold$ (where $threshold$ is initially set at 0.10) cannot be rejected at usual confidence levels are classified as "college" ones. Those for which this hypothesis is rejected are classified as "noncollege". Finally, occupations where more than 90% of employees are college graduates are classified as "college" occupations and those where less than 5% of employees are college graduates are classified as "noncollege" occupations.

7 Estimation of the Influence of College Supply on Allocation of College Graduates Across Occupations

7.1 Data Description

For the purpose of the empirical analysis I use the Czech national employer survey, ISPV. This is a linked employee-employer dataset (LEED) gathered and processed according to the requirements of the Czech Ministry of Labor and the European Union. Information is collected from a sample of more than 3500 firms in the private sector, which report wages and other information for about 1.3 million workers (about

[23]The sample used to classify occupations contains all college and high school educated workers not older than 35. The sample choice is discussed in more detail in the next section.

a third of the whole employment). This dataset is a repeated cross-section; the data is collected at the firm level and individual workers are not explicitly followed.

The main advantage of the dataset is its size. In order to apply the Gottschalk and Hansen (2003) methodology of classifying occupations, it is necessary to have no fewer than 100 observations of workers with high school or higher level of education in each occupation. In the ISPV dataset there are about 35,000 young college graduates, defined as individuals with at least a bachelor degree, below 35 years of age, and 65,000 young high school graduates, defined as individuals below 35 years of age[24] who have passed a maturity exam, for each of the years in the 2000 – 2008 period. This is enough to carry out the analysis at the level of 3-digit occupations.

The variables reported in the dataset include age, gender, and education level of each employee. Moreover, one can find the characteristics of the firm (location, industry, size, ownership structure, etc.) and occupation in which an individual is employed, and her monthly earnings together with the number of hours worked. The last two variables allow me to calculate the hourly wage, which is defined as the average pay per hour during the first quarter of a year.

Occupations are coded in the ISPV dataset according to a local system which follows the International Standard Classification of Occupations (ISCO). For the purpose of this study, I use occupations defined at the 3-digit level. This is the precision also used by Gottschalk and Hansen (2003). Occupations defined by 3-digit codes are detailed enough to capture quite narrowly-defined jobs and are at the same time wide enough to include the number of workers, allowing me to perform the estimations. Nevertheless, some occupations had to be merged in order to achieve a larger sample size, in which case the aggregation was kept the same for each year of the analysis.

District- and region-specific data on population and labor force structure are taken from the Czech Labor Force Survey (LFS). This survey is representative at the

[24] Card and Lemieux (2001) show that younger and older workers are not perfect substitutes. I work just with young workers to avoid this issue.

regional (NUTS-3) level. To get district-level information, 1991 and 2001 Census data are used. 2001 values are extrapolated to other years of the analysis using region-specific growth rates calculated from the LFS. Additionally, the district information on registered unemployment gathered by the Czech Ministry of Labor is used to calculate gender- and employment-specific unemployment rates.

7.2 Cross-sectional Estimation at the District Level

This section presents the second-stage estimates of the relationship between the relative number of college graduates in the population and the fraction working in "noncollege" occupations, as described by equation (17b), in the cross-district dimension. As shown in Table 2, this analysis supplies some evidence that the productivity spillover from a high concentration of skills is strong enough to create improved employment possibilities for college graduates in districts where their stock is relatively high. The table reports the estimates of θ_1 obtained using different models (OLS and IV) and different sets of districts. Prague and Brno, the two major cities of the Czech Republic, are eliminated from the estimation because they have an incomparably large share of college graduates in the local population and a high concentration of businesses. Additionally, I remove districts characterized by a high migration of college-educated citizens, as discussed in Section 3.

Table 2 indicates that the estimates of the influence of the relative number of college graduates in a district population on the fraction of them working in "noncollege" occupations are significantly negative when the OLS estimation method is applied. These results are, however, biased downwards due to the simultaneity in the determination of these two variables. Thus, we should expect the true relationship not to be that negative. Indeed, when instrumenting the 2001 share of college graduates in the district population with the same measure as of the end of communism, estimates closer to zero are obtained. The relationship between the relative stock of college graduates in the district population and the fraction of them working in "noncollege" occupations is estimated to be different from zero with only 85% confidence. Nevertheless, it is not estimated to be positive, which would be the expected result when no spillover effects are present.[25] Actually, the economic significance of the coefficient by $CollShare$ is quite strong – a one percentage point increase in the share of college graduates in the local labor market is estimated to cause a 0.9 percentage point decrease in the fraction of college graduates working in "noncollege" occupations. This gives us some evidence to support the hypothesis that a larger number of college graduates attracts advanced technologies and in this way improves the situation of highly-educated workers in the district labor market.

Table 2: Determinants of the share of college graduates in "noncollege" occupations across Czech districts in 2001

	(1) OLS	(2) OLS	(3) OLS	(4) IV	(5) IV	(6) IV
$CollShare$	-1.241**	-1.250**	-1.241**	-0.890	-0.897	-0.908
(p-value)	(0.030)	(0.028)	(0.030)	(0.150)	(0.146)	(0.142)
Prague & Brno	Yes	No	No	Yes	No	No
High migration	Yes	Yes	No	Yes	Yes	No
Observations (distr. cells)	71	69	67	71	69	67

Notes: The dependent variable is individual young college graduate's probability of working in a "noncollege" occupation (defined as paying a college premium higher than 10%). *CollShare* is the 2001 share of college graduates in a respective district's young population; as an IV for this variable, I use the share of college graduates in the district population as of the end of communism (1991). Young workers are defined as being younger than 35. Columns (1) - (3) report OLS estimation results, while columns (4) - (6) report IV estimation results. P-values are in parentheses.

It needs to be stressed that the effect identified in this section comes solely from different allocation of college graduates across the same set of "college" and "noncollege" occupations and not from different classification of occupations across districts. This is because the classification of occupations is defined on the national level. To investigate whether the presence of many skilled workers triggers changes in production technologies within some occupations, one should classify occupations into the two groups separately for each district, which is not possible to do in this analysis

[25] Recall that, according to equation (15), $\delta_2 > 0$. Thus a non-positive estimate of $\theta_1 = \delta_1 + \delta_2$ implies that $\delta_1 < 0$, i.e. that the spillover effect exists.

due to data limitations. Nevertheless, a similar effect is analyzed in the next section, where the classification of occupations varies from year to year.

7.3 Estimation on the Panel of Districts

To complete the picture, estimates of the relationship between the relative number of college graduates in the population and their fraction working in "noncollege" occupations in cross- and within-district dimension should be examined. Table 3 presents OLS and fixed-effect (FE) estimates of θ_1 obtained using different sets of districts. As in the case of cross-district analysis, separate analyses were performed excluding Prague, Brno, as well as high migration districts.

Table 3: Determinants of the share of college graduates in "noncollege" occupations in Czech districts over the 2000-2008 period

	(1) OLS	(2) OLS	(3) OLS	(4) FE	(5) FE	(6) FE
$CollShare$	0.170**	0.191***	0.214***	0.211**	0.270**	0.272**
(p-value)	(0.014)	(0.010)	(0.004)	(0.029)	(0.025)	(0.028)
Prague& Brno	Yes	No	No	Yes	No	No
High migration	Yes	Yes	No	Yes	Yes	No
Observations (distr.-year cells)	639	621	603	639	621	603

Notes: The dependent variable is individual young college graduate's probability of working in a "noncollege" occupation (defined as paying a college premium higher than 10%). *CollShare* is the year-specific share of college graduates in a respective district's young population. Young workers are defined as being younger than 35. Columns (1) - (3) report OLS estimation results, while columns (4) - (6) report fixed-effect estimation results. P-values are in parentheses.

In the over-time dimension, estimates of the relationship between the share of college graduates in the district population and the fraction of them working in "noncollege" occupations are positive even under OLS. The fixed-effect estimates are even higher, as expected. This suggests that the supply effect is stronger than the spillover effect[26] and that an increase in the relative stock of college graduates in the local labor market worsens their employment situation.

[26] Movement along a downward sloping demand curve is larger in scale than the shift of this curve.

The contrasting results of cross-sectional and over-time analysis might be interpreted in the following way. Districts with a historically determined higher supply of college graduates have attracted skill-complementing capital and offer more employment possibilities in "college" occupations. Thus, the situation of college graduates is better in these regions. Nevertheless, by stimulating an increase in the stock of college graduates from year to year, districts are not able to attract enough capital to compensate for the supply effect, and thus over time we observe a positive relationship between the share of college graduates in a district population and the fraction of them working in "noncollege" occupations. These could be thought of as long-run and short-run effects. Positive spillovers from a high concentration of college graduates are found to be significant only in the long-run context.

Additional insight is provided by repeating the above analysis with the classification of occupations held constant for each year, which captures changes in the fraction of college graduates working in "noncollege" occupations due to reallocation within the same set of occupations. This exercise results in significantly higher estimates of the relationship between the share of college graduates in the district population and the fraction of them working in "noncollege" occupations. Thus, we can conclude that reclassification of occupations plays an important role in determining the fraction of skilled workers working in "noncollege" occupations.

7.4 Robustness Check

It could be argued that the results presented above are specific to the definition of "college" occupations. Recall that an occupation is defined to be "college" when the wage premium it pays to college graduates exceeds 10% or when the proportion of college graduates working there exceeds 90%. These thresholds have been chosen specifically to reflect the conditions of the Czech economy. To show that the results are not driven by the chosen thresholds, I present the outcomes of analogous estimations performed using an alternative definition of a "college" occupation, i.e., with the wage premium threshold set at 15% and the proportion threshold at 85%. These are the values used in previous research to distinguish between "college" and “noncollege” occupations. As seen in Tables 4 - 5, the use of an alternative definition leads to qualitatively the same results.

Table 4: Determinants of the share of college graduates in “noncollege” occupations across Czech districts in 2001

	(1) OLS	(2) OLS	(3) OLS	(4) IV	(5) IV	(6) IV
$CollShare$	-0.985*	-0.994*	-0.990*	-0.737	-0.746	-0.753
(p-value)	(0.065)	(0.065)	(0.069)	(0.203)	(0.201)	(0.202)
Prague&Brno	Yes	No	No	Yes	No	No
High migration	Yes	Yes	No	Yes	Yes	No
Observations (distr. cells)	71	69	67	71	69	67

Notes: The dependent variable is individual young college graduate’s probability of working in a “noncollege” occupation (defined as paying a college premium higher than 15%). *CollShare* is the 2001 share of college graduates in a respective district’s young population; as an IV for this variable, I use the share of college graduates in the district population as of the end of communism (1991). Young workers are defined as being younger than 35. Columns (1) - (3) report OLS estimation results, while columns (4) - (6) report IV estimation results. P-values are in parentheses.

Additionally, I check whether the noisy character of district-level data does not influence the results of panel estimations. As explained in Section 7.1, district-level data for non-census years are derived from the Czech Labor Force Survey (LFS) which is not representative at the district level. Thus, I repeat the panel estimation on the regional level (a region aggregates 5 districts, on average), for which data derived from the LFS is more reliable. The relevant estimates are presented in Table 6. They are qualitatively the same as district-level regressions.

Other robustness checks involved including different forms of $\frac{L_C}{L_N}$ in the regressions and repeating the analysis on a panel of firms subsample. Neither of these brought additional insight to the analysis.

Table 5: Determinants of the share of college graduates in "noncollege" occupations in Czech districts over the 2000-2008 period

	(1) OLS	(2) OLS	(3) OLS	(4) FE	(5) FE	(6) FE
CollShare	0.149**	0.179***	0.201***	0.148*	0.231**	0.234**
(p-value)	(0.017)	(0.006)	(0.003)	(0.096)	(0.037)	(0.038)
Prague&Brno	Yes	No	No	Yes	No	No
High migration	Yes	Yes	No	Yes	Yes	No
Observations (distr.-year cells)	639	621	603	639	621	603

Notes: The dependent variable is individual young college graduate's probability of working in a "noncollege" occupation (defined as paying a college premium higher than 15%). *CollShare* is the year-specific share of college graduates in a respective district's young population. Young workers are defined as being younger than 35. Columns (1) - (3) report OLS estimation results, while columns (4) - (6) report fixed-effect estimation results. P-values are in parentheses.

8 Conclusion

In this study I argue that the fraction of college graduates employed in "noncollege" occupations offers a useful measure for investigating forces shaping the labor market. Analysis of the evolution of this measure over time in the U.S. (Gottschalk and Hansen, 2003), Portugal (Cardoso, 2007), the U.K. (Grazier et al., 2008) and the Czech Republic (this study) reveals a consistent pattern. In every country the fraction of college graduates employed in "noncollege" occupations has been decreasing over time despite a significant growth in the relative number of college-educated workers in the labor market. This phenomenon could be driven by two forces: (1) exogenous technological shocks simultaneously triggering shifts in the demand for and supply of college graduates, or (2) a higher number of college graduates attracting advanced technologies and thus endogenously shifting the demand for skilled workers.

These forces are not mutually exclusive; most probably they act simultaneously. Nevertheless, from the policy point of view it is important to know how strong the endogenous effect is as compared to the exogenous effect. In the absence of the endogenous effect, college enrolments should reflect the trend in technological progress of the economy; while the existence of this effect implies that increasing the educa-

Table 6: Determinants of the share of college graduates in "noncollege" occupations in Czech regions over the 2000-2008 period

	(1) OLS	(2) OLS	(4) FE	(5) FE
CollShare	0.050	0.024	0.178***	0.239***
(p-value)	(0.524)	(0.766)	(0.081)	(0.043)
Prague	Yes	No	Yes	No
Observations (reg.-year cells)	112	104	112	104

Notes: The dependent variable is individual young college graduate's probability of working in a "noncollege" occupation (defined as paying a college premium higher than 10%). *CollShare* is the year-specific share of college graduates in a respective region's young population. Young workers are defined as being younger than 35. Columns (1) - (2) report OLS estimation results, while columns (3) - (4) report fixed-effect estimation results. P-values are in parentheses.

tional attainment of the local population could be used as a tool to attract advanced technologies and increase the skill bias of the economy.

Results presented in this paper confirm the presence of a negative influence of the number of skilled workers on the fraction of them working in "noncollege" occupations across NUTS-4 districts of the Czech Republic. This is in line with the findings of Acemoglu (2003), who shows that a high supply of skilled labor shifts the skill bias of the local economy. On the other hand, in the within-district setup the relationship between the number of skilled workers and the fraction of them working in "noncollege" occupations is found to be positive. This could be caused by market frictions which delay the reaction of firms to the observed high concentration of skilled labor. Altogether, the findings of this paper suggest that in the long run, districts should be able to positively stimulate their labor markets by providing higher education to a larger fraction of their population (explanation 2). Nevertheless, in the short run the supply of college seats should be a response to the observed level of demand for skills (explanation 1).

Two challenges for future research follow. First, this study documents a positive relationship between the relative number of college graduates and their situation in the labor market, while Jurajda (2004) finds no influence of the concentration of college graduates in local labor markets on their wages. This implies that the Czech labor market reacts to an increased supply of skilled labor by offering more workplaces for college graduates and keeping their wage constant, on average. This observation could be used in further research to discriminate between alternative models of labor allocation between "college" and "noncollege" occupations, as proposed in the appendix. Second, while the presented analysis sheds some light on the within-country patterns observed in Figure 1, the cross-country differences remain unexplained. Understanding these differences would require a measure of college skills usage that is comparable across countries, development of which could be a topic for further research.

References

[1] Acemoglu D. (2002) Technical change, inequality, and the labor market, Journal of Economic Literature, 40(1): 7-72.

[2] Acemoglu D. (2003) Patterns of skill premia, The Review of Economic Studies, 70(243):199-230.

[3] Acemoglu D., Angrist J. (2000) How large are human capital externalities? Evidence from compulsory schooling laws, NBER Macroeconomics Annual, MIT Press: Boston.

[4] Bauer T. (2002) Educational mismatch and wages: a panel analysis, Economics of Education Review, 21(3):221-229.

[5] Bound J., Green J., Kezdi G., Turner S. (2004) Trade in university training: cross-state variation in the production and stock of college-educated labor, Journal of Econometrics, 121(1-2): 143-173.

[6] Bound J., Johnson G. (1992) Changes in the structure of wages in the 1980's: An evaluation of alternative explanations, American Economic Review, 82: 371-92.

[7] Brunello G., Comi S., Lucifora C. (2000) The college wage gap in 10 European countries: Evidence from two cohorts, IZA Discussion Paper No. 228.

[8] Card D., Lemieux T. (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.

[9] Cardoso A.R. (2007) Jobs for young university graduates, Economic Letters, 94(2): 271-277.

[10] CSO, Czech Statistical Office (2009)
http://www.czso.cz/eng/redakce.nsf/i/education_lide

[11] CSO, Czech Statistical Office (2010), Historie a vyvoj vysokeho skolstvi (History and development of higher education)
http://www.czso.cz/csu/2010edicniplan.nsf/p/3314-10

[12] Donald S G., Lang K. (2007) Inference with difference in differences and other panel data,The Review of Economics and Statistics, 89(2): 221–233.

[13] Eurostat (2009)
http://nui.epp.eurostat.ec.europa.eu/nui/show.do?dataset=lfsa_pgaed&lang=en

[14] Fidrmuc J. (2004) Migration and regional adjustment to asymmetric shocks in transition economies, Journal of Comparative Economics, 32: 230–247.

[15] Filer R., Jurajda S., Planovsky J. (1999) Education and wages in the Czech and Slovak Republics during transition, Labour Economics, 6(4): 581-593.

[16] Flabbi L., Paternostro S., Tiongson E.R.(2008) Returns to education in the economic transition: A systematic assessment using comparable data, Economics of Education Review, 27(6): 724-740.

[17] Fortin N.M. (2006) Higher education policies and the college wage premium: Cross-state evidence from the 1990's, American Economic Review, 96(4): 959-987.

[18] Grazier S., O'Leary N., Sloane P. (2008) Graduate employment in the UK: An application of the Gottschalk-Hansen Model, IZA Discussion Paper No. 3618.

[19] Gottschalk P., Hansen M. (2003) Is the proportion of college workers in noncollege jobs increasing? Journal of Labor Economics, 21(2): 449-471.

[20] Jurajda S. (2004) Are there increasing returns to local concentration of skills? Evidence on wages and returns to education in transition, mimeo

[21] Jurajda S. (2005) Czech returns to schooling: does the short supply of college education bite? Czech Journal of Economics and Finance (Finance a úvěr), 55(1-2): 83-95.

[22] Jurajda S., Terrell K. (2009) Regional unemployment and human capital in transition countries. Economics of Transition, 17(2): 241-274.

[23] Katz L., Murphy K. (1992) Changes in relative wages, 1963-1987: Supply and demand factors, The Quarterly Journal of Economics, 107(1): 35-78.

[24] McGuinness S. (2006) Overeducation in the labor market, Journal of Economic Surveys, 20(3): 387-418.

[25] McGuinness S., Bennett J. (2007) Overeducation in the graduate labour market: A quantile regression approach, Economics of Education Review, 26(5): 521-531.

[26] Moretti E. (2004) Estimating the social return to higher education: Evidence from longitudinal and repeated cross-sectional data, Journal of Econometrics, 121(1-2): 175-212.

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

[28] OECD (2006) Thematic Review of Tertiary Education - Country Background Report for Czech Republic, http://www.oecd.org/dataoecd/19/18/36443831.pdf

[29] OECD (2009) Education at a Glance 2008: OECD Indicators, www.oecd.org/edu/eag2008

[30] Pryor F., Schaffer D. (1997) Wages and the university educated: a paradox resolved, Monthly Labor Review, 120(7): 3-14.

