Working Paper Series, 2010, Vol. 20(1), 41-87

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

Barbara Gebicka
CERGE-EI†

WP 407

Abstract

Public funding drives much of the recent growth of college degree supply in Europe, but few indicators are available to assess 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.

Abstrakt

Veřejné financování v současné době dominantně určuje růst nabídky vysokoškolsky vzdělaných lidí v Evropě, ačkoli existuje jen velmi málo indikátorů pro určení její optimální úrovně. V tomto článku zkoumám indikátor použití vysokoškolsky vzdělané pracovní síly – podíl absolventů vysokých škol na "vysokoškolských" pozicích. Gottschalk a Hansen (2003) navrhují identifikovat "vysokoškolské" pozice na základě platové 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 použitím vzdělané pracovní síly. Empirické výsledky založené na datech z NUTS-4 pro české regiony naznačují pozitivní vztah, tudíž potvrzují přítomnost endogenního vlivu počtu vzdělané pracovní síly na poptávce po ní.

Keywords: occupational allocation, demand for skills, productivity spillovers
JEL Classification: J20, J24

*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.
This paper was supported by Center of Advanced Political Economy Research [Centrum pro pokročilá politicko-ekonomická studia], No. LC542, (2005-2009)

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 is to analyze social returns[1] to higher education (Acemoglu and Angrist 2001, Moretti 2004), which directly captures the benefits of educating people, however, are difficult to measure. An alternative is offered by the overskilling literature (see McGuiness 2006 for a review), which investigates employment of college graduates in the so-called “noncollege” occupations (Pryor and Schaffer 1997, McGuiness 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

[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.

number of college graduates attracts firms 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, 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),

[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.

which could be reflected in more occupations being classified as "college" in these countries. More robust and more interesting is the positive 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 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 sys-

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

tem 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 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 significant wage punishment (McGuiness 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 the 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 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 paper 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, 2008). 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.

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

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

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.

District-specific production of college graduates is almost directly translated into the number of skilled workers in the local labor markets because of low cross-district migration of graduates.[6] While young Czechs move across districts to obtain a college education, they are much less likely to move after graduation. Low migration within the Czech Republic has been already documented by Fidrmuc (2004). This author, however, did not distinguish education-specific migration. To fill this gap, I compare the 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. with 40-44 and 45-50 years-old).[7] This comparison, presented in Table 2 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. 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

[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] The districts of Prague and Brno, the outliers in the number of college graduates, have been removed from this analysis.

final analysis is conducted with and without the high migration districts.

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.[8] 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.[9] The authors do not need to model this because they

[8]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.

[9]The supply functions of college and high school graduates to the "college" sector used

do not 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:

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

$$L^S_{C2} = L_C - L^S_{C1} \tag{5}$$

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

$$L^S_{N2} = L_N - L^S_{N1}, \tag{7}$$

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)[10] 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)

by Gottschalk and Hansen (2003) are the following: $L^S_{C1} = \lambda_C + \beta_C \frac{w_{C1}}{w_{C2}}$ and $L^S_{N1} = \lambda_N + \beta_N \frac{w_{N1}}{w_{N2}}$. Note that they do not explicitly account for the total amount of college- and high school-educated labor in the economy.

[10]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.

and demand conditions (labor productivities, α_{ij}) on the equilibrium fraction of college graduates working in the "noncollege" sector (π_C^*).

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

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[11] 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 2002, Moretti 2004). In this paper I use a simple linear relationship:

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

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

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:

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

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 us 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. Alternatively, the amount of capital complementing college-educated workers might be limited, which sanctions the number of college graduates who can be employed in "college" occupations. Discussion of these models is not within the scope of this paper. Let me note, however, that each of the alternative explanations 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).

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

where 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.[12]

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.[13]

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,[14]

---

[12] 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.

[13] 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.

[14] The dependent variable is at the individual level, while the explanatory variable of

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 get over this problem. I follow this procedure by first estimating 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.[15] 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, the expansion of hi-

interest is at the group (district) level.

[15] 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).

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.[16]

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[17] 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

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

[17] 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.

some colleges opened in the 1950's and 1960's 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[18] 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.[19]

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 amount 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[20] – would remove some of the unobservable demand from the error term and

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

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

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

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 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.[21] 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 use a higher threshold (15%). Nevertheless, as presented in Section 7.4, the qualitative results are insensitive to the chosen threshold.

[21] 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.

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 lower this threshold to 85%. Additionally, I classify occupations where more than 95% of workers have only a high school diploma as "noncollege" occupations.

The procedure of classifying occupations can be described as follows. For each 3-digit occupation where college graduates constitute between 5% and 85% of all employees, I estimate the following wage equation:

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

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.[22] 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 0.15) 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 85% 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.