[31] Walker I., Zhu Y. (2005) The College Wage Premium, Overeducation, and the Expansion of Higher Education in the UK, IZA Discussion Paper No. 1627.

[32] Wheeler C.H. (2001) Search, sorting, and urban agglomeration, Journal of Labor Economics,19(4): 879-899.

Appendix

A1. Derivations

The baseline model

The model proposed by Gottschalk and Hansen (2003) and adopted for the purpose of this study assumes that firms in "college" and "noncollege" sectors have the following production functions:

$$Q_1 = F_1(\alpha_{C1}L_{C1} + \alpha_{N1}L_{N1}) \tag{A1}$$

$$Q_2 = F_2(\alpha_{C2}L_{C2} + \alpha_{N2}L_{N2}), \tag{A2}$$

and workers allocate themselves across these sectors "*based on their heterogenous preferences and the relative wages available to them across sectors.*" (p. 5) What makes sector 1 a "college" sector is the relative productivity of college to high school graduates ($\frac{\alpha_{C1}}{\alpha_{N1}}$) which is higher than in sector 2 ($\frac{\alpha_{C2}}{\alpha_{N2}}$).

While Gottschalk and Hansen (2003) introduce linear aggregate supply functions of college and high school graduates to the "college" sector, I propose log-linear supply functions. These can be derived from workers' preferences, as described below.

Assume that workers of both types, subscribed by $k = C, N$, are endowed with one unit of labor which could be supplied to the "college" or "noncollege" sector for a respective wage. They derive utility from consumption of the uniform good produced by either sector and from working in the preferred sector. To account for heterogeneous preferences, the difference in utility between working in the "college" and "noncollege" sector (denoted as ε_i) is distributed across workers with education level k according to a probability distribution function $G_k(\cdot)$ defined for both positive and negative values. Individuals with positive values of ε_i prefer to work in the technologically advanced sector (e.g., they value social status and challenges of the job performed), while individuals with negative values of ε_i prefer to work in the simple sector (e.g., they value easiness of the job performed).

The workers' utility maximization problem looks as follows:

$$\max_j U(X_i, j_i) = U(X_i) + \varepsilon_i \cdot I(j = 1) \tag{19}$$
$$\text{s.t. } X_i = w_{kj},$$

where X_i denotes consumption of an individual worker, j_i is the sector where she works ($j_i = 1, 2$), $U(\cdot)$ is a twice-differentiable utility function with $U('\cdot) > 0$ and $U''(\cdot) < 0$, and $I(j = 1)$ is the indicator function equal to one for workers employed in the "college" sector.

Under this setup workers choose to work in the simple sector if $U(w_{k1}) + \varepsilon_i < U(w_{k2})$. So the total supply of workers to the "college" sector can be expressed as

$$L_{k1} = G_k[U(w_{k1}) - U(w_{k2})] \cdot L_k. \tag{20}$$

Assuming a logarithmic utility function, $U(x) = \ln(x)$, which is the most often used utility function in static models with one type of good, this translates to

$$\frac{L_{k1}}{L_k} = G_k\left[\ln\frac{w_{k1}}{w_{k2}}\right]. \tag{21}$$

$G(\cdot)$, being a cumulative distribution function, is increasing in its argument $\ln \frac{w_{k1}}{w_{k2}}$. Within its limits, it could be approximated by a linear function

$$\frac{L_{k1}}{L_k} = \lambda_k + \beta_k \ln\left(\frac{w_{k1}}{w_{k2}}\right),$$

where λ_k and β_k depend on the shape of G_k.

Workers with Heterogeneous Ability

An alternative to this approach is to assume that workers have heterogeneous ability to use college-gained skills. For simplicity, let me assume that this does not affect the "noncollege" sector, which continues to produce the uniform good according to the production function specified in equation (A1). In the "college" sector, highly educated workers have heterogeneous productivity, which for an individual i could be expressed as $\alpha_{C1} + \varepsilon_i$, where α_{C1} is the sector-specific productivity (given by the technology used there) and ε_i is individual-specific ability to use that technology. An individual's ability is drawn from a distribution $G(\varepsilon)$. It could enhance or downturn the sector-specific productivity. To summarize, the production function in "college" sector looks as follows:

$$Q_1 = F_1\left(\int_{\varepsilon=\varepsilon*}^{\infty} (\alpha_{C1} + \varepsilon)\, dG(\varepsilon) + \alpha_{N1} L_{N1}\right), \tag{A3}$$

where $\varepsilon*$ is the ability to use college-gained skills of the last college graduate employed in the "college" sector. This value is a characteristic of equilibrium in the labor market. L_{Cj} and L_{Nj} are, as before, the amounts of college- and high school-educated labor in sector j, while α_{ij} are productivities of labor type i in sector j. Under these conditions, profit maximizing-firms selling their output at price normalized to 1, pay workers their marginal products expressed as follows:

$$w_{C1}(\varepsilon_i) = (\alpha_{C1} + \varepsilon_i)\frac{\partial F_1}{\partial L_1}, \quad w_{N1} = \alpha_{N1}\frac{\partial F_1}{\partial L_1} \tag{A4}$$

$$w_{C2} = \alpha_{C2}\frac{\partial F_2}{\partial L_2}, \quad \text{and} \quad w_{N2} = \alpha_{N2}\frac{\partial F_2}{\partial L_2}, \tag{A5}$$

where $L_1 = \int_{\varepsilon=\varepsilon*}^{\infty} (\alpha_{C1} + \varepsilon)\, dG(\varepsilon) + \alpha_{N1} L_{N1}$ is the total labor aggregate used in sector 1 and $L_2 = \alpha_{C2} L_{C2} + \alpha_{N2} L_{N2}$ is the total labor aggregate used in sector 2.

The supply-side of this model looks in the following way. Workers choose to work in the sector which pays them a higher wage. As all high school graduates are identical, they all should be paid the same wage no matter which sector they work in or we would observe college workers exclusively in one sector. All college graduates are paid an equal wage in the "noncollege" sector and a wage reflecting their individual ability in the "college" sector. The last (marginal) college graduate employed in the "college" sector gets the same wage in either sector. This leads to the specification of the following conditions:

$$\begin{aligned}
L^*_{C1} &= \int\limits_{\varepsilon=\varepsilon*}^{\infty} dG(\varepsilon), \text{ with } w_{C1}(\varepsilon^*) = w_{C2} \\
L_{C2} &= L_C - L_{C1} \\
L_{N1} &= \begin{cases} L_N & \text{if } w_{N1}(L_{N1}, L_{C1}) > w_{N2}(L_{N2}, L_{C2}) \\ L_{N1} & \text{if } w_{N1}(L_{N1}, L_{C1}) = w_{N2}(L_{N2}, L_{C2}) \\ 0 & \text{if } w_{N1}(L_{N1}, L_{C1}) < w_{N2}(L_{N2}, L_{C2}) \end{cases} \\
L_{N2} &= L_N - L_{N1},
\end{aligned}$$

where $w_{N1}(L_{N1}, L_{C1})$ and $w_{N2}(L_{N2}, L_{C2})$ are given by equations (A4) and (A5), respectively, and $\varepsilon*$ is given by:

$$\begin{aligned}
(\alpha_{C1} + \varepsilon*)\frac{\partial F_1}{\partial L_1} &= w_{C1}(\varepsilon^*) = w_{C2} = \alpha_{C2}\frac{\partial F_2}{\partial L_2} \\
\frac{(\alpha_{C1} + \varepsilon*)}{\alpha_{C2}} &= \underbrace{\frac{\partial F_2}{\partial L_2} / \frac{\partial F_1}{\partial L_1} = \frac{\alpha_{N1}}{\alpha_{N2}}}_{\text{if } w_{N1}(L_{N1}, L_{C1}) = w_{N2}(L_{N2}, L_{C2})} \\
\varepsilon* &= \frac{\alpha_{C2}}{\alpha_{N2}}\alpha_{N1} - \alpha_{C1}
\end{aligned}$$

The fraction of college graduates employed in "noncollege" occupations defined as $\pi_C \equiv \frac{L_{C2}}{L_C}$ in equilibrium is equal to

$$\pi^*_C \equiv 1 - \frac{L^*_{C1}}{L_C} = 1 - \frac{\int_{\varepsilon=\varepsilon*}^{\infty} dG(\varepsilon)}{L_C} = f(\varepsilon*, L_C) = f(\underset{+}{L_C}, \underset{-}{\alpha_{C1}}, \underset{+}{\alpha_{N1}}, \underset{+}{\alpha_{C2}}, \underset{-}{\alpha_{N2}}). \tag{A6}$$

Similarly like in the baseline model, here we also observe that the average wage of college graduates is higher in the "college" sector than in the "noncollege" sector, which is shown below:

$$\begin{aligned}
w_{C1} &\equiv \overline{w_{C1}} = (\alpha_{C1} + \overline{\varepsilon_i}|\varepsilon_i > \varepsilon*)\frac{\partial F_1}{\partial L_1} \\
\frac{w_{C1}}{w_{N1}} &= \frac{(\alpha_{C1} + \overline{\varepsilon_i}|\varepsilon_i > \varepsilon*)}{\alpha_{N1}} > \frac{\alpha_{C1} + \varepsilon*}{\alpha_{N1}} = \frac{\alpha_{C2}}{\alpha_{N2}} = \frac{w_{C2}}{w_{N2}}.
\end{aligned}$$

Thus, the methodology of classifying occupations into "college" and "noncollege" as proposed by Gottschalk and Hansen (2003) is applicable also in this case.

Table 7: Summary statistics of the ISPV data

Year	Total	Education		Gender	
		College	High school	Male	Female
2000	123669	22%	78%	56%	44%
2001	134441	22%	78%	56%	44%
2002	134249	23%	77%	54%	46%
2003	138142	25%	75%	56%	44%
2004	164288	27%	73%	55%	45%
2005	173972	22%	78%	55%	45%
2006	185375	23%	77%	56%	44%
2007	220025	25%	75%	56%	44%
2008	231037	26%	74%	57%	43%

Note: The above table presents summary statistics of the sample of young workers, i.e., workers under 35 years of age.

Table 8: Colleges and universities in Czech NUTS-3 regions in 2008

Region (NUTS-3)	Colleges and universities		Enrolled students	
	public	private	public	private
Hlavni mesto Praha	32	24	49941	11392
Stredocesky kraj	3	3	33785	7709
Jihocesky kraj	5	3	21093	1921
Plzensky kraj	1	0	15687	1461
Karlovarsky kraj	1	1	7090	1594
Ustecky kraj	2	1	22020	3635
Liberecky kraj	1	0	11882	1490
Kralovehradecky kraj	1	0	17039	1227
Pardubicky kraj	1	0	16048	1050
Vysocina	2	1	17353	1285
Jihomoravsky kraj	12	7	38900	3280
Olomoucky kraj	3	2	21853	1952
Zlinsky kraj	2	1	22399	1265
Moravskoslezsky kraj	5	2	42936	3365

Note to **Table 9**: The entries in this table represent the absolute numbers of college graduates of given 5-year-wide age cohorts in Czech districts and the percentage changes in these numbers between 1991 and 2001. Only two age cohorts (of age 30-34 and 35-39 in the year 1991) are chosen, because younger cohorts might have still been in school in 1991 and older cohorts could be out of the labor force in 2001. The last row of the table presents the country's average change in the number of college graduates in given age cohorts. The majority of district-specific changes do not differ much from the country average, which is reflected in low variance of district-specific changes. There are only two outlying districts experiencing a decrease in the number of college graduates (Jindrichuv Hradec and Sumperk) and four districts experiencing a very large increase in this number (Benesov, Blansko, Rakovnik, and Uherske Hradiste).

Table 9: Changes in cohort-specific sizes of college-educated population by district from 1991 to 2001

	Born in 1961-1965			**Born in 1956-1960**		
	1991	**2001**	**Change**	**1991**	**2001**	**Change**
Benesov	539	654	21%	540	675	25%
Beroun	470	519	10%	420	500	19%
Blansko	620	801	29%	596	756	27%
Breclav	728	817	12%	644	728	13%
Bruntal	531	558	5%	555	611	10%
Ceska Lipa	549	672	22%	526	613	17%
Ceske Budejovice	1945	2136	10%	1888	2044	8%
Cesky Krumlov	347	394	14%	360	407	13%
Cheb	459	570	24%	497	589	19%
Chomutov	565	655	16%	507	561	11%
Chrudim	640	715	12%	555	607	9%
Decin	499	595	19%	524	644	23%
Domazlice	341	344	1%	310	350	13%
Frydek Mistek	1621	1880	16%	1518	1773	17%
Havlickuv Brod	639	718	12%	538	587	9%
Hodonin	973	1091	12%	871	969	11%
Hradec Kralove	1726	1887	9%	1819	1919	5%
Jablonec nad Nysou	657	722	10%	600	650	8%
Jicin	464	529	14%	487	557	14%
Jihlava	833	926	11%	684	754	10%
Jindrichuv Hradec	726	655	-10%	708	592	-16%
Karlovy Vary	791	884	12%	703	814	16%
Karvina	1770	1959	11%	1696	1845	9%
Kladno	967	1137	18%	1056	1195	13%
Klatovy	627	675	8%	580	627	8%
Kolin	542	629	16%	517	628	21%
Kromeriz	767	914	19%	743	838	13%
Kutna Hora	566	586	4%	489	531	9%
Liberec	1303	1382	6%	1189	1300	9%
Litomerice	609	668	10%	603	704	17%
Louny	543	543	0%	510	545	7%
Melnik	569	632	11%	555	611	10%
Mlada Boleslav	649	768	18%	693	802	16%
Most	620	668	8%	594	631	6%
Nachod	687	765	11%	602	693	15%
Novy Jicin	1082	1172	8%	961	1081	12%

	Born in 1961-1965			Born in 1956-1960		
	1991	**2001**	**Change**	**1991**	**2001**	**Change**
Nymburk	535	648	21%	464	563	21%
Olomouc	2209	2482	12%	2079	2358	13%
Opava	1175	1394	19%	1159	1318	14%
Ostrava-mesto	3010	3143	4%	3137	3315	6%
Pardubice	1415	1531	8%	1438	1500	4%
Pelhrimov	447	483	8%	399	470	18%
Pisek	607	614	1%	525	559	6%
Plzen	2021	2061	2%	2112	2204	4%
Plzen-jih	365	433	19%	323	383	19%
Plzen-sever	369	435	18%	293	357	22%
Prachatice	340	356	5%	307	343	12%
Prerov	1052	1118	6%	946	1005	6%
Pribram	912	921	1%	804	837	4%
Prostejov	766	849	11%	708	759	7%
Rakovnik	296	375	27%	325	371	14%
Rokycany	318	336	6%	226	281	24%
Rychnov nad Kneznou	497	554	11%	457	502	10%
Semily	445	534	20%	468	515	10%
Sokolov	334	399	19%	350	384	10%
Strakonice	516	521	1%	447	480	7%
Sumperk	1091	939	-14%	904	801	-11%
Svitavy	599	676	13%	521	580	11%
Tabor	918	953	4%	1034	1035	0%
Tachov	350	366	5%	276	287	4%
Teplice	490	600	22%	545	660	21%
Trebic	882	955	8%	780	878	13%
Trutnov	637	758	19%	638	743	16%
Uherske Hradiste	774	1077	39%	636	942	48%
Usti nad Labem	716	801	12%	739	813	10%
Usti nad Orlici	786	920	17%	717	845	18%
Vsetin	1121	1223	9%	988	1085	10%
Vyskov	620	700	13%	596	635	7%
Zdar nad Sazavou	787	881	12%	751	800	7%
Zlin	1742	1878	8%	1558	1700	9%
Znojmo	669	706	6%	651	709	9%
TOTAL	66575	74564	12%	63964	71640	12%
		Variance	0.0068		Variance	0.0068

Table 10: Estimates of occupation-specific college wage premia and classification into "college" and "noncollege" occupations.

	2000		2008	
Occupation group	premium	"college"	premium	"college"
Mathematicians, statisticians and related professionals	N/A	1	N/A	1
Legal professionals	N/A	1	N/A	1
Other department managers	0.379	1	0.651	1
Production and operations department managers	0.498	1	0.636	1
Directors and chief executives	0.194	1	0.583	1
Ship and aircraft controllers and technicians	0.638	1	0.535	1
Social science and related professionals	N/A	1	0.522	1
Business professionals	N/A	1	0.511	1
Material-recording and transport clerks	0.099	0	0.480	1
Numerical clerks	0.490	0	0.459	1
Travel attendants and related workers	0.094	0	0.459	1
Life science professionals	N/A	1	0.455	1
Finance and sales associate professionals	0.692	1	0.453	1
Archivists, librarians and related information professionals	N/A	1	0.447	1
Computer associate professionals	0.325	1	0.446	1
Safety and quality inspectors	0.283	1	0.414	1

Occupation group	2000 premium	2000 "college"	2008 premium	2008 "college"
Administrative associate professionals	0.339	1	0.412	1
Social work, artistic, entertainment and sports associate professionals	0.021	0	0.406	1
General managers	0.376	1	0.395	1
Business services agents and trade brokers	0.316	1	0.372	1
Optical and electronic equipment operators	0.333	1	0.356	1
Secretaries and keyboard-operating clerks	0.360	0	0.351	1
Housekeeping and restaurant services workers	0.263	0	0.341	1
Architects, engineers and related professionals	N/A	1	0.337	1
Cashiers, tellers and related clerks	0.318	0	0.332	1
Physicists, chemists and related professionals	N/A	1	0.326	1
Physical and engineering science technicians	0.289	1	0.325	1
Professional administrative workers	0.208	1	0.321	1
Computing professionals	0.243	1	0.312	1
Mining and mineral-processing plant operators	-0.021	0	0.289	1
Market-oriented skilled agricultural workers	0.102	0	0.275	1
Client information clerks	0.085	0	0.257	1

	2000		2008	
Occupation group	premium	"college"	premium	"college"
Other office clerks	0.304	1	0.241	1
Writers and creative or performing artists	0.227	1	0.236	1
Food processing and related trades workers	-0.067	0	0.235	1
Modern health associate professionals (except nursing)	0.570	1	0.227	1
Life science technicians and related associate professionals	0.250	1	0.198	1
Metal-processing plant operators	0.085	0	0.160	1
Power-production and related plant operators	-0.066	0	0.135	1
Protective services workers	-0.088	0	0.133	1
Building frame and related trades workers	0.837	1	0.109	1
Teaching and religious associate professionals	0.423	1	0.102	1
Metal processing workers	0.162	0	0.443	0
Automated assembly-line and industrial-robot operators	N/A	0	0.239	0
Electrical and electronic equipment mechanics and fitters	0.225	0	0.213	0
Printing-, binding- and paper-products machine operators	-0.287	0	0.201	0

	2000		2008	
Occupation group	premium	"college"	premium	"college"
Building finishers and related trades workers	0.374	0	0.190	0
Assemblers	-0.100	0	0.183	0
Precision workers in metal and related materials	0.178	0	0.146	0
Manufacturing laborers	0.126	0	0.145	0
Messengers, porters, doorkeepers and related workers	0.728	0	0.137	0
Library, mail and related clerks	0.031	0	0.134	0
Rubber- and plastic-products machine operators	-0.110	0	0.133	0
Chemical-processing plant operators	0.146	0	0.133	0
Chemical-products machine operators	0.541	0	0.131	0
Textile, garment and related trades workers	0.163	0	0.118	0
Potters, glass-makers and related trades workers	0.140	0	0.116	0
Locomotive engine drivers, ships' deck crews and related workers	0.131	0	0.112	0
Motor vehicle drivers	0.079	0	0.091	0
Glass, ceramics and related plant operators	0.203	0	0.083	0
Mining and construction laborers	-0.401	0	0.075	0
Railway and train technicians	0.077	0	0.065	0

Occupation group	2000 premium	2000 "college"	2008 premium	2008 "college"
Wood-processing and papermaking plant operators	0.121	0	0.055	0
Machinery mechanics and fitters	0.019	0	0.049	0
Other machine operators and assemblers	0.388	0	0.037	0
Textile-, fur- and leather-products machine operators	0.158	0	0.022	0
Models, salespersons and demonstrators	0.298	0	0.007	0
Handicraft workers in wood, textile, leather and related material	N/A	0	0.005	0
Blacksmiths, tool-makers and related trades workers	0.210	0	-0.016	0
Food and related products machine operators	0.139	0	-0.022	0
Metal-and mineral-products machine operators	-0.187	0	-0.026	0
Transport laborers and freight handlers	0.149	0	-0.051	0
Miners, shotfirers, stone cutters and carvers	-0.386	0	-0.107	0
Agricultural and other mobile plant operators	-0.099	0	-0.150	0
Street services elementary occupations, cleaners and launderers	-0.254	0	-0.225	0
Printing and related trades workers	-0.027	0	-0.305	0
Market-oriented forestry and fishery workers	0.165	1	-0.312	0

Note: This table presents occupation-specific college wage premia and classification of occupations into "college" (1) and "noncollege" (0), as discussed in section 6 of this paper. An occupation is classified as "noncollege" if the share of college graduates among its employees is less than 5% even if the college wage premium is above the 0.1 threshold.

CHAPTER 2

Measuring Skill Intensity of Occupations with Imperfect Substitutability Across Skill Types

Barbara Pertold-Gębicka*

CERGE-EI†

Abstract

In the absence of a model-based measure of occupational skill intensity, the literature on wage inequality cannot consistently track technological progress on occupational level — a key ingredient of recent theories of labor market polarization. In this paper, I use the March CPS data from 1983 to 2002 to estimate such a measure corresponding to occupation-specific relative productivities of college and high-school educated. With imperfect substitution across skill types, the measurement of relative productivities requires estimation of substitution elasticities, and I propose a simple strategy to obtain these. The resulting measure is used to shed light on the modified skill-biased technological change hypothesis proposed by Autor et al. (2006).

*I would like to thank Alena Bičáková, David Card, Randall Filer, Štěpán Jurajda, Peter Katuščák, Enrico Moretti and Andrea Weber for valuable comments and suggestions. Part of this paper was written while visiting the Department of Economics at the University of California at Berkeley and I would like to thank the department for its hospitality. All remaining errors are mine. This research was supported by the Grant Agency of Charles University funding No. 35209.
E-mail: barbara.gebicka@cerge-ei.cz

†A joint workplace of the Center for Economic Research and Graduate Education, Charles University, Prague, and the Economics Institute of the Academy of Sciences of the Czech Republic. Address: CERGE-EI, Politickych veznu 7, Prague 11121, Czech Republic.

1 Introduction

The literature on wage inequality was able to successfully account for the majority of wage-structure shifts of the twentieth century (including the rising returns to education in the face of rising educational attainment in the 1980s) by employing a framework with high- and low-skilled workers supplying labor to a homogeneous labor market with factor-augmenting technology (Katz and Murphy, 1992; Bound and Johnson, 1992). Nevertheless, the varying extent of skill-biased technological progress (Card and DiNardo, 2002) or the recently documented earnings growth polarization (Goos and Manning, 2007; Autor et al., 2006) are not accounted for by this framework. The search for a coherent explanation for recent trends has led researchers to analyze the labor market partitioned at the level of occupations, which allows for a natural way to introduce the differential speed of technological progress (Firpo et al., 2011). This literature, however, so far has not employed a model-based measure of occupation-specific technological progress or skill intensity.

The occupation-focused literature began with the works of Autor et al. (2003), Goos and Manning (2007) and Autor et al. (2006), who propose a modified version of the skill-biased technological change (SBTC) hypothesis. They argue that new technologies have a heterogeneous impact on workers. In particular, technologies complement workers performing non-routine cognitive tasks and substitute for workers performing routine tasks. Also, a chapter of the recent Handbook of Labor Economics by Acemoglu and Autor (2010) highlights that the mix of tasks performed by a worker defines the impact of technological progress on her productivity. Thus, to the extent that occupations capture the task content of work, the occupation-level analysis offers a key towards understanding the impact of technological change on wage structures. In the context of technological progress and the related demand for skilled labor, it is helpful to link occupations to their skill intensity. The latter would translate the occupation-specific task mix into the demand for skills defined by an occupation-specific production function.

Currently there is no consensus on how to capture the skill intensity of occupations and the literature offers several simple alternatives. One strategy is to rely on the description of skills, tasks and work activities associated with individual occupations as reported in occupation dictionaries such as the Occupational Information Network (O*NET), the replacement for the Dictionary of Occupational Titles (DOT). This comprehensive source of information about occupations is widely used in the literature in the context of income inequality and wage structures (Autor and Dorn, 2009) and overeducation (McGoldrick and Robst, 1996). The multidimensional description of occupations given by the dictionaries provides a a valuable insight into the changing structure of job tasks and its relation to the observed wage structures (Firpo et al., 2011; Acemoglu and Autor, 2010). On the other hand, while capturing IT use or manual task involvement is relatively simple using occupation dictionaries, other dimensions of the SBTC are not captured systematically and in a harmonized way across occupations. For example, it is less convenient to use the O*NET for studying the implications of technological progress on the demand for educated labor in an international context, as it lacks information on occupation-specific demand for education and it is not available outside the U.S.

A clear definition of occupations' demand for skills allowing for straightforward cross-country comparability is offered by several alternative skill intensity measures, although at the cost of being derived from the data used to investigate wage structures. Some studies of technological progress on the occupational level use employees' average years of schooling as a proxy for the skill content of occupations (Goos and Manning, 2007; Autor et al., 2006). This approach relies on a strong assumption that the employment structure of occupations correctly reflects their skill requirements; one can easily imagine violation of this assumption in occupations that are in the process of rapidly adjusting their skill requirements. Despite these shortcomings, the average years of schooling measure is also used to define occupations' demand for educated labor when studying the fraction of college graduates underutilizing their skills, i.e., working in the so-called "noncollege" occupations (Pryor and Schaffer, 1997).

In this line of research, Gottschalk and Hansen (2003) – further referred to as GH – offer a model-based approach for defining occupation-specific demand for educated labor. Their methodology is used in this paper as a starting point for defining a measure of skill intensity of occupations. Assuming that production technologies are homogeneous within occupations, the occupation-specific relative productivity of differently skilled workers reflects the utilization of their skills, thus offering a continuous model-based measure of occupation-specific skill intensity. GH assume perfect substitutability between differently educated workers which allows them to use the wage gap to measure the relative productivity of college and high school graduates. However, there are many studies estimating the market-wide elasticity of substitution between more and less educated labor in the U.S. to be around 1.4,[1] which requires imperfect within-occupation substitution between the two types of workers and/or outputs of individual occupations being not well substitutable.

This study generalizes the GH approach by estimating the within-occupation elasticity of substitution between high school and college graduates. Following the common practice in the literature, I assume that occupation-specific production functions are of the constant elasticity of substitution (CES) type, and I use a modification of the strategy proposed by Card (2001) to estimate their elasticity parameters. These, combined with the observed relative employment and wages, allow me to derive occupation-specific relative productivities of college and high school graduates, which provide a measure of skill intensity of occupations. It can be used, for example, to track the technological progress of individual occupations or to derive the demand for educated labor within different groups of occupations. In this study I use the measure of skill intensity of occupations to analyze the recent polarization of earnings growth in the U.S.

The rest of the paper is organized as follows. Section 2 briefly explains the idea behind using occupation-specific relative productivities of differently skilled labor as a measure of skill intensity of occupations. In the next section, I present a model of worker allocation across occupations characterized by different skill intensity. This model is further used for empirical analysis. Section 4 describes econometric procedures used to identify occupation-specific elasticities of substitution between college

[1] Ciccone and Peri (2005) offer a review of these studies.

and high school graduates that allow for estimation of the skill intensity of occupations. The next section presents the results of these estimations. In section 6, I use the estimated occupation-specific skill intensities to analyze the earnings growth polarization. The last section concludes.

2 The measure of skill intensity

Within-occupation relative productivity of college and high school graduates, where college graduates represent highly-skilled labor and high school graduates represent less-skilled labor, can be used as a proxy for occupation-specific skill intensity. Let me illustrate this point using a relatively general occupation-specific production function – the constant elasticity of substitution (CES) aggregate of college- and high school-educated labor, as specified in Equation (1).

$$Y_j = \left(\alpha_{Cj} L_{Cj}^{\gamma_j} + \alpha_{Nj} L_{Nj}^{\gamma_j}\right)^{\frac{1}{\gamma_j}}, \tag{1}$$

where Y_j is the output of occupation j, L_{Cj} is the number of college graduates, L_{Nj} is the number of high school graduates employed in occupation j, and γ_j is a parameter describing the substitutability between these two labor types (the elasticity of substitution is $\sigma_j = \frac{1}{1-\gamma_j}$). In this context, $\frac{\alpha_{Cj}}{\alpha_{Nj}}$ describes the occupation-specific relative productivity of differently educated workers. In occupations where this parameter assumes high values, college graduates are much more productive than high school graduates, which could be attributed to the skill difference among differently educated workers. That is why $\frac{\alpha_{Cj}}{\alpha_{Nj}}$ describes the skill intensity of an occupation. It tells us how crucial college-gained skills are for the tasks performed within a specific occupation.

Under the simplifying assumption made by GH, i.e., when the elasticity of substitution between college and high school graduates is infinite ($\gamma_j = 1$), $\frac{\alpha_{Cj}}{\alpha_{Nj}}$ is fully reflected in the relative wage of the two education groups. This is why GH classify occupations according to the college wage premia that they pay. The perfect substitutability assumption is, however, questionable. One could easily come up with examples of occupations where the elasticity of substitution between college and high school graduates is zero (e.g., medical doctors) or where it is highly limited (e.g., financial advisors). Relaxing the infinite elasticity of substitution assumption (i.e., allowing for $\gamma_j < 1$) and rearranging the first-order conditions for firms' profit maximization problem gives

$$\frac{\alpha_{Cjt}}{\alpha_{Njt}} = \frac{w_{Cjt}}{w_{Njt}} \left(\frac{L_{Cjt}}{L_{Njt}}\right)^{1-\gamma_j} = \frac{w_{Cjt}}{w_{Njt}} \left(\frac{L_{Cjt}}{L_{Njt}}\right)^{-\frac{1}{\sigma_j}}. \tag{2}$$

Thus, in the setup where college and high school graduates are allowed to be imperfect substitutes, one needs to know the elasticity of substitution between them in order to derive the occupation-specific relative productivity.[2]

3 A model of labor allocation across occupations

In this section I outline a theoretical model describing the allocation of differently skilled labor across occupations characterized by different skill intensity and different substitutability between skill types. The model explains why observationally similar people are found in different (and differently paying) occupations. It also provides the baseline for an econometric specification used to estimate occupation-specific elasticity of substitution between college and high school graduates.

3.1 Demand for labor

Let us assume that the economy produces one uniform good which sells at price p. This good is produced using J different occupations with production technology described by a twice-differentiable function $G(\cdot)$:

$$Y = G\left(L_1, L_2, ...L_J\right).$$

Each occupation could be described as a technology aggregating two labor types: college and high school graduates. The "output" of occupation j is labor aggregate L_j being a CES combination of college- and high school-educated labor. Occupations differ in their skill intensity ($\frac{\alpha_{C_j}}{\alpha_{N_j}}$) and in the elasticity of substitution between college

[2]Note that setting $\gamma_j = 1$ in the occupation-specific production function (1), one gets $\frac{\alpha_{Cjt}}{\alpha_{Njt}} = \frac{w_{Cjt}}{w_{Njt}}$, as in GH.

and high school graduates ($\sigma_j = \frac{1}{1-\gamma_j}$). As before, the production function used by occupation j could be summarized in the following way:

$$L_j = \left(\alpha_{Cj} L_{Cj}^{\gamma_j} + \alpha_{Nj} L_{Nj}^{\gamma j}\right)^{\frac{1}{\gamma j}}, \tag{3}$$

where L_{Cj} and L_{Nj} are the amounts of college- and high school-educated labor employed in occupation j.

In a competitive market, under the above-specified functions, wages of each education group in occupation j should be equal to their marginal products, as expressed by the following first-order conditions:

$$\begin{aligned} w_{Cj} &= p\frac{\partial Y}{\partial L_j}\frac{\partial L_j}{\partial L_{Cj}} = p\frac{\partial Y}{\partial L_j} L_j^{1-\gamma_j} \alpha_{Cj} L_{Cj}^{\gamma_j - 1}; \\ w_{Nj} &= p\frac{\partial Y}{\partial L_j}\frac{\partial L_j}{\partial L_{Nj}} = p\frac{\partial Y}{\partial L_j} L_j^{1-\gamma_j} \alpha_{Nj} L_{Nj}^{\gamma_j - 1}. \end{aligned}$$

These equations lead to the formulation of the relative wage of college and high school graduates in occupation j:

$$\frac{w_{Cj}}{w_{Nj}} = \frac{\alpha_{Cj}}{\alpha_{Nj}} \left(\frac{L_{Nj}}{L_{Cj}}\right)^{1-\gamma_j}, \tag{4}$$

which, after rearrangement and substitution of $\sigma_j = \frac{1}{1-\gamma_j}$, gives

$$\ln\left(\frac{L_{Cj}}{L_{Nj}}\right) = \sigma_j \ln\left(\frac{\alpha_{Cj}}{\alpha_{Nj}}\right) - \sigma_j \ln\left(\frac{w_{Cj}}{w_{Nj}}\right). \tag{5}$$

Equation (5) describes the relative labor demand in occupation j. It depends on the relative wages of the two education groups, their relative productivities and the elasticity of labor substitution within occupation j.

3.2 Supply of labor

Let us assume now that there are N_{Cj} college-educated workers and N_{Nj} high school-educated workers who could potentially supply labor to occupation j (N_{Cj} and N_{Nj} describe labor markets specific to occupation j). The notion of occupation-specific labor markets, introduced by Card (2001), is used to accommodate the observation that a worker usually looks for employment in a specific occupation; however, she has some flexibility to switch occupations as a reaction to productivity shocks affecting the labor market. In this context, N_{Cj} and N_{Nj} capture all workers who would supply labor to occupation j under favorable labor market conditions. Only some of these people are actually observed working in occupation j because workers differ in their occupation-specific reservation wage. This leads to the formulation of the supply of labor to occupation j as a fraction of the total size of this occupation's specific labor market:[3]

$$\begin{aligned} \ln\left(\frac{L_{Cj}}{N_{Cj}}\right) &= \beta_j \ln w_{Cj} \qquad (6) \\ \ln\left(\frac{L_{Nj}}{N_{Nj}}\right) &= \beta_j \ln w_{Nj}. \end{aligned}$$

Log-linear aggregate labor supply functions are commonly used when describing the supply of workers to different units of production, usually occupations (Card, 2001; Gottschalk and Hansen, 2003). The occupation-specific elasticity of labor supply, $\beta_j > 0$, represents workers' aggregate preferences towards occupation j. It is assumed to be the same for each education group within the occupation-specific labor market. This assumption is crucial for the model to have a closed-form solution. Despite being strong, this assumption is actually less restrictive than the relaxed assumption about $\gamma = 1$, where the labor supply can be any but equilibrium values are determined from the total demand.

The above specified supply functions can be combined into one equation describing the relative supply of labor into occupation j:

$$\ln\left(\frac{L_{Cj}}{L_{Nj}} / \frac{N_{Cj}}{N_{Nj}}\right) = \beta_j \ln\left(\frac{w_{Cj}}{w_{Nj}}\right), \qquad (7)$$

which depends on the relative wages of the two education groups and the occupation-specific elasticity of labor supply.

3.3 Equilibrium

Equations (5) and (7) describe the relative demand and supply of labor for occupation j. Equalizing supply with demand and rearranging, one arrives at a system capturing

[3]Let me note that $\frac{L_{Cj}}{N_{Cj}}$ and $\frac{L_{Nj}}{N_{Nj}}$ are restricted not to exceed 1, which is not captured by the presented functions. I do not incorporate these restrictions in the model because in reality they never bind.

the equilibrium relative wages and relative employment in each occupation:

$$\begin{cases} \ln\left(\frac{w_{Cj}}{w_{Nj}}\right) = \frac{\sigma_j}{\sigma_j+\beta_j}\ln\left(\frac{\alpha_{Cj}}{\alpha_{Nj}}\right) - \frac{1}{\sigma_j+\beta_j}\ln\left(\frac{N_{Cj}}{N_{Nj}}\right) \\ \ln\left(\frac{L_{Cj}}{L_{Nj}}\right) = \frac{\sigma_j\beta_j}{\sigma_j+\beta_j}\ln\left(\frac{\alpha_{Cj}}{\alpha_{Nj}}\right) + \frac{\sigma_j}{\sigma_j+\beta_j}\ln\left(\frac{N_{Cj}}{N_{Nj}}\right) \end{cases}. \tag{8}$$

Let us note that both relative wages and relative employment depend on occupation-specific supply factors (total relative amounts of college- and high school-educated workers in occupation-specific labor markets) and demand factors (relative productivity of college and high school graduates). The shape of these dependencies is described jointly by the occupation-specific elasticity of labor supply and the elasticity of substitution between the two labor types.

The system derived above describes how the observed occupation-specific employment structure and wages depend on the relative number of college and high school graduates ready to supply labor to that occupation. These formulas strongly rely on the functional forms assumed, i.e., on the shape of the production function and the shape of the labor supply. Nevertheless, the CES production function and the log-linear supply function are the functional forms most widely used in the context of labor-labor substitutability and occupational choice; as such, they constitute a good baseline for this study. Observing occupation-specific labor allocation, relative wages and the structure of this occupation's labor market, one can use the derived system to estimate the elasticity of substitution between more and less educated labor as well as the occupation-specific elasticity of labor supply.

4 Econometric approach

Under the assumption that the occupation-specific elasticities of substitution between college and high school graduates and the elasticities of labor supply do not change over time, I can use the above presented model to estimate them. To do so, let me analyze an economy, as described in the previous section, over several consecutive periods (subscribed by t). In each period the occupation-specific supply and demand factors are different. The relative amounts of college- and high school-educated workers in occupation-specific labor markets vary with the sociodemographic structure of the population, current popularity of occupations, and the fraction of college graduates in the total population. The relative productivity of college and high school graduates varies with the SBTC.[4] These movements of the relative supply and demand curves lead to the observation of different equilibrium values of occupation-specific relative wages and employment, which can be used to estimate the system of equations as presented in (8).

To completely specify the model, let me decompose the (unobserved) variation in the relative productivity of labor into three components: occupation-specific (characteristic of a given occupation, constant over time), year-specific (common for all occupations) and occupation-year specific effects. It is usual to assume that the occupation-specific component is deterministic, while the other two are stochastic (Card, 2001), which can be expressed as $\ln\left(\frac{\alpha_{Cjt}}{\alpha_{Njt}}\right) = \ln(\alpha_j) + \varepsilon_t + \varepsilon_{jt}$. Using this notation, the system (8) can be rewritten into the following econometric model:

$$\begin{cases} \ln\left(\frac{w_{Cjt}}{w_{Njt}}\right) = c_{j0} + c_{j1}\ln\left(\frac{N_{Cjt}}{N_{Njt}}\right) + v_t + v_{jt} \\ \ln\left(\frac{L_{Cjt}}{L_{Njt}}\right) = d_{j0} + d_{j1}\ln\left(\frac{N_{Cjt}}{N_{Njt}}\right) + \mu_t + \mu_{jt} \end{cases}, \tag{9}$$

where $c_{j0} = \frac{\sigma_j}{\sigma_j+\beta_j}\ln(\alpha_j)$, $c_{j1} = -\frac{1}{\sigma_j+\beta_j}$, $v_t = \frac{\sigma_j}{\sigma_j+\beta_j}\varepsilon_t$, $\nu_{js} = \frac{\sigma_j}{\sigma_j+\beta_j}\varepsilon_{jt}$ and $d_{j0} = \frac{\sigma_j\beta_j}{\sigma_j+\beta_j}\ln(\alpha_j)$, $d_{j1} = \frac{\sigma_j}{\sigma_j+\beta_j}$, $\mu_s = \frac{\sigma_j\beta_j}{\sigma_j+\beta_j}\varepsilon_t$, $\mu_{js} = \frac{\sigma_j\beta_j}{\sigma_j+\beta_j}\varepsilon_{jt}$.

This model describes the simultaneous determination of occupation-specific relative wages and relative employment as a function of the relative numbers of college- and high school-educated workers in occupation-specific labor markets in a given time period t. Note that the occupation-specific elasticity of substitution between college and high school graduates, σ_j, could be expressed as $\sigma_j = -\frac{d_{j1}}{c_{j1}}$. Thus, consistent estimation of c_{j1} and d_{j1} allows for the identification of σ_j. Before turning to the estimation, however, one has to acknowledge several important features of the model and data used in the analysis.

First, consider the endogenous nature of occupation-specific labor markets. As a result of a positive skill-biased productivity shock affecting occupation j, relative

[4] Note that according to the modified version of the SBTC, technological progress might have a positive influence on the relative productivity in some occupations while having a negative effect on others.

wages and relative employment of college graduates in this occupation increase. At the same time, however, more college graduates enter this occupation-specific labor market, as they see a possibility of high returns to education. Due to this effect, the OLS estimates of c_{j1} and d_{j1} are likely to be biased upwards. In the existing literature, such a problem is commonly dealt with by assuming that the time evolution of relative productivity is log-linear (Katz and Murphy, 1992; Card and DiNardo, 2002; Autor et al., 2008), i.e., that $\varepsilon_t + \varepsilon_{jt}$ can be approximated by a linear time trend. Although this does not capture all the unobservable shocks to relative labor productivity, it captures the ones that can be anticipated by workers and thus might influence the structure of the occupation-specific labor market.

Second, the explanatory variable $\frac{N_{Cjt}}{N_{Njt}}$, is not directly observable in the data. Estimating this variable using fitted values from a multinomial logit model introduces a measurement error satisfying the classical error-in-variables (CEV) assumptions. To mitigate this problem, I rely on two alternative approaches to estimate the sizes of occupation-specific labor markets. As discussed in the next section, the measurement errors of these estimates are uncorrelated. In the final estimation one measure is used as an instrument for the other to reduce the attenuation bias (Griliches and Mason, 1972).

Finally, the disturbance terms from the relative wage and relative employment equations for a single occupation are expected to be correlated between themselves, as they are both derived from the stochastic part of the relative productivity, $\varepsilon_t + \varepsilon_{jt}$. While this feature does not invalidate the estimates of the model coefficients, taking it into account can greatly improve the estimation efficiency. Thus, I estimate the elasticity parameters of each occupation using a 2-equation system of seemingly unrelated regressions (SUR).

Taking into account the above-discussed properties, the final econometric model is specified in the following way:

$$\begin{cases} \ln\left(\frac{w_{Cjt}}{w_{Njt}}\right) = c_{j0} + c_{j1}\ln\left(\frac{N_{Cjt}}{N_{Njt}}\right) + c_{j2}t + \zeta_{jt} \\ \ln\left(\frac{L_{Cjt}}{L_{Njt}}\right) = d_{j0} + d_{j1}\ln\left(\frac{N_{Cjt}}{N_{Njt}}\right) + d_{j2}t + \xi_{jt} \end{cases}, \tag{10}$$

where $c_{j2}t+\zeta_{jt}=v_t+v_{jt}$ and $d_{j2}t+\xi_{jt}=\mu_t+\mu_{jt}$, with ζ_{jt} and ξ_{jt} being uncorrelated with the true value of $\ln\left(\frac{N_{Cjt}}{N_{Njt}}\right)$. When estimating this model, I use an estimate of the relative size of the occupation-specific labor market, $\ln\left(\frac{N_{Cjt}}{N_{Njt}}\right)^A$, which is instrumented by an alternative measure, $\ln\left(\frac{N_{Cjt}}{N_{Njt}}\right)^B$. The whole system is estimated using the SUR approach.

Under the assumption that predictable shocks to occupation-specific relative labor productivity follow a linear trend and the measurement errors in the two estimates of occupation-specific labor markets are uncorrelated, the above presented approach leads to consistent estimation of $\widehat{c_{j1}}$ and $\widehat{d_{j1}}$. These estimates are further used to calculate the elasticity of substitution between more and less educated labor: $\widehat{\sigma_j}=-\frac{\widehat{d_{j1}}}{\widehat{c_{j1}}}$. Finally, one can combine $\widehat{\sigma_j}$'s estimated separately for each occupation with occupation-specific estimates of the college wage premium and relative employment to calculate the relative productivities as

$$\frac{\widehat{\alpha_{Cjt}}}{\alpha_{Njt}}=\frac{w_{Cjt}}{w_{Njt}}\left(\frac{L_{Cjt}}{L_{Njt}}\right)^{-\frac{1}{\widehat{\sigma_j}}}. \tag{11}$$

This is the measure used in this study to define the skill intensity of occupations.

5 Data and measurement issues

The data used in this study come from the 1983-2002 March Supplement to the Current Population Survey (March CPS), which means that I observe earnings for the years 1982 through 2001. This is the longest time span with consistent occupational coding, which is crucial for the analysis.[5] Due to the limited number of observations offered by March CPS, three consecutive years had to be merged to obtain sample sizes large enough to allow the data-hungry occupation-level analysis to be conducted. This means that data used to analyze year t are composed of $t-1$, t and $t+1$ March CPS samples. Thus, I can effectively analyze years 1983 - 2000. This time period

[5]In 1983, CPS started to use the 1980 Census occupation codes. These were later substituted by 1990 Census occupation codes which, however, introduced only minor changes. The 2000 Census occupational classification introduced to CPS in 2003 differs substantially from the previous ones.

covers a decade of rapid increase in the college-high school wage gap as well as the later slowdown in the rate of growth of this gap. Thus, it should be enough to capture any interesting phenomena in the labor market.

In order to make my analysis comparable to GH, I apply the same restrictions to the data as these authors do. Only male and female workers with at least a high school diploma and no more than a college degree are included in the sample. I do not construct college equivalents and high school equivalents, as many studies do. Instead, I focus on occupational allocation of college graduates with no higher degree as compared to high school graduates not having a college diploma. To avoid the issue of imperfect substitutability between experience groups, as discussed by Card and Lemieux (2001), GH and I concentrate on recent school leavers defined as individuals with 10 or fewer years of potential labor market experience.[6] Both full-time and part-time workers are included in the sample to ensure a sufficient number of observations. However, self-employed individuals are excluded from the sample as are those with reported working hours per week of zero or above 98. The earnings measure used in this analysis is the log of weekly earnings defined as yearly wage and salary income divided by weeks worked last year. Earnings are expressed in 2000 dollars.

I deal with earnings censoring by assigning the cell-means of earnings to the top-coded individuals. Starting in 1996, the cell-means are reported in the March CPS, while the cell-means for years 1983-1995 are calculated by Larrimore et al. (2008). Re-coding of occupations due to the switch from the 1980 to the 1990 Census occupational classification is done according to the scheme proposed by Meyer and Osborne (2005). Finally, for the earlier years, when March CPS reported the years spent in education instead of the highest degree obtained, I use the sample of individuals having 12-17 years of education (Jeager, 1997). Those with 16 or 17 years of education are assumed to be college graduates. Occupations are defined at a 3-digit level. However, some of the 3-digit categories had to be merged with other 3-digit

[6]Potential labor market experience is calculated as $age - years\ of\ schooling - 6$.

categories to ensure sufficient sample sizes.[7]

5.1 Relative wages and relative employment measures

To calculate the relative wages of college and high school graduates, I use the regression-adjusted wages of individuals. The controls included in the log-wage regressions, widely used to estimate returns to college, are experience, gender, race, education, full-time work status, and dummies for years $t-1$ and $t+1$.

Relative employment is calculated as the ratio of the numbers of college and high school graduates observed in a given occupation in a given year weighted by individual sample weights.

5.2 Occupation-specific labor markets

Occupation-specific labor markets, N_{Cjt} and N_{Njt}, are not directly observed in the data. They are composed of all workers who would supply labor to occupation j in period t if the labor market conditions were favorable enough. As one never knows what fraction of potential employees actually supplies labor to occupation j, it is not possible to measure the sizes of occupation-specific labor markets precisely and the measurement error associated with predicting the size of such a labor market might be correlated with the observed number of employees, i.e., in times economically favorable for a given occupation we might overestimate the size of this occupation's labor market. To mitigate the effect of measurement error, I rely on two alternative approaches to estimate N_{Cjt} and N_{Njt}. First, I draw on Card (2001), who proposes considering an individual's occupation as a probabilistic outcome that depends on her underlying characteristics. Let me call the obtained variable the *probabilistic measure.* Second, I construct transition matrices which define overlaps between occupation-specific labor markets to obtain the *overlapping markets measure.*

[7] Gottschalk and Hansen (2003) give a detailed description of occupational coding and aggregation.

5.2.1 The probabilistic measure

Card's (2001) idea is that individuals with a given education level choose which occupation labor market to enter based on their predispositions and the expected labor market conditions. These predispositions (proxied by observable demographic and other characteristics) determine the probabilities (π_{ij}) of choosing each occupation given the expectations about the labor market conditions. Under these assumptions, the number of people who could potentially work in occupation j at time t can be expressed as the sum of π_{ij}'s across the active population.

The probability of working in occupation j should be estimated against all other occupations, as these are competing choices. An obvious choice in this context is to apply the multinomial logit. This model is, however, very computationally demanding and difficult to track when the number of possible choices is large. While Card (2001) deals with 6 broad occupational categories, this study analyses 90 3-digit occupations. To overcome this problem, I propose that for each occupation (and at each education level) a group of so-called "neighboring" occupations is defined. This group contsists of all occupations *from which* we observe a significant number of workers switching to occupation j and *to which* a significant number of workers from occupation j switch. To find these occupations, I utilize the information on individuals' current and previous occupation as provided in March CPS.[8]

Once "neighboring" occupations are defined for each occupation at each education level, the multinomial logit model of occupational choice is estimated. For each employed individual[9] with a given education level, I estimate the probability of choosing occupation j from among all the "neighboring" occupations as a function of her demographic characteristics, such as gender, age and race, as well as the region where she lives and a time trend which controls for the predictable shifts in

[8] An occupation k is defined as a neighbor of occupation j in year t for college (high school) graduates if more than 5% of occupation j's employees with a college (high school) diploma were employed in occupation k in year $t-1$ or more than 5% of those employed in occupation j in year $t-1$ are working in occupation k in year t.

[9] The unemployed are not taken into account in this study. It is expected to have a negligible effect on the results because I analyze relatively highly-educated individuals.

occupation attractiveness. The estimated equation is as follows:

$$prob(occ_{it} = j) = G\left(X_{it}\beta_j + \psi_1 t + \eta_{itj}\right), \tag{12}$$

where the dependent variable equals one if an individual i works in occupation j at time t, X_{it} contains individual demographic characteristics and regional dummies, t is the time trend, and η_{itj} captures individual unobservable effects. This approach allows me to estimate the importance of each characteristic for working in occupation j's labor market given the expected labor market conditions (proxied by the time trend). The estimate of β_j is then used to predict the individual-specific probability of working in occupation j, i.e., $\widehat{\pi_{ij}}$, cleared of time effects. The year-specific sum of these fitted values over all individuals represents occupation j's specific labor market in the given year. This measure could be thought of as the number of people who would work in occupation j in year t if the productivity shocks experienced by this occupation exactly followed the expected trend. As such, this measure is independent of yearly deviations from the time trend which drive the variation in relative wages and relative quantities of labor actually employed in a given occupation.

5.2.2 The overlapping markets measure

The alternative measure of occupation-specific labor markets is based on aggregate trends rather than individual predispositions. It assumes that occupation-specific labor markets overlap to a well-defined extent. One can understand the overlap between two occupations' markets as the fraction of people employed in occupation k who belong to occupation's j labor market. Knowing these fractions one can easily calculate the sizes of occupation-specific labor markets as the sum of employment in all occupations weighted by the respective overlaps.

Assuming that the extent of the cross-occupational overlap of the labor markets follows a linear time trend (with slight variations caused by year-specific shocks), I can use the pooled data from the whole time period covered in this study to construct education-specific transition matrices, T_{Ct} and T_{Nt}, whose elements in the k-th row and j-th column represent the average fraction of workers in occupation k who move to occupation j within a year. The elements of these matrices are treated as proxies for the fraction of workers observed in occupation k who also belong to the labor market of occupation j. That is why the elements on the diagonal are set to be 1.[10]

With the transition matrices in hand, one can retrieve the total number of college and high school graduates ready to supply labor to each occupation j by observing employment in all 90 occupations. Under the assumptions stated above, the occupation-specific labor market at time t can be defined as the weighted sum of all workers with a given education level employed in each occupation in the given year. The weights are composed of the elements of the j-th columns of the education-specific transition matrices:

$$\begin{aligned} N_{Cjt} &= T_{Ctj} \times L_{Ct} \\ N_{Njt} &= T_{Ntj} \times L_{Nt}, \end{aligned}$$

where T_{Ctj} and T_{Ntj} are the j-th columns of matrices T_{Ct} and T_{Nt}, and L_{Ct} and L_{Ct} are the horizontal vectors of employment of college and high school graduates in all occupations in year t.

The two approaches to measure N_{Cjt} and N_{Njt} result in similar estimates of occupation-specific labor markets.[11] Nevertheless, they are based on different assumptions and are disturbed by different factors: the overlapping markets measure is identified with occupation switchers, while the probabilistic measure is defined with stayers. Thus, I use one measure to instrument for the other to reduce the measurement error bias when estimating system (10).

6 Skill-intensity estimates

This section presents step-by-step results leading to the estimation of occupation-specific skill intensities. As explained in Section 4, the main challenge of this analysis,

[10] Transition matrices were constructed using observations of occupation-switchers. Observed number of switchers from occupation k to occupation j between years t and $t+1$ was regressed on a constant and a linear time trend. Predicted values of this regression were used as the average fraction of workers in occupation k who move to occupation j within in during year t.

[11] The correlation between these two measures is 0.855.

and the main contribution of this study, is the estimation of occupation-specific elasticities of substitution between college and high school graduates.

Estimation of the substitution elasticities using the system of equations (10) can be implemented for occupations employing significant amounts of both labor types, which in this study are defined as occupations with at least 10% of employees having only a high school diploma and at least 5% of employees being college graduates. Occupations where college and university graduates constitute the wide majority of employees are treated as licensed occupations (i.e., occupations where holding a degree is required by law), which implies an elasticity of substitution between college and high school graduates of zero.[12] Occupations where hardly any college graduates are employed are treated as not attractive for highly-educated workers, which implies perfectly inelastic labor supply and does not allow for estimation of the within-occupation substitution elasticities.[13] For the remaining 73 occupations, the system (10) is estimated and the estimates of c_{j1} and d_{j1} are recorded.

For many occupations c_{j1} is found not to be statistically different from zero. These are plausible values. The parameter c_{j1} is expected to be zero for occupations where college and high school graduates are perfect substitutes ($\sigma_j = \infty$) or where workers supply labor perfectly elastically ($\beta_j = \infty$). In the latter case, d_{j1} should also be zero, while in the former, d_{j1} is expected to be one. This property can be used to distinguish between the two cases. Additionally, d_{j1} is expected to be zero (but c_{j1} significant and negative) for occupations where it is impossible to substitute between college and high school graduates ($\sigma_j = 0$). For all other occupations, the substitutability between workers with different education levels is found to be finite and positive. The full list of the estimates of substitution elasticities ($\widehat{\sigma}_j = -(\widehat{d}_{j1}/\widehat{c}_{j1})$) is found in the second column of Table 1 in the appendix. Note that only 28 of all

[12] These occupations include architects, biological and life scientists, health diagnosing occupations, judges, lawyers, postsecondary teachers, secondary school teachers, elementary school teachers, special education teachers, and speech therapists.

[13] These occupations include cashiers, food preparation and service occupations, freight, stock, material handlers, and service station occupations, mail and message distributing occupations, mechanics and repairers, vehicle and industrial machinery, transportation and material moving occupations, waiters and waitresses.

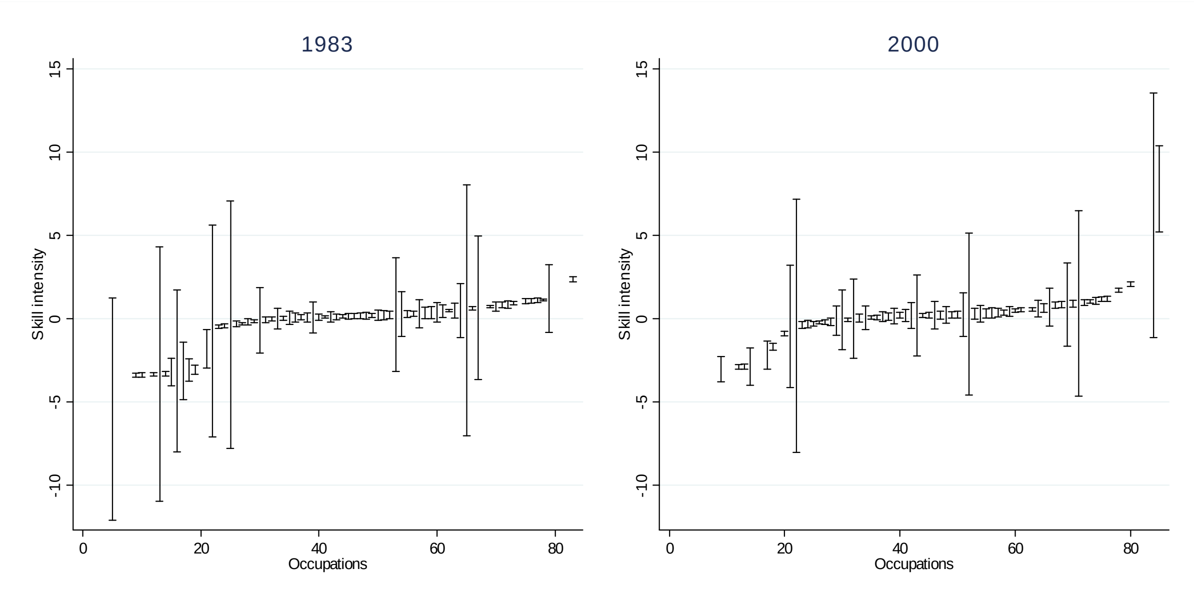
90 analyzed occupations are characterized by the elasticity of substitution between college and high school graduates being non-zero and finite. These are, however, the occupations for which skill (or educational) requirements are often discussed – sales workers, record processing occupations, or computer technicians, for example – which strengthens the argument that the elasticity of substitution between different labor types is crucial when analyzing the skill intensity of occupations.

As I am not aware of any other study estimating the occupation-specific substitution elasticities, I can only compare my estimates to previous economy-wide measures. These estimate the elasticity of substitution between differently educated workers to be between 1 and 3 (Ciccone and Peri, 2005). My finite estimates of the within-occupation elasticity of substitution between college and high school graduates are of the same order of magnitude – they vary from 0.5 to 10, with a median of 2.7. Occupations with the highest substitution elasticities include artists, sales workers, record processing and service occupations. They involve jobs that can be performed well by college and high school graduates. Occupations with the lowest, but still finite, elasticity of substitution include more specialist jobs like legal assistants, purchasing agents or insurance specialists. It is intuitive to think about these jobs as not equally performed by college and high school graduates. Further specialist occupations like therapists, health assistants and some management related occupations are found to have the elasticity of substitution between college and high school graduates equal to zero, while among the occupations characterized by perfect substitutability, we can find all types of office and administrative occupations.

The estimated elasticities of substitution are further used to calculate occupation-time specific relative productivities of college and high school graduates – the measure of the skill intensity of occupations. These are calculated for each occupation-year cell separately according to equation (11). Occupations with zero elasticity of substitution between the two worker types are assigned the relative productivity of college and high school graduates equal to the relative employment, and occupations with infinite substitution elasticity are assigned the relative productivity equal to the college – high school wage premium. While it is difficult to present here all 1620 estimates

(90 occupations in 18 years), point estimates of occupation-specific skill intensity for the years 1983 and 2000 (the first and last year of the sample) are presented in Table 1 in the appendix and visualized in Figure 1 together with the estimated confidence intervals. The full list of this measure is available from the author upon request.

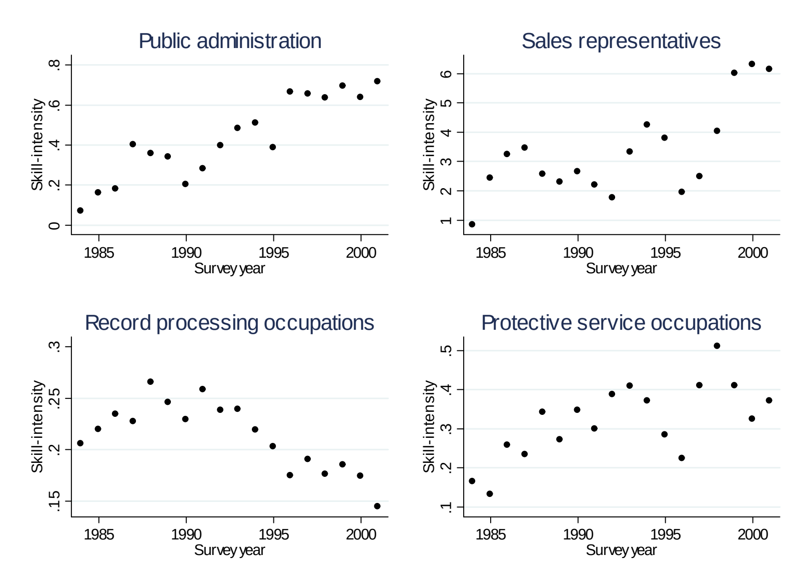
Figure 1: Distribution of skill-intensity across occupations



Note that a great majority of occupations experienced an upgrade in their skill intensity between 1983 and 2000, which is consistent with the skill-biased technological change hypothesis. Nevertheless, some occupations became significantly less skill intensive during the analyzed period. This mainly concerns the occupations involving high precision mechanical tasks, like records processing, laboratory technicians and farm occupations, that originally required high skills but gradually become substituted by machines. There is also a group of occupations in which the relative productivity of college and high school graduates remained constant. The most interesting trends in occupation-specific skill intensities are presented in Figure 2.

Note the extensive increase in skill intensity among public administration officers and sales representatives. These occupations used to be relatively un-intensive in college skills in the mid 1980s but popularization of personal computers increased their

Figure 2: Evolution of log of skill intensity in selected occupations (1983-2000)



Note: Log of skill intensity is defined as $ln(\frac{\alpha_{Cjt}}{\alpha_{Njt}}) = ln(\frac{w_{Cjt}}{w_{Njt}}) - \frac{1}{\sigma_j} ln(\frac{L_{Cjt}}{L_{Njt}})$.

skill requirements. The opposite trend is observed in records processing occupations, which are an example of occupations where computers substituted skilled labor. Finally, the bottom right panel of Figure 2 presents protective service occupations, which were not affected by recent technological progress.

The measure of skill intensity of occupations developed in this study relies on the estimated within-occupation elasticity of substitution between college and high school graduates. An alternative to this approach would be to assume that all occupations are characterized by the same well-defined elasticity of substitution and use it to derive occupation-specific skill intensities. As discussed in Section 2, one of the common practices in the literature is to assume infinite elasticity of substitution between college and high school graduates ($\sigma_j = \infty$). While the assumption of perfect substitutability between the two skill types is questionable, I propose using the estimate of market-wide elasticity of substitution between more and less educated labor instead ($\sigma_j = 1.4$). The estimates of the skill intensity of occupations based on these assumptions are presented in columns 5-6 (for $\sigma_j = 1.4$) and 7-8 (for $\sigma_j = \infty$)

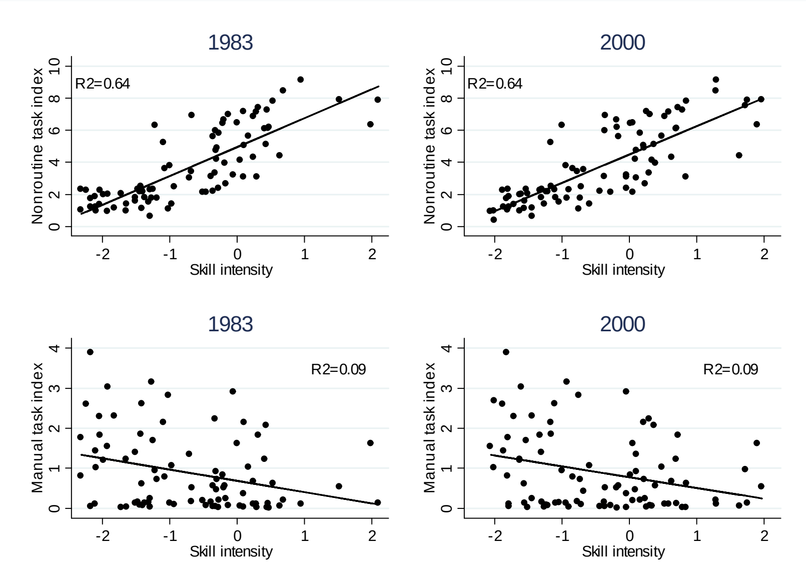
of Table 1 in the appendix.

The correlation between the measure developed in this paper and the one based on perfect substitutability between skill types is 0.319, while the correlation between the new measure and the one based on $\sigma_j = 1.4$ is 0.618. Obviously, the lower the elasticity of substitution between college and high school graduates, the more important the educational structure of the occupation is for the determination of its skill intensity. This is the source of major differences between the three measures presented in Table 1. Note, for example the secretaries, for whom the estimated elasticity of substitution is infinity. If one assumes the substitution elasticity of 1.4, secretaries become much more skill-intensive because few college graduates perform this occupation. In the case of accountants and auditors, the elasticity of substitution is estimated to be 1.03. If one assumes infinite elasticity, their skill intensity significantly drops, which is caused by the large number of college graduates employed in this occupation.

That the proposed measure of skill intensity pictures reality the best is confirmed by its correlation with job characteristics reported by occupation dictionaries. Such a comparison is presented in Figure 3, which plots the non-routine tasks index and manual tasks index, as derived from the DOT, against the skill intensity.

Note that there is a strong positive correlation between the non-routine tasks index and the measure of skill intensity, but hardly any relationship is observed between the manual tasks index and the measure of skill intensity. This is intuitive as the productivity advantage of college graduates should come from their ability to perform non-routine tasks while performance in manual tasks should not depend on level of education. When regressing the non-routine task index on the measure of skill intensity based on infinite elasticity of substitution, one gets the R-squared of 0.22, and when regressing it on the measure based on $\sigma_j = 1.4$, one gets the R-squared of 0.16.

Figure 3: Comparison of the measure of skill intensity with the DOT routine and manual tasks index



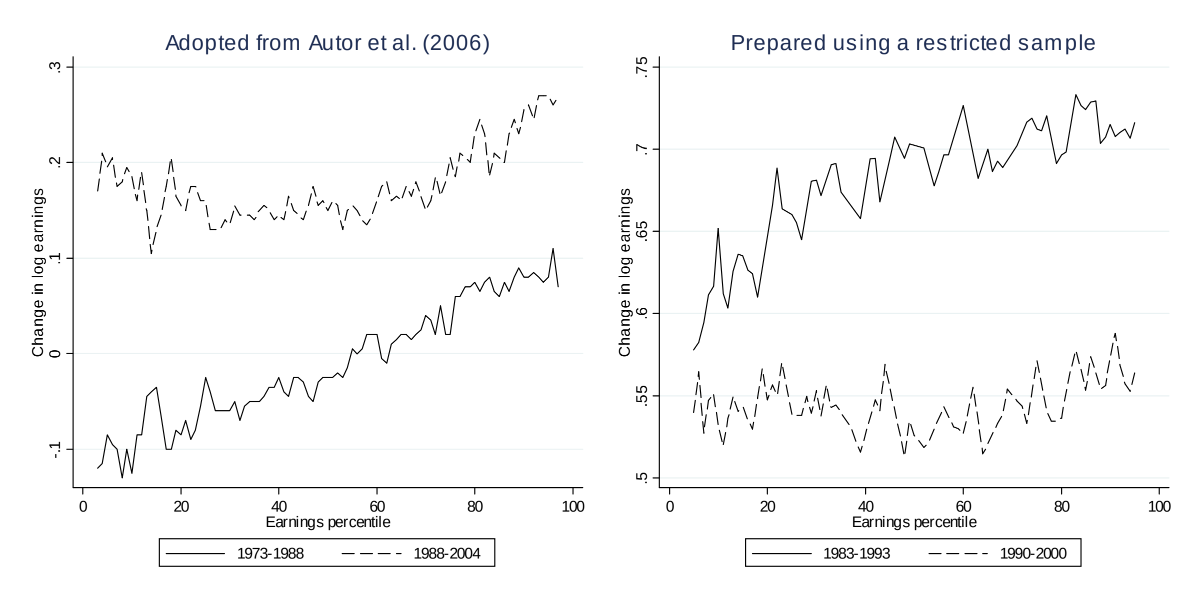
7 Applications of the measure of skill intensity of occupations

The measure of skill intensity of occupations derived in this study can be used, for example, to track the technological progress of individual occupations or derive the demand for educated labor within different types of occupations. This section presents another application: an analysis of the recent polarization of earnings growth in the U.S.

7.1 Polarization of earnings growth

The pattern in earnings growth changes observed in the last decade of the twentieth century, when the wage growth in the bottom and top part of the earnings distribution was faster than in the middle part, known as earnings growth polarization, was documented by Autor et al. (2006). This observation is especially interesting when contrasted with earlier periods when earnings at the low end of the distribution were falling and those at the top end were increasing, which is illustrated in the left panel of Figure 4, adopted from Autor et al. (2006). The same pattern, although with higher growth rates for the whole distribution, is present in the sub-sample of the U.S. labor force investigated in this study, i.e., among college and high school graduates with no more than 10 years of labor market experience, as presented in the right panel of Figure 4.

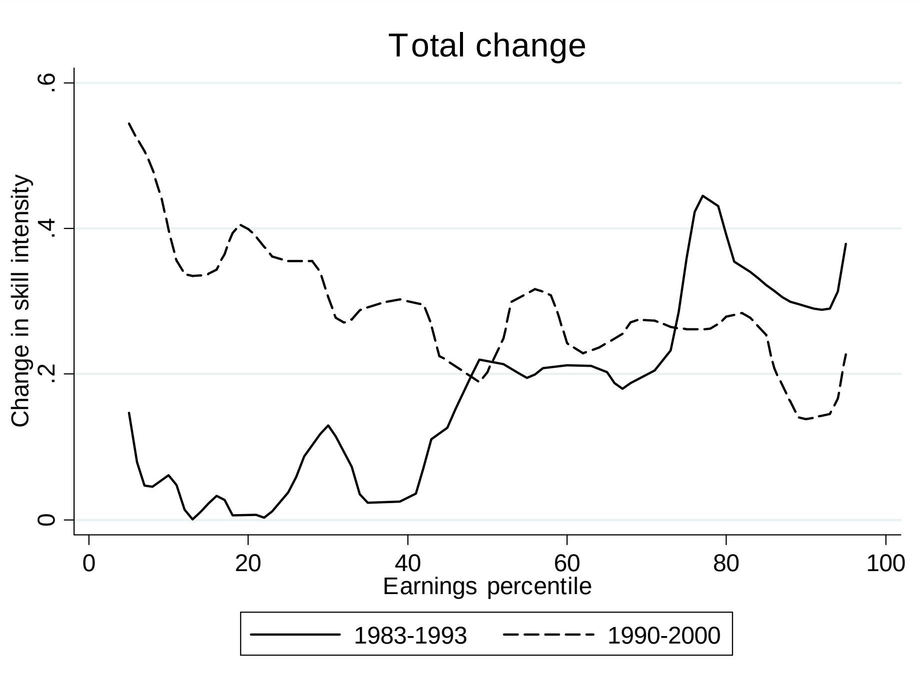
Figure 4: Changes in log earnings by earnings percentile



Note: The left panel is adopted from Autor et al. (2006); the right panel is obtained using the sample of young college and high school graduates described in Section 5.

Recent literature explains the changes in the growth profile documented above by the varying impact that new technologies have on different job tasks (Autor and Dorn, 2009; Firpo et al., 2011; Acemoglu and Autor, 2010). In particular, it is argued that modern technologies complement workers performing non-routine cognitive tasks and substitute for workers performing routine tasks. Assuming that the task content of work is homogeneous within occupations, this statement can be verified using the measure of skill intensity of occupations defined in this chapter. Recall that skill intensity is defined as the within-occupation relative productivity of college and high school graduates and, as such, it measures occupation-specific skill bias. If changes in earnings inequality observed in the last decades of the twentieth century are indeed driven by the heterogeneous impact of technologies on different occupations, plotting changes in the average skill intensity of occupations employing workers from each percentile of the earnings distribution should reveal patterns similar to those in Figure 4.

Figure 5: 1983-1993 and 1990-2000 changes in log occupational skill intensity by earnings percentile



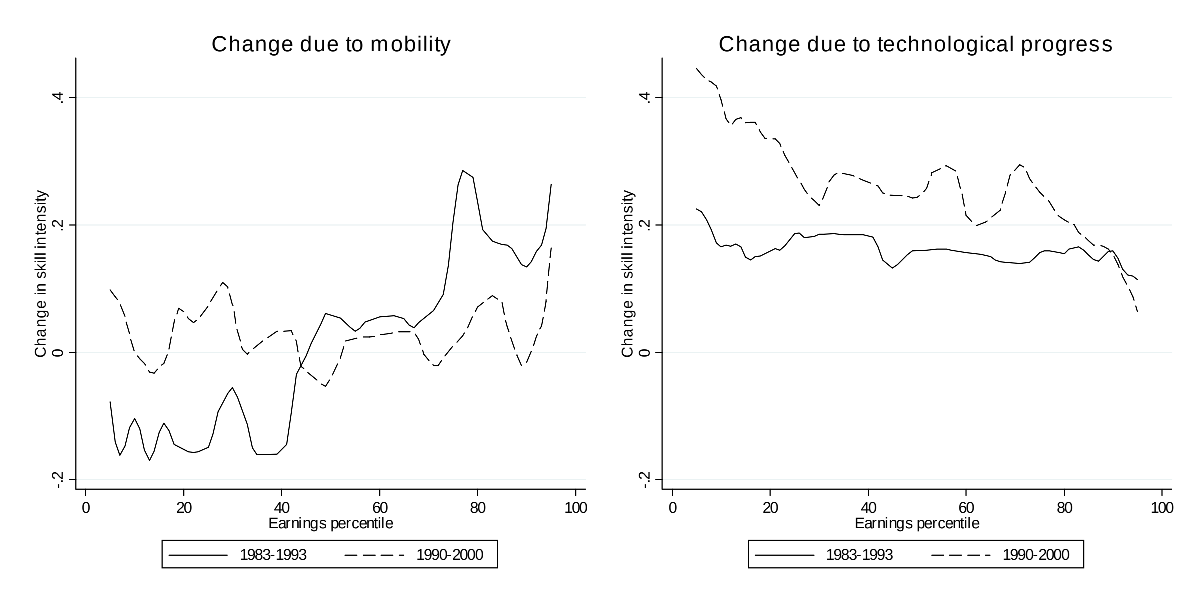
Note: This figure plots total changes in log of average skill intensity of occupations performed by young college and high school graduates from each percentile of the earnings distribution. Log of skill intensity is defined as $ln(\frac{\alpha_{Cjt}}{\alpha_{Njt}}) = ln(\frac{w_{Cjt}}{w_{Njt}}) - \frac{1}{\sigma_j} ln(\frac{L_{Cjt}}{L_{Njt}})$.

In the above figure we observe that between 1983 and 1993 workers from the lower part of the earnings distribution did not experience any significant change in the average skill intensity of the occupations in which they were employed, while workers with above median earnings experienced strong skill-biased technological progress. Interestingly, between 1990 and 2000 the growth in skill intensity of occupations performed by the top 50% of earners was roughly the same as in the previous period, while people with below-median earnings experienced much more spectacular improvement in the skill intensity of occupations in which they were employed. The increase in skill intensity of occupations employing workers from the lower part of the earnings distribution could be in part responsible for the earnings growth polarization.

To better understand the nature of the patterns observed in Figure 5, let me decompose the difference in skill intensity of occupations where workers from each percentile of the earnings distribution are employed into occupation-specific technological progress and mobility (of workers across occupations and of occupations across earnings percentiles). Worker mobility happens when fewer people become employed in certain types of occupations (e.g., low skill-intensive) and more of them find employment in other types of occupations (e.g., high skill-intensive), and thus we observe shifts in the skill intensity of occupations employing the reallocated workers. Occupation-earnings mobility happens when certain types of occupations pay relatively more (or less) than they used to pay (e.g., due to changes in total factor productivity) and thus shift to a different earnings percentile. By fixing the skill intensity of all 90 occupations at their initial level (i.e., at the level from 1983 or 1990, respectively) one can observe the changes in average skill intensity of occupations performed by workers from different percentiles of the earnings distribution which are solely due to mobility. In other words, occupation-specific skill intensities are forced to be constant and thus any changes observed in the distribution of skill intensities over the earnings distribution have to be attributed to workers' and occupation-earnings mobility. Changes due to shifts in occupation-specific skill intensity constitute the difference between the total changes depicted in Figure 5 and changes due to mobility, i.e., this is the residual variation. The resulting decomposition is pictured in the two panels of Figure 6.

When abstracting from occupation-specific technological progress, which is presented in the left panel of Figure 6, the pattern of skill intensity changes observed across the earnings distribution in the 1980s is to a great extent preserved; however, the pattern from the 1990s disappears. We observe that during the 1980s the occupation mix for the bottom 40% of earners shifted towards less skill intensive occupations,

Figure 6: A decomposition of 1983-1993 and 1990-2000 changes in log occupational skill intensity by earnings percentile



Note: The left panel illustrates changes in the log of average skill intensity of occupations due to different composition of occupations performed by workers from each percentile of the earnings distribution; the right panel illustrates changes in average skill intensity due to technological change. Figures were obtained using the sample of young college and high school graduates described in Section 5. The log of skill intensity is defined as $ln(\frac{\alpha_{Cjt}}{\alpha_{Njt}}) = ln(\frac{w_{Cjt}}{w_{Njt}}) - \frac{1}{\sigma_j} ln(\frac{L_{Cjt}}{L_{Njt}})$.

while the occupation mix for the top 30% of earners shifted towards more skill intensive occupations. Interestingly, no changes in the occupation mix were observed in the 1990s, which suggests that workers were neither changing occupations nor were occupations switching places in the earnings distribution (or these two cancelled each other out). On the other hand, when plotting changes in the skill intensity driven purely by occupation-specific technological progress, only the relationship observed in the latter period is mimicked. The right panel of Figure 6 shows that in the 1980s occupations employing workers from all earnings percentiles experienced the same technological progress, on average, while in the 1990s, occupations employing the bottom earners were subject to a much larger increase in their skill intensity than other occupations. This suggests that the differences between the 1983-1993 and 1990-2000 periods can be attributed to the changing nature of the technological progress.

Specifically, in the earlier period across-the-board computerization concurred with strong reallocation of the top earners towards more computerized occupations, as there appeared more work opportunities involving complex tasks (for example, the demand for IT specialists increased). The least earning (and, supposedly, the least skilled) workers moved towards less skill-intensive occupations either because they were substituted by machines or because they did not know how to operate them. In the later stages of computerization these effects were not observed because the young labor force was already prepared to meet new technologies. During this time we observe an above-average increase in the skill intensity of occupations employing the least earning workers, which could be caused by the gradual computerization of simple job tasks.

How to reconcile the above findings with the modified SBTC hypothesis? As argued above, the mobility of workers across occupations documented in the 1980s could be driven by the heterogeneous impact of technologies on different job tasks; and the fast growing skill intensity of occupations employing the least earning workers could be caused by the technological improvement of simple job tasks.

8 Conclusion

In this study I propose a model-based approach for determining the skill intensity of occupations. This measure can be used to track technological progress on the occupational level — a key ingredient of recent theories of labor market polarization. I argue that a good proxy for occupation-specific skill intensity is the relative productivity of college and high school graduates. This parameter of the production function captures the importance of college-gained skills for the tasks performed within a specific occupation.

When proposing a new measure of skill intensity of occupations, I relax the assumption of the elasticity of substitution between college and high school graduates being the same across occupations, but still assume that occupation-specific substitution elasticities do not change over time. Keeping the elasticity constant over time is one of the identifying assumptions of the econometric model used to estimate σ_j. Relaxing this one is a challenge for future research.

When estimating occupation-specific relative productivities, it is important to take into account the elasticity of substitution between college and high school graduates. This parameter in many studies is *ex ante* assumed to be infinite. I estimate the elasticity of substitution between differently educated workers and find that many occupations are characterized by imperfect substitutability between college and high school graduates. Not taking that into account would bias the estimates of relative productivities.

Let me acknowledge the fact that estimating the skill intensity of occupations is a data-hungry process. This limits the application of the methodology developed in this study to economies which have sizeable worker-level data. One alternative solution would be to take advantage of the findings of Kezdi (2003) who shows that the skill bias in Hungary follows global skill-biased changes. Extrapolating these findings would suggest that the occupation-specific relative productivity of college and high school graduates (occupation-specific skill bias) is similar in all open economies. Thus, skill intensities calculated for the U.S. in this study could be, with some care, also applied in other countries. Another alternative involves assuming that all occupations are characterized by the same elsticity of substituion (let us call it σ^*) between college and high school graduates and calculate occupation-specific skill intensities by substituting σ^* to equation (11). The results of this exercise are presented in the last two columns of Table 1.

The proposed measure of skill intensity of occupations has multiple applications. This paper discusses one of them. I show that the measure of skill intensity could be used to analyze the recently observed polarization of earnings growth, as documented by Goos and Manning (2007) for the U.K. and Autor et al. (2006) for the U.S. The presented results are in line with the hypothesis proposed by Autor that the technological change in the 1980s had a positive effect on the high earners, while in the 1990s also the low end of the earnings distribution benefited from it. This paper also brings new evidence about the changing nature of the technological progress.

I show that in the earlier phase the technological progress was equally distributed across occupations from all the earnings distribution, but high earners sorted to more skill-intensive occupations and low earners sorted to less skill-intensive occupations. In the latter phase, there was no further reallocation and the least-paying occupations experienced greater technological progress. The observed reallocation of workers across occupations is in line with Acemoglu and Autor (2010), who argue that technological progress changed the task composition of occupations and thus their demand for skills.

References

[1] Acemoglu D., Autor D. (2010) Skills, tasks and technologies: Implications for employment and earnings, in: Ashenfelter, O., Card, D. (eds.), Handbook of Labor Economics, Vol. 4, Amsterdam.

[2] Autor D., Dorn D. (2009) This job is "getting old": Measuring changes in job opportunities using occupational age structure, American Economic Review Papers and Proceedings, 99(2): 45-51.

[3] Autor D., Katz L., Kearney M. (2006) The polarization of the U.S. labor market, American Economic Review, 96(2): 189–194.

[4] Autor D., Katz L., Kearney M. (2008) Trends in U.S. wage inequality: Revising the revisionists, The Review of Economics and Statistics, 90(2): 300–323.

[5] Autor D., Levy, F., Murnane R. (2003) The skill content of recent technological change: An empirical exploration, Quarterly Journal of Economics, 118(4): 1279-1334.

[6] Bound, J., Johnson, G. (1992) Changes in the structure of wages in the 1980's: An evaluation of alternative explanations, American Economic Review, 82: 371-92.

[7] Card D., 2001. Immigrant inflows, native outflows and the local labor market impacts of higher immigration, Journal of Labor Economics, 19(1): 22-64.

[8] Card D., DiNardo J. (2002) Skill-biased technological change and wage inequality: Some problems and puzzles, Journal of Labor Economics, 20(4): 733-783.

[9] Card D., Lemieux T. (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.

[10] Ciccone A., Peri G. (2005) Long-run substitutability between more and less educated workers: Evidence from U.S. states, 1950–1990, The Review of Economics and Statistics, 87(4): 652 - 663.

[11] Firpo S., Fortin N., Lemieux T. (2011) Occupational tasks and changes in the wage structure, IZA Discussion Paper No. 5542.

[12] Goos M., Manning A. (2007) Lousy and lovely jobs: The rising polarization of work in Britain, The Review of Economics and Statistics, 89(1):118-133

[13] Gottschalk, P., Hansen, M. (2003) Is the proportion of college workers in noncollege jobs increasing? Journal of Labor Economics, 21(2): 449-471.

[14] Griliches Z., Mason W.M. (1972) Education, income, and ability, The Journal of Political Economy, 80(3): S74-S103.

[15] Jeager D. (1997) Reconciling the old and new Census Bureau education questions: Recommendations for researchers, Journal of Business & Economic Statistics 15(3): 300-309.

[16] Katz L., Murphy K. (1992) Changes in relative wages, 1963-1987: Supply and demand factors, The Quarterly Journal of Economics, 107(1): 35-78.

[17] Kezdi G. (2003) Imported technology and increasing demand for skill: The case of Hungary, mimeo, Central European University.

[18] Larrimore J., Burkhauser R., Feng S., Zayatz L. (2008) Consistent cell means for topcoded incomes in the public use March CPS (1976-2007), Journal of Economic and Social Measurement, 33(2-3): 89-128.

[19] McGoldrick, K.M , Robst, J. (1996) Gender differences in overeducation: A test of the theory of differential overqualification, American Economic Review, 86(2): 280-84.

[20] Meyer P., Osborne A. (2005) Proposed category system for 1960-2000 Census occupations. U.S., Bureau of Labor Statistics Working Paper No. 383.

[21] Pryor F., Schaffer D. (1997) Wages and the university educated: A paradox resolved, Monthly Labor Review, 120(7): 3-14.

Table 1: Estimates of occupation-specific elasticities of substitution between college and high school graduates and the imputed relative productivities

Occupation group	$\widehat{\sigma}_j$	Estimated σ_j		Assumed $\sigma_j = 1.4$		Assumed $\sigma_j = \infty$	
		$ln\widehat{\frac{\alpha_{Cjt}}{\alpha_{Njt}}}_{1983}$	$ln\widehat{\frac{\alpha_{Cjt}}{\alpha_{Njt}}}_{2000}$	$ln\widehat{\frac{\alpha_{Cjt}}{\alpha_{Njt}}}_{1983}$	$ln\widehat{\frac{\alpha_{Cjt}}{\alpha_{Njt}}}_{2000}$	$ln\widehat{\frac{\alpha_{Cjt}}{\alpha_{Njt}}}_{1983}$	$ln\widehat{\frac{\alpha_{Cjt}}{\alpha_{Njt}}}_{2000}$
Sales representatives, commodities except retail	∞	0.324	0.575	0.534	0.690	0.324	0.575
Sales-related occupations	∞	0.326	0.492	1.187	1.439	0.326	0.492
Miscellaneous management-related occupations	∞	0.296	0.481	0.540	0.696	0.296	0.481
Securities and financial services sales occupations	∞	0.158	0.447	0.136	0.300	0.158	0.447
Real estate managers	∞	0.495	0.426	0.979	0.783	0.495	0.426
Supervisors and proprietors, sales occupations	∞	0.282	0.400	0.724	0.850	0.282	0.400
Supervisors, production occupations	∞	0.278	0.377	1.228	1.480	0.278	0.377
Nursing aides	∞	0.321	0.352	0.036	2.771	0.321	0.352
Painters, sculptors, and photographers	∞	0.054	0.343	0.095	0.454	0.054	0.343
Extractive and precision production occupations	∞	0.265	0.316	0.042	2.188	0.265	0.316
Engineers, n.e.c.	∞	0.264	0.310	0.201	0.180	0.264	0.310
Managers, marketing and advertising	∞	0.415	0.300	0.492	0.238	0.415	0.300
Miscellaneous financial officers	∞	0.250	0.299	0.364	0.295	0.250	0.299
Other mechanics and repairers	∞	0.052	0.293	0.412	1.466	0.052	0.293
Miscellaneous professional specialty occupations	∞	0.260	0.284	0.222	0.187	0.260	0.284

Occupation group	$\widehat{\sigma_j}$	Estimated σ_j		Assumed $\sigma_j = 1.4$		Assumed $\sigma_j = \infty$	
		$ln\frac{\widehat{\alpha_{Cjt}}}{\alpha_{Njt}}_{1983}$	$ln\frac{\widehat{\alpha_{Cjt}}}{\alpha_{Njt}}_{2000}$	$ln\frac{\widehat{\alpha_{Cjt}}}{\alpha_{Njt}}_{1983}$	$ln\frac{\widehat{\alpha_{Cjt}}}{\alpha_{Njt}}_{2000}$	$ln\frac{\widehat{\alpha_{Cjt}}}{\alpha_{Njt}}_{1983}$	$ln\frac{\widehat{\alpha_{Cjt}}}{\alpha_{Njt}}_{2000}$
Material recording, scheduling, & distr. clerks	∞	0.249	0.281	1.376	1.353	0.249	0.281
Science technicians	∞	0.299	0.227	0.000	0.846	0.299	0.227
Financial managers	∞	0.351	0.224	0.450	0.215	0.351	0.224
Engineering technologists and technicians	∞	0.102	0.209	0.325	0.593	0.102	0.209
Stenographers and typists	∞	0.140	0.184	0.823	0.772	0.140	0.184
General office clerks	∞	0.174	0.181	0.833	0.724	0.174	0.181
Administrative support occupations	∞	0.194	0.178	0.917	0.629	0.194	0.178
Public administration	∞	0.069	0.113	0.104	0.148	0.069	0.113
Secretaries	∞	0.087	0.112	0.522	0.563	0.087	0.112
Information clerks	∞	0.092	0.102	0.558	0.559	0.092	0.102
Farm occupations	∞	0.475	0.084	2.264	0.292	0.475	0.084
Median	∞	0.262	0.296	0.507	0.660	0.262	0.296

Occupation group	$\widehat{\sigma_j}$	Estimated σ_j		Assumed $\sigma_j = 1.4$		Assumed $\sigma_j = \infty$	
		$ln\widehat{\frac{\alpha_{Cjt}}{\alpha_{Njt}}}_{1983}$	$ln\widehat{\frac{\alpha_{Cjt}}{\alpha_{Njt}}}_{2000}$	$ln\widehat{\frac{\alpha_{Cjt}}{\alpha_{Njt}}}_{1983}$	$ln\widehat{\frac{\alpha_{Cjt}}{\alpha_{Njt}}}_{2000}$	$ln\widehat{\frac{\alpha_{Cjt}}{\alpha_{Njt}}}_{1983}$	$ln\widehat{\frac{\alpha_{Cjt}}{\alpha_{Njt}}}_{2000}$
Sales workers, retail	9.72	0.300	0.391	1.408	1.291	0.315	0.415
Writers, artists, and related workers	7.21	0.272	-0.084	0.299	-0.079	0.226	-0.083
Service occupations, n.e.c.	6.79	0.335	0.184	1.239	0.430	0.221	0.105
Records processing occupations, except financial	6.33	0.231	0.333	0.920	0.908	0.203	0.251
Financial records processing occupations	5.34	0.268	0.100	0.706	0.363	0.146	0.097
Carpenters, electricians, and painters	5.30	-0.007	0.351	0.044	0.411	-0.036	0.059
Sales occupations, advertising & other services	5.10	0.425	0.720	0.627	0.918	0.370	0.625
Construction trades, n.e.c.	4.91	0.081	0.889	0.047	2.657	0.230	0.371
Miscellaneous managers and administrators	4.50	0.451	0.474	0.523	0.561	0.311	0.388
Public relations specialists, announcers	4.45	0.242	0.207	0.260	0.207	0.230	0.208
Designers	4.45	0.255	0.333	0.362	0.359	0.216	0.295
Health technologists and technicians	3.61	0.429	0.504	1.024	0.945	0.202	0.235
Personnel, training, and labor relations specialists	3.31	0.054	0.275	0.488	0.420	0.283	0.351
Computer equipment operators	3.18	0.289	0.856	0.408	1.267	0.102	0.362

		Estimated σ_j		Assumed $\sigma_j = 1.4$		Assumed $\sigma_j = \infty$	
Occupation group	$\widehat{\sigma}_j$	$ln\widehat{\frac{\alpha_{Cjt}}{\alpha_{Njt}}}_{1983}$	$ln\widehat{\frac{\alpha_{Cjt}}{\alpha_{Njt}}}_{2000}$	$ln\widehat{\frac{\alpha_{Cjt}}{\alpha_{Njt}}}_{1983}$	$ln\widehat{\frac{\alpha_{Cjt}}{\alpha_{Njt}}}_{2000}$	$ln\widehat{\frac{\alpha_{Cjt}}{\alpha_{Njt}}}_{1983}$	$ln\widehat{\frac{\alpha_{Cjt}}{\alpha_{Njt}}}_{2000}$
Pre-kindergarten and kindergarten teachers	2.94	0.250	0.653	0.224	0.689	0.267	0.566
Clinical laboratory technologists and technicians	2.81	0.278	0.403	0.299	0.490	0.256	0.338
Cooks	2.79	0.101	1.877	0.039	2.165	0.186	0.277
Computer programmers	2.59	0.269	0.168	0.286	0.153	0.238	0.180
Fabricators and assemblers, production occs.	2.16	0.087	1.091	0.033	1.614	0.110	0.215
Real estate sales occupations	1.69	0.155	0.195	0.187	0.240	0.107	0.157
Supervisors, administrative support occupations	1.64	1.011	0.428	1.177	0.433	0.403	0.170
Insurance adjusters, examiners, & investigators	1.62	0.417	0.396	0.473	0.407	0.258	0.213
Insurance sales occupations	1.38	0.478	0.641	0.404	0.601	0.242	0.392
Accountants and auditors	1.03	0.120	0.182	0.217	0.244	0.263	0.372
Child-care workers	0.92	1.038	1.277	0.716	0.924	0.168	0.183
Purchasing agents and buyers	0.58	1.182	1.364	0.580	0.740	0.310	0.360
Legal assistants	0.50	1.302	1.357	0.224	0.313	0.134	0.181
Median	2.72	0.272	0.403	0.404	0.490	0.230	0.251

Occupation group	$\widehat{\sigma}_j$	Estimated σ_j		Assumed $\sigma_j = 1.4$		Assumed $\sigma_j = \infty$	
		$ln\widehat{\frac{\alpha_{Cjt}}{\alpha_{Njt}}}_{1983}$	$ln\widehat{\frac{\alpha_{Cjt}}{\alpha_{Njt}}}_{2000}$	$ln\widehat{\frac{\alpha_{Cjt}}{\alpha_{Njt}}}_{1983}$	$ln\widehat{\frac{\alpha_{Cjt}}{\alpha_{Njt}}}_{2000}$	$ln\widehat{\frac{\alpha_{Cjt}}{\alpha_{Njt}}}_{1983}$	$ln\widehat{\frac{\alpha_{Cjt}}{\alpha_{Njt}}}_{2000}$
Editors and reporters	0	0.741	1.035	0.065	0.431	0.110	0.907
Social workers	0	0.487	0.735	0.082	0.212	0.116	0.315
Electrical and electronic engineers	0	0.680	0.712	0.172	0.216	0.195	0.323
Therapists, n.e.c.	0	0.574	0.704	0.364	0.166	0.316	0.217
Recreation and religious workers	0	0.709	0.640	-0.011	0.348	-0.008	0.318
Registered nurses	0	0.431	0.585	0.094	0.142	0.064	0.162
Technicians, n.e.c.	0	0.450	0.481	0.005	0.329	0.003	0.278
Counselors, librarians, archivists, and curators	0	0.006	0.364	0.243	0.289	0.244	0.375
Health assessment and treating occupations	0	0.731	0.321	0.387	0.988	0.415	0.502
Police and detectives	0	0.189	0.231	0.771	0.459	0.223	0.182
Drafting occupations & surveying and mapping	0	0.222	0.191	0.423	0.589	0.136	0.192
Miscellaneous adjusters and investigators	0	0.217	0.189	0.560	1.207	0.185	0.398
Protective service occupations	0	0.155	0.357	0.682	1.248	0.163	0.350

Occupation group	$\widehat{\sigma}_j$	Estimated σ_j		Assumed $\sigma_j = 1.4$		Assumed $\sigma_j = \infty$	
		$ln\widehat{\frac{\alpha_{Cjt}}{\alpha_{Njt}}}_{1983}$	$ln\widehat{\frac{\alpha_{Cjt}}{\alpha_{Njt}}}_{2000}$	$ln\widehat{\frac{\alpha_{Cjt}}{\alpha_{Njt}}}_{1983}$	$ln\widehat{\frac{\alpha_{Cjt}}{\alpha_{Njt}}}_{2000}$	$ln\widehat{\frac{\alpha_{Cjt}}{\alpha_{Njt}}}_{1983}$	$ln\widehat{\frac{\alpha_{Cjt}}{\alpha_{Njt}}}_{2000}$
Agricultural, forestry, fishing, and hunting	0	0.114	0.121	1.064	1.522	0.211	0.354
Dental assistants and health aides	0	0.152	0.085	1.195	1.332	0.266	0.231
Machine operators	0	-3.314	-2.957	2.042	1.282	0.286	0.105
Handlers and laborers	0	-3.346	-2.899	0.034	1.697	0.300	0.214
Median	0	0.222	0.357	0.364	0.459	0.195	0.315

Note: The 2nd and 3rd columns of this table present the estimated elasticity of substitution between college and high school graduates followed by its standard error calculated using the delta method. For occupations with zero or infinite elasticity of substitution, the standard errors are unavailable because the substitution elasticity is inferred from the observed properties of occupations rather than estimated from the data. Columns 4 and 5 present logs of the estimated relative productivities of college and high school graduates in years 1983 and 2000, and columns 6 and 7 present logs of the estimated relative productivities when the elasticity of substitution is assumed to be 1.4 for all occupations.

CHAPTER 3
Job Market Polarization and Employment Protection in Europe

Barbara Pertold-Gębicka*

CERGE-EI†

Abstract

Although much attention has been paid to the polarization of national labor markets, with employment and wage growth occurring in both low- and high- but not middle-skilled occupations, there is little consistent evidence on cross-country differences in this process. I analyze job polarization in 12 European countries using an occupational skill-intensity measure, which is independent of country-specific labor supply conditions. Extensive north-south differences in the extent and skewness of polarization correspond to variation in economic growth and to dissimilarities in employment protection.

*I would like to thank Alena Bičáková, Randall Filer, Štěpán Jurajda, and Filip Pertold for valuable comments and suggestions. E-mail: barbara.gebicka@cerge-ei.cz

†A joint workplace of the Center for Economic Research and Graduate Education, Charles University, Prague, and the Economics Institute of the Academy of Sciences of the Czech Republic. Address: CERGE-EI, Politickych veznu 7, Prague 11121, Czech Republic.

1 Introduction

Polarization of the labor market, defined as employment and wage growth in low- and high-skilled occupations at the cost of middle-skilled occupations, was first documented by Goos and Manning (2007) in the U.K.[1] Further analyses of the British and American labor markets confirm this trend and suggest some explanations of its causes. Autor et al. (2006) propose that labor market polarization observed since the 1990s can be accounted for by "routinization", i.e., the substitution of routine job tasks by modern technologies.[2] Firpo et al. (2011) suggest that offshoring certain job tasks to low-wage countries can also be partially responsible for polarization in the U.S. Finally, Acemoglu and Autor (2010) note that the allocation of workers to occupational tasks might be influenced by labor market imperfections and institutions, thus challenging the polarization pattern in some countries.

This has raised the question of whether labor market polarization is unique within the Anglo-Saxon countries, among which the U.S. is known as the pioneer in technological progress and the largest outsourcer of manufacturing and remote consumer service jobs. In answer to this question, recent research suggests that polarization can be observed across the majority of developed economies. For example, studies by Spitz-Oener (2006) and Dustmann et al. (2009) show that polarization is present in another leading economy, Germany. Most importantly, Goos et al. (2009) provide evidence of this phenomenon across 16 European countries.[3]

Nevertheless, the international analysis of labor market polarization is not complete. First, the European evidence is based on a crude measure of the skill requirements of occupations – the average wage. As argued in Chapter 2 of this dissertation, this approach implicitly assumes that within occupations differently skilled workers are perfect substitutes, which is likely not to be the case. Second, cross-country

[1] Goos and Manning first used the term "polarization" to describe employment growth in low- and high-skilled occupations at the cost of middle-skilled occupations in the 2003 Working Paper version of this publication.

[2] The term "routinization" was introduced by Autor et al. (2003).

[3] These countries are: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Italy, Ireland, Luxembourg, the Netherlands, Norway, Portugal, Spain, Sweden and the U.K.

differences in the shape of employment change distribution (which is used to picture polarization), while documented, have not been given much attention. These differences might be caused by cross-country heterogeneity in the supply of skills, variation in economic cycles, or distinct labor market legislations. The interaction of the latter with technological progress on the occupation level has been recognized by Acemoglu and Autor (2010) in their chapter of the recent Handbook of Labor Economics as a fruitful area for further research. Finally, while in the U.S. polarization has been measured in employment changes as well as in earnings changes, the existing international analysis focuses only on employment changes, i.e., it documents so-called job polarization as opposed to wage polarization.[4] Studying wage polarization would give additional insight into the structure of the European labor market.

This paper addresses the first two issues. I use the European Union Labor Force Survey (EULFS) to report differences in the extent of job polarization across European countries, adopting the measure of skill requirements of occupations developed by Pertold-Gebicka (2010). This is a preferable measure to document polarization across countries, as it is independent of supply conditions in local labor markets. The discussion and examples provided in the current study confirm this statement. With the use of the skill requirements measure, I provide extensive evidence on cross-country differences in the extent of polarization. Specifically, one can observe that polarization is the strongest in Southern European countries and Ireland, while it is somewhat weaker in Northern Europe. As a potential explanation for this observation, I suggest differences in economic growth and educational attainment of their populations. The remaining cross-country variation in the extent of polarization is shown to be partially driven by dissimilarities in labor market institutions. This latter finding suggests that strong employment protection might impede or slow down the market mechanisms observed in non-regulated countries, such as substitution

[4] Wage polarization is known as the pattern of the earnings growth in the bottom and top percentiles of earnings distribution with a simultaneous decrease of earnings in the middle of the earnings distribution. Job polarization is known as growth of employment in high- and low-skilled occupations with simultaneous decrease (or stagnation) of employment in middle-skilled occupations. See Acemoglu and Autor (2010) for a summary of the terminology used in the polarization literature.

of certain job tasks by computers (Acemoglu and Autor, 2010), which flattens the polarization patterns.

The rest of this chapter is organized as follows: Section 2 describes the skill intensity measure used to order occupations according to their skill requirements. Section 3 describes the data used in this analysis and Section 4 presents some evidence on the incidence of labor market polarization in Europe and compares it to the results obtained using alternative measures of occupational skill requirements. Section 5 discusses cross-country differences in the extent of polarization and proposes an explanation for this observation. Finally, conclusions are presented in Section 6.

2 The measure of skill requirements of occupations

The term job polarization is used in the literature to indicate growth of employment in high- and low-skilled occupations with a simultaneous decrease (or stagnation) of employment in middle-skilled occupations (Goos and Manning, 2007). Thus, the key ingredient of any analysis of labor market polarization is a measure of the skill requirements of occupations.

The recent literature uses several alternative measures of the skill requirements of occupations. The most often encountered are the average educational achievement of workers (Autor et al., 2006, for the U.S.; Goos and Manning, 2007, for the U.K.) and the average wage (Firpo et al., 2011; Goos et al., 2009), although both approaches are based on implicit assumptions that are likely to be violated. For the employment structure of occupations to correctly reflect their skill requirements, we need to face zero within-occupation substitutability between workers of different skills. On the other hand, wages are good predictors of occupational skill requirements when differently skilled workers are perfect substitutes. With imperfect substitutability between skill types, occupation-specific employment structures are driven not only by skill requirements (i.e., the demand for skills) but also by the supply of differently skilled workers. In this case wages are the equilibrium outcome of the interaction between these two forces. Thus, neither wages nor employment can be used to identify occupational skill requirements.

To deal with this lack of identification, I use the measure of skill requirements of occupations (called the skill intensity of occupations) developed in my earlier paper (Pertold-Gebicka, 2010). This alternative measure is based on estimating the relative productivity of more and less skilled workers employed within each occupation. Thus, it measures how crucial workers' skills are for the tasks performed within a specific occupation. I propose that each occupation uses a relatively general labor-aggregating technology of the constant elasticity of substitution (CES):

$$Y_j = \left(\alpha_{Hj} L_{Hj}^{\gamma_j} + \alpha_{Lj} L_{Lj}^{\gamma j}\right)^{\frac{1}{\gamma j}} \tag{1}$$

where Y_j is the output of occupation j, L_{Hj} is the amount of high-skilled labor, L_{Lj} is the amount of low-skilled labor employed in occupation j, and γ_j is a parameter describing substitutability between these two labor types (the elasticity of substitution is $\sigma_j = \frac{1}{1-\gamma_j}$). In this context, $\frac{\alpha_{H_j}}{\alpha_{L_j}}$ describes the occupation-specific relative productivity of differently skilled workers.

Under perfect competition, occupation-specific employment (L_{Hj} and L_{Lj}) and equilibrium wages (w_{Hj} and w_{Lj}) have to satisfy

$$\frac{\alpha_{Hj}}{\alpha_{Lj}} = \frac{w_{Hj}}{w_{Lj}} \left(\frac{L_{Hj}}{L_{Lj}}\right)^{1-\gamma_j} = \frac{w_{Hj}}{w_{Lj}} \left(\frac{L_{Hj}}{L_{Lj}}\right)^{-\frac{1}{\sigma_j}}.$$

Thus, in the setup where more- and less-skilled workers are imperfect substitutes (i.e., where $0 < \sigma_j < \infty$), it is necessary to combine the relative employment of differently skilled workers (the average educational attainment), relative wages, and the elasticity of substitution between more and less skilled workers to determine occupation-specific relative productivity. The measure of the skill intensity of occupations proposed in Pertold-Gebicka (2010) incorporates all of these ingredients.

Occupation-specific average wages and employment can be easily retrieved from worker-level data, e.g., the European Community Household Panel (ECHP) or EU-LFS; however, the substitution elasticities need to be carefully estimated. In my earlier paper, I propose a strategy for estimating this parameter by employing data on individual workers, and estimate occupation-specific elasticities of substitution between college and high school educated workers using the U.S. Current Population Survey (CPS) data. As this is a very data-hungry process, it is not possible to replicate the estimations using the EULFS data, which does not provide information about individual workers' earnings, or using the ECHP, which is too small to allow estimations on the occupational level. However, assuming that the characteristics of occupations in the U.S. and Europe are similar, I can match the estimates obtained for the U.S. with European data. With today's extent of globalization and spillover of technologies between the U.S. and Europe, the above-made assumption is realistic. Nevertheless, in future work I plan to test it using the Danish register data to estimate the elasticities of substitution between college and high school educated workers in Denmark.

3 Data

Throughout this paper I use the 1993-2001 waves of the EU LFS microdata for scientific purposes. This is a collection of harmonized labor force surveys conducted at national levels in all EU member states and associated countries. The availability of this dataset for all European economies, its comparability across countries and over time, and its representativeness at the 2-digit occupation level makes it the most applicable for this study.

The chosen time span corresponds to the time period when polarization has been documented (Goos and Manning, 2007) and to the availability of the skill intensity measure. Due to the limited time consistency of the occupational coding in the US CPS data, I could only estimate occupation-specific elasticities of substitution between more and less educated labor for the 1983-2001 period.

Given the limitations in data availability for some countries, this study investigates 12 Western European economies: Denmark, Finland, Greece, Ireland, Iceland, Luxembourg, Netherlands, Norway, Portugal, Spain, Sweden and the United Kingdom. Central and Eastern European countries are not analyzed, because the assumption about the similarity of occupations' characteristics between these countries and the U.S. is likely to fail. This study covers the whole working population of the above-mentioned 12 countries.

In the anonymized version of the EULFS, occupations are coded using 2-digit ISCO codes, while the estimates of the substitution elasticities (and thus the estimates of occupation-specific skill intensities) are available at the 3-digit level of the U.S. Census occupational classification. To merge the EULFS data with the U.S. occupational characteristics, 3-digit occupations from both datasets are matched according to an algorithm based on Elliott and Gerova (2005)[5] and skill intensities are averaged at the 2-digit level. This procedure leaves me with 20 occupations listed in Table 1. Throughout the paper occupation-specific employment is measured as the usual weekly man hours worked. For countries with shorter time spans,[6] man hours worked in each of the 20 2-digit occupations were extrapolated on the basis of average annual growth rates in occupation-specific employment.

4 Job polarization in Europe

Job polarization across European countries was first documented by Goos et al. (2009). These autors report changes in employment share for 21 ISCO occupations ranked according to their 1993 mean European wage. Employment structure is calculated by Goos et al. (2009) using the EULFS, and data for 1993 mean wages come from the ECHP dataset. I complement this study by providing evidence on job polarization using the skill intensity measure introduced in Chapter 2.

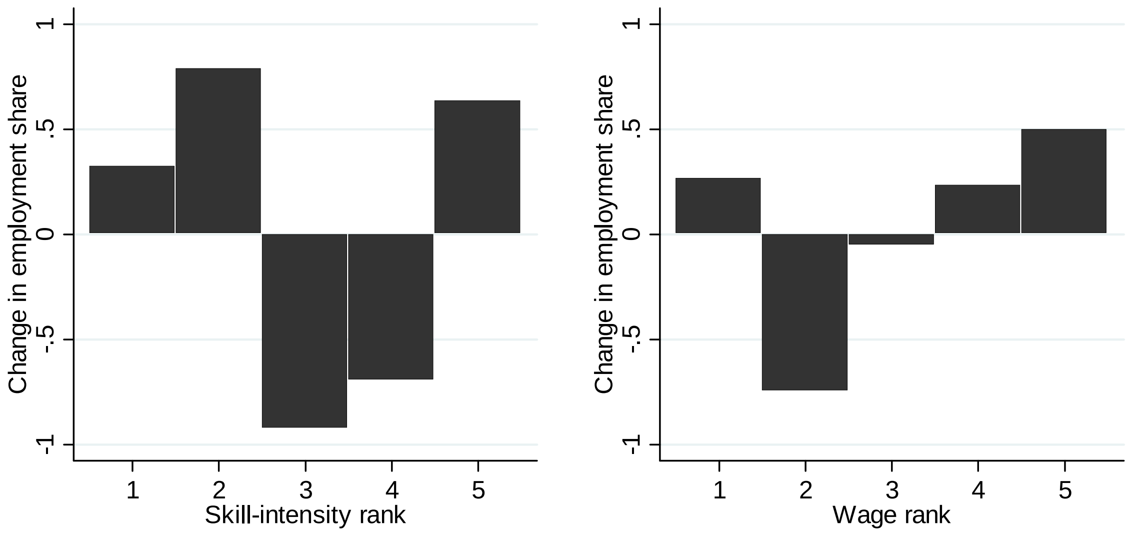
Figure 1 depicts job polarization in Europe with high- and low-skilled occupations

[5] Elliott and Gerova (2005) propose a crosswalk between the 2000 census occupational classification and 88-ISCO, while the skill-intensity measure is available for the 1990 census occupational classification.

[6] Finland has data available from 1997, Iceland from 1995, Norway from 1996, Sweden from 1997, and Spain till 2000.

experiencing a significant employment expansion and middle-skilled occupations experiencing a decrease or stagnation of employment between 1993 and 2001 using the pooled data on occupation-specific employment across 12 European countries. The two panels of this figure use different measures of the skill requirements of occupation: the left panel employs skill intensity proposed by Pertold-Gebicka (2010) and the right panel employs average wage, as proposed by Goos et al. (2009).[7]

Figure 1: Changes in employment share in Europe between 1993-2001 by occupational skill intensity and wage rank



Note: Both graphs were obtained using the EULFS data. For countries with shorter time spans (Finland, Iceland, Norway, Spain and Sweden), man hours worked were imputed on the basis of average annual growth rates. Skill intensity rank corresponds to the position of each occupation in the skill intensity distribution (5 = the most skilled); the wage rank corresponds to the position of each occupation in the U.S. wage distribution (5 = the highest wage).

Polarization is present in both graphs, although there are significant differences between the two. First, the location of the minimum is skewed towards occupations

[7] Although Goos et al. (2009) use the 1993 mean European wage, here occupations are ordered according to the 1993 average US wage to ensure consistency with ordering according to the skill intensity. Nevertheless, there are only minor differences between ordering of occupations according to the US and European average wage.

higher in the skill requirement distribution when using the skill intensity measure; and second, the variation in employment share changes across occupations from different skill requirement ranks is significantly lower when the average wage is used. These differences are driven by the characteristics of the two measures used to capture the skill requirements of occupations. Specifically, with the growing supply of skilled labor, which is well-documented in Europe, some highly-skilled occupations might pay relatively low wages and thus can be classified as less dependent on skills than they actually are. This would be the case for the most "popular" occupations (i.e., the ones with a high supply of skilled workers) characterized by decreasing marginal productivity of highly-skilled labor. Note that the CES occupation-specific production function introduced in Section 2 captures this behavior.

Table 1: Comparison of occupational ranking using the 1993 skill intensity and 1993 average wage measures

Skill intensity rank	Wage rank	Occupation
1	3	Life science and health professionals
2	2	Physical, mathematical and engineering science professionals
3	8	Life science and health associate professionals
4	1	Corporate managers
5	4	Other professionals
6	5	Managers of small enterprises
7	6	Physical and engineering science associate professionals
8	18	Models, salespersons and demonstrators
9	13	Customer service clerks
10	7	Other associate professionals
11	12	Office clerks
12	17	Personal and protective services workers
13	14	Extraction, shot firers, stone cutters and carvers
14	11	Metal, machinery and related workers
15	10	Precision, handcraft, craft printing and related trades workers
16	16	Other craft and related trades workers
17	9	Stationary plant and related operators
18	15	Machine operators and assemblers
19	19	Laborers in mining, construction, manufacturing and transport
20	20	Sales and services elementary occupations

Note: The skill intensity rank has been computed by the author; the wage rank is adapted from Goos et al. (2009).

To better understand the differences between the two measures of skill requirements of occupations, Table 1 documents the ranking of 2-digit ISCO occupations using skill intensity and average wage. If occupations were characterized by perfect or zero substitutability between more and less skilled workers, these two measures of skill requirements would give the same ranking of occupations.

Note that there are substantial differences in ranking of occupations prepared according to the two alternative skill requirement measures. These concern occupations such as corporate managers, which in 1993 paid higher wages than professional occupations because of the short supply of workers educated in management; or sales and service occupations, which paid relatively low wages due to the high supply of potential workers.

Table 2: Changes in employment share over 1993-2001 for low-skilled, middle-skilled, and high-skilled occupations

	Skill intensity			Wage		
	Low-skill	Mid-skill	High-skill	Low-wage	Mid-wage	High-wage
EU average	2.04	-3.55	4.78	1.58	-7.77	6.19
Denmark	4.45	-0.60	0.18	-0.96	-7.16	8.13
Spain	4.18	-6.10	4.73	0.96	-7.04	6.07
Finland	3.07	-2.33	-1.48	6.66	-6.54	-0.12
France	-5.46	-12.55	19.07	-0.74	-12.07	12.81
Greece	3.02	-14.38	12.62	1.75	-6.08	4.34
Ireland	7.29	-8.08	-0.14	6.19	-5.47	-0.72
Luxembourg	-1.90	-8.24	0.74	-1.66	-8.45	10.1
Netherlands	-0.23	-4.62	3.02	2.27	-4.68	2.41
Norway	6.51	-4.84	-1.85	4.96	-6.52	1.57
Portugal	3.64	-3.31	-0.09	2.39	-1.13	-1.26
Sweden	3.57	-2.32	2.25	1.9	-6.93	5.03
UK	2.91	-6.05	5.04	5.77	-10.32	4.55

Note: Classification of occupations into low-, middle-, and high-skilled using the skill intensity measure was done by the author; classification using the average wage is adopted from Goos et al. (2009).

As major differences in the alternative ranking of occupations appear in the middle of the skill requirement distribution, higher aggregation of rank should lead to more similar patterns across the two measures. Indeed, Table 2 shows that once occupations are classified into three (as opposed to five) groups according to their skill requirements, the patterns revealed by both measures are similar.

Table 2 also shows that job polarization is present in all analyzed European economies, although its extent varies significantly across these economies. The next section provides further evidence and suggests some explanations for this finding.

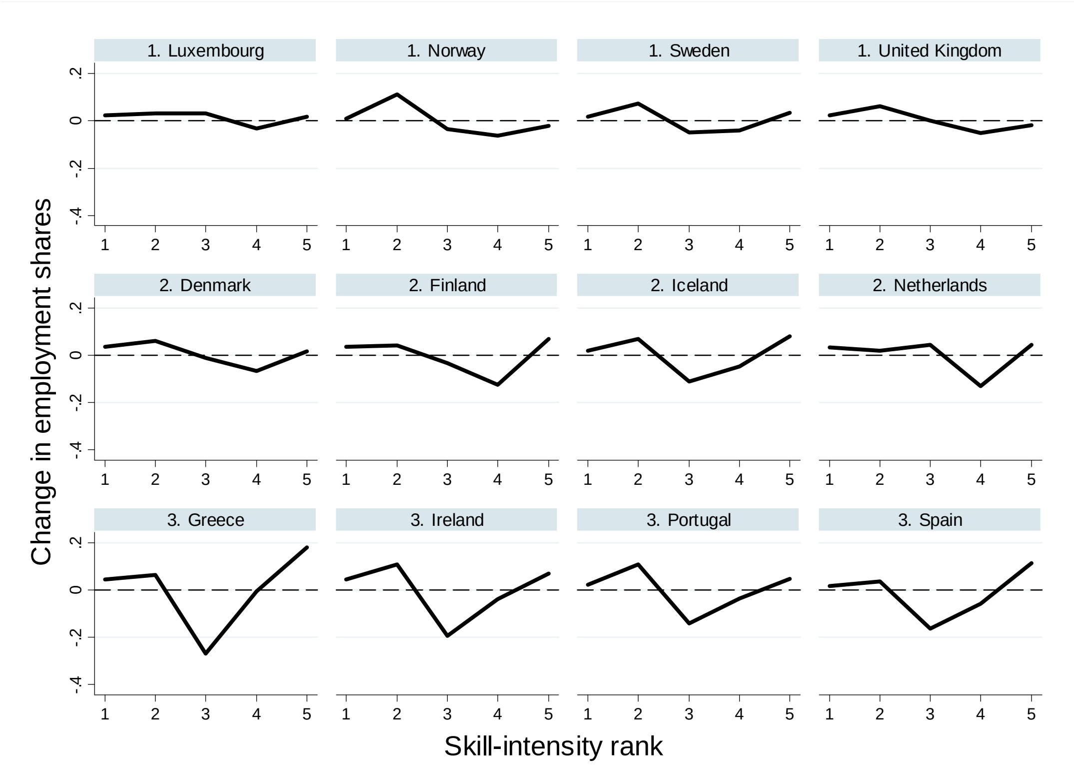
5 Explaining cross-country differences in the extent of polarization

Cross-country differences in the extent of job polarization are analyzed using five ranks of skill requirements, and are presented in Figure 2 illustrating employment changes in 12 European economies over the 1993-2001 period. Although all the analyzed countries experienced job polarization over the analyzed time interval, the differences across them are striking.

Note that in Benelux, Nordic and Anglo-Saxon countries (Denmark, Finland, Luxembourg, the Netherlands, Norway, Sweden, and the United Kingdom) the minimum is skewed towards occupations higher in the skill requirement distribution, while in Southern Europe (Greece, Portugal and Spain) plus Ireland and Iceland the minimum employment change happens in occupations close to the median in the skill requirement distribution. Additionally, the extent of polarization, measured as the difference between the lowest change in employment share for middle-skilled occupations and the highest change in employment share for low-skilled occupations, varies substantially across Europe. One observes the strongest polarization in the Southern European countries and Ireland, while the weakest polarization is observed in Nordic and Anglo-Saxon countries.

The polarization literature discusses two main sources of polarization. First, the decrease of employment in middle-skilled occupations is attributed to "routinization" (Autor et al., 2006), i.e., substitution of routine job tasks by modern technologies. Since machines carry out routine, precision tasks previously performed by administrative clerks or production workers, the demand for workers in occupations involving

Figure 2: Changes in employment share across European countries between 1993-2001 by occupational skill-intensity rank



Note: Source: European Union Labor Force Survey. For countries with shorter time spans (Finland, Iceland, Norway, Spain and Sweden), man hours worked were imputed on the basis of average annual growth rates. Skill intensity rank corresponds to the position of each occupation in the skill intensity distribution.

these tasks drops. The second hypothesized reason for the contraction of employment in middle-skilled occupations lies in offshoring (Acemoglu and Autor, 2010). The development of communication and transport technologies makes it cheaper to outsource certain job tasks to low-wage countries, which decreases the demand for occupations involving these tasks in the developed economies. Additionally, Goos et al. (2009) show that routine tasks content[8] has a negative influence on occupation-specific employment changes, while abstract tasks content has a positive influence on occupation-specific employment changes. Although Goos et al. (2009) do not find

[8] The routine tasks index is reported in the Occupational Information Network dataset (ONET).

any effects of offshorability[9] on employment changes in the U.K., Firpo et al. (2011) show that offshorability[10] is a strong determinant of the development of occupational wages in the U.S.

The channels through which "routinization" and offshorability are expected to affect allocation of labor across occupations might be strongly influenced by the economic cycle. Fast-growing countries are adopting new technologies at higher rates than other countries and thus "routinization" might have a greater impact on their labor markets. Additionally, the increase in the average educational achievement of a country's workforce might strengthen the polarization effect. First, this means that there are more high-skilled workers to implement new technologies, and, second, there are fewer people to work in middle-skilled occupations. Thus, we expect a positive correlation between both GDP growth and average educational attainment growth and the extent of polarization.

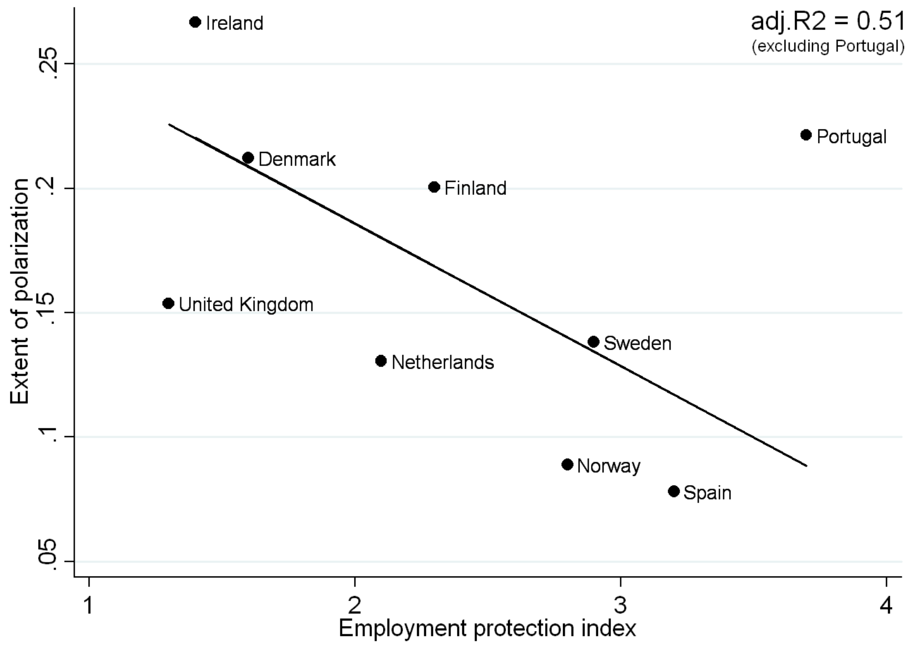
In addition to the above-discussed forces, the extent to which "routinization" and offshoring are expected to affect the shape of job polarization might be influenced by labor market institutions. In countries with high employment protection, it is more difficult to adjust employment to the prevailing technological conditions (Samaniego, 2006; Kugler and Pica, 2008) and thus the possibility of substituting workers with machines might be limited there. On the other hand, in countries with flexible labor markets employment adjusts to the changing structure of occupational skill requirement. Additionally, as high employment protection is supposed to slow down the process of adjusting the labor market to current economic and technological conditions, we might observe that polarization affects different occupations in countries with different degrees of employment protection. This leads me to the formulation of two hypotheses: (i) the extent of polarization should be negatively correlated with the strength of employment protection (once the economic cycle is controlled for); and (ii) in countries with strong employment protection the largest employment drop

[9] Goos et al. (2009) measure offshorability as the number of occurrences in the European Restructuring Monitor.

[10] Firpo et al. (2011) measure offshorability as an index based on ONET information about the necessity of face-to-face contact on site work, and decision-making for each occupation.

should be observed in occupations close to the median of the global skill requirement distribution.

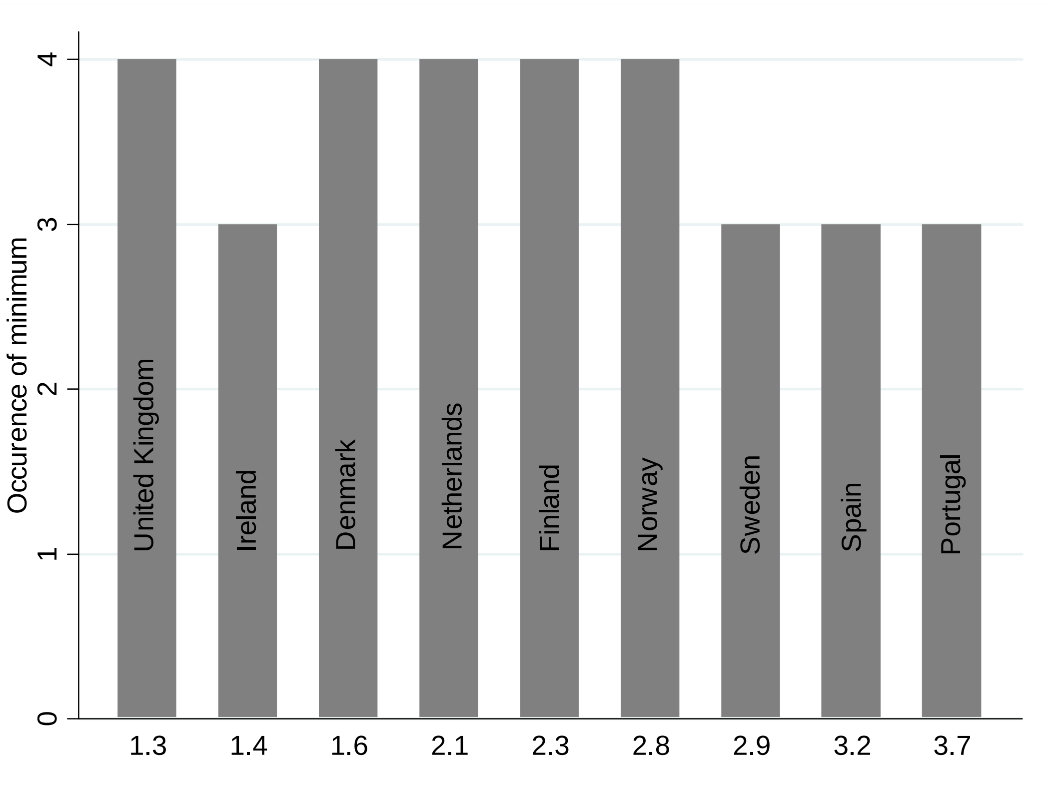
Figure 3: Correlation between the extent of polarization and employment protection



Note: This graph is constructed controlling for country-specific average educational achievement and GDP growth. The extent of polarization is measured as the difference between the lowest change in employment share (occurring either at 3rd or 4th fifth of the occupational skill-intensity distribution) and the highest change in employment share (occurring either at 2nd or 3rd fifth of the occupational skill intensity distribution). The employment protection index is an index decreasing on the {0, 5} range developed by Allard (2005) on the basis of the OECD methodology. This index is unavailable for Greece, Iceland and Luxembourg.

Using the measure of occupational skill intensity which is exogenous to the European labor market and the same for all analyzed economies, I can perform a consistent cross-country comparison of the polarization patterns to verify the above-stated hypotheses. This is illustrated in Figures 3 and 4. Using the employment protection index developed by Allard (2005) and reported in Nickell (2006), Figure 3 plots the correlation between the extent of polarization (after controlling for country-specific average educational achievement and GDP growth) and employment protection. As expected, countries with strong employment protection – the Southern European and Scandinavian countries – experience stronger polarization than other countries. Specifically, the conditional correlation between Allard's employment protection index and the extent of polarization is -0.37.

Figure 4: The occurrence of minimum change in employment share change and employment protection



Note: The occurrence of minimum is equal to the skill intensity rank at which the strongest drop in employment share over the 1993-2001 period is observed. The employment protection index is an index decreasing on the {0, 5} range developed by Allard (2005) on the basis of the OECD methodology. This index is unavailable for Greece.

Figure 4 illustrates the relationship between the occurrence of the minimum in the distribution of employment share changes and the employment protection index. It is clearly visible that countries with the strongest employment protection (Sweden, Spain and Portugal) experience the largest drop in employment in occupations around the median skill intensity, while the remaining countries (except Ireland) experience the strongest decrease of employment in occupations higher in the skill intensity distribution. Note that occupations around the median skillintensity are characterized by the strongest automation (Firpo et al., 2011), which makes them the most prone to "routinization".

6 Conclusion

Polarization of the labor market is a new phenomenon. Furter research is needed to better understand its causes and to draw conclusions for the future development of the labor market, as Acemoglu and Autor (2010) point out in their recent chapter in the Handbook of Labor Economy. This study applies a new measure of the skill requirements of occupations, which is independent of local labor market conditions, to analyze job polarization across Europe and reveals extensive cross-country differences in polarization patterns. Specifically, it is observed that polarization is stronger in Southern European countries and Ireland, while somewhat weaker in Northern Europe. By exploring the exogeneity of the skill requirement measure, I show that these differences in the extent and skewness of polarization are not only correlated with country-specific GDP and educational achievement growth, but also with the strength of employment protection.

The latter finding is especially interesting, as it indirectly confirms the existing theories explaining polarization – the "routinization" and offshoring hypotheses. According to these hypotheses, polarization is driven by workers employed in middle-skilled occupations being substituted by modern technologies or by a cheaper workforce in distant locations. Employment protection limits the possibility to adjust these firms' workforce in response to technological change and thus dampens the polarization effect. The natural next step in the development of the polarization literature would be to explicitly model the interaction between labor market institutions and occupational allocation of workers.

References

[1] Acemoglu D., Autor D. (2010) Skills, tasks and technologies: implications for employment and earnings. In: Ashenfelter, O., Card, D. (Eds.), Handbook of labor economics, vol. 4, Amsterdam: Elsevier.

[2] Allard G. (2005) Measuring job security over time: In search of a historical indicator, Instituto de Empressa Working Paper WP-05.

[3] Autor D., Katz L., Kearney M. (2006) The polarization of the U.S. labor market, American Economic Review, 96(2): 189—194.

[4] Autor D., Levy F., Murnane R. (2003) The skill content of recent technological change: An empirical exploration, Quarterly Journal of Economics, 118(4): 1279-1334

[5] Dustmann C., Ludsteck J., Schonberg U. (2009) Revisiting the German wage structure, Quarterly Journal of Economics, 142(2): 843-881.

[6] Elliott J., Gerova V. (2005): Translation from US OCC 2000 to ISCO-88, Center for Longitudinal Studies.

[7] Firpo S., Fortin N., Lemieux T. (2011) Occupational tasks and changes in the wage structure, IZA Discussion Paper No. 5542.

[8] Goos M., Manning A. (2007) Lousy and lovely jobs: The rising polarization of work in Britain, The Review of Economics and Statistics, 89(1):118-133.

[9] Goos M., Manning A., Salomons A. (2009) Job polarization in Europe, American Economic Review Papers and Proceedings, 99(2): 58-63.

[10] Kugler A., Pica G. (2008) Effects of employment protection on worker and job flows: Evidence from the 1990 Italian reform, Labour Economics, 15(1): 78-95.

[11] Nickell W. (2006) The CEP-OECD institutions data set (1960-2004), Centre for Economic Performance Discussion Paper No. 759.

[12] Pertold-Gebicka B. (2010) Measuring skill intensity of occupations with imperfect substitutability across skill types, CERGE-EI Working Paper No. 421.

[13] Samaniego R.M. (2006) Employment protection and high-tech aversion, Review of Economic Dynamics, 9(2): 224-24.

[14] Spitz-Oener A. (2006) Technical change, job tasks, and rising educational demands: Looking outside the wage structure, Journal of Labor Economics, 24(2): 235-270.

Conclusion

This dissertation focuses on the occupational allocation of workers. The first chapter analyzes the employment possibilities of college graduates in the Czech Republic, the second chapter concentrates on the occupational allocation of workers from different percentiles of the earnings distribution in the U.S., and the third chapter compares changes in occupation-specific employment across European countries.

The occupation-centered literature has recently been given increased attention due to its ability to account for the heterogeneous speed of technological progress (Firpo et al., 2009) and to accommodate the task-based approach in modeling the labor market (Acemoglu and Autor, 2010). The central point of this line of research is the observation of labor market polarization, i.e. employment and wage growth in low- and high-skilled occupations at the cost of middle-skilled occupations. This dissertation supplements the literature by developing an empirical model-based measure of the skill content of occupations, which is further used to analyze differences in workers' occupational allocation and polarization patterns both over time and across counties .

While Acemoglu and Autor (2010) suggest that the preferable framework to analyze recent developments in the labor market should consider workers with three distinct skill levels, in this dissertation I work with two levels of skills. This simplification allows for tractable modelling and facilitates empirical analysis without loss of generality. Using a two-skill language, one can easily define the third, intermediate skill level as an interaction of the two baseline levels: low-skilled occupations value mainly low skills, middle-skilled occupations value both skill levels to a similar extent, and high-skilled occupations predominantly value high skills.

The major empirical contributions of this work include (i) demonstrating that an increased supply of skilled workers boosts their productivity and thus causes a positive shift in the demand for skills; (ii) identifying the differences between earnings growth patterns in the 1980s and 1990s as being caused by strong across-occupation reallocation of workers in the earlier period; and (iii) documenting the link between employment protection legislation and the extent of polarization. These results bring important information to the economic literature. First, positive spillovers from a high concentration of skills should be taken into account when modelling interactions in the labor market. Second, occupations differ significantly in the extent to which different labor inputs can be used in their production process. Finally, as the extent to which employment reacts to technological progress is shown to be influenced by labor market institutions, there is a need to carefully model these interactions.

Results achieved in this dissertation also have important policy implications. The first chapter suggests that in the long run, districts should be able to positively stimulate their labor markets by providing higher education to a larger fraction of their population, while the last chapter shows that the evolution of labor markets in economies with strong employment protection lags behind countries with weaker labor market regulations.