[22] The sample used to classify of 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.

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 commercial sector, which report wages and other information for about 1.3 million workers (about 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[23] 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.

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

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 on a 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 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 3, 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 3 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 "non-college" 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.[24] 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.

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

[24] 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.

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 4 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.

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[25] and that an increase in the relative stock of college graduates in the local labor market worsens their employment situation.

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

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

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 15% or when the proportion of college graduates working there exceeds 85%. 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 10% and the proportion threshold at 90%. These are the values used in previous research to distinguish between the "college" and "noncollege" occupations. As seen in Tables 5 - 6, the use of an alternative definition leads to qualitatively the same results.

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 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 7. 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.

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), UK (Grazier 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 educational 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 shed some light on the within-countries patterns observed in Figure 1, the cross-countries 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] 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.

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

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

[7] 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.

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

[9] CSO, Czech Statistical Office (2008)
http://www.czso.cz/eng/redakce.nsf/i/people_and_society_s1m

[10] 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.

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

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

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

[14] 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.

[15] 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.

[16] 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.

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

[18] 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.

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

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

[21] Katz L., Autor D. (1999) Changes in the wage structure and earnings inequality, in: Ashenfelter O., Card D. (eds.), Handbook of Labor Economics, Amsterdam.

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

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

[24] 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.

[25] 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.

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

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

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

[29] Roback J. (1982) Wages, rents and the quality of life, *Journal of Political Economy*, 90(6): 1257-1278.

[30] Roback J. (1988) Wages, rents and amenities: Differences among workers and regions, *Economic Inquiry*, 26(1): 23-41.

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

Appendix

A1. Derivations

The model proposed by Gottschalk and Hansen (2003) and adapted 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 heterogeneous 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 ($\frac{\alpha_{C2}}{\alpha_{N2}}$).

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 the "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 characteristics of equilibrium in the labor market. L_{N1} is, as before, the amount of high school-educated labor in the "college" sector, while α_{i1} are productivities of labor type i in the "college" sector. 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 as follows. 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:

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

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

$$\pi_C^{*} \equiv 1 - \frac{L_{C1}^{*}}{L_C} = 1 - \frac{\int_{\varepsilon=\varepsilon*}^{\infty} dG(\varepsilon)}{L_C} = f(\varepsilon*, L_C) = f(\underset{+}{L_C}, \underset{-}{\alpha_{C1}}, \underset{+}{\alpha_{N1}}, \underset{+}{\alpha_{C2}}, \underset{-}{\alpha_{N2}}). \qquad \text{(A6)}$$

Like in the baseline model, also here we observe that the average wage of college graduates is higher in the "college" sector than in the "noncollege" sector, what 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 also applicable in this case.

College-skills Complementing Capital

Allocation of workers between "college" and "noncollege" occupations could also be driven by the capital structure. For simplicity, I assume that "noncollege" firms do not use any capital, while "college" firms use capital-complementing skilled workers. This can be expressed using the Leontieff function:

$$Q_1 = F_1(\alpha_{C1} \min\{L_{C1}, K_1\} + \alpha_{N1} L_{N1}), \tag{A7}$$

where K_1 is the amount of skill-complementing capital in the "college" sector. 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:

$$w_{C1} = \begin{cases} \alpha_{C1} \frac{\partial F_1}{\partial L_1} & \text{for up to } K_1 \text{ workers} \\ 0 & \text{for every additional worker} \end{cases} \qquad w_{N1} = \alpha_{N1} \frac{\partial F_1}{\partial L_1} \tag{A8}$$

$$w_{C2} = \alpha_{C2} \frac{\partial F_2}{\partial L_2} \quad \text{and} \quad w_{N2} = \alpha_{N2} \frac{\partial F_2}{\partial L_2}, \tag{A9}$$

where $L_1 = \alpha_{C1} \min\{L_{C1}, K_1\}$ 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.

Workers choose to work in the sector which pays them a higher wage. As all high school graduates are identical, they should be paid the same wage no matter which sector they work in or we would observe college workers exclusively in one sector. College graduates are also all identical, but there is a limit on how many of them are productive in the "college" sector. A positive wage can be paid only to a limited number of college-educated workers and thus even if it is higher than in the "noncollege" sector, not all college graduates will work there. This can be expressed by the following conditions:

$$\begin{aligned}
L_{C1} &= \begin{cases} \min\{L_C, K_1\} & \text{if } w_{C1}(L_{N1}, L_{C1}) > w_{C2}(L_{N2}, L_{C2}) \\ L_{C1}: \ w_{C1}(L_{N1}, L_{C1}) = w_{C2}(L_{N2}, L_{C2}) \\ 0 & \text{if } w_{C1}(L_{N1}, L_{C1}) < w_{C2}(L_{N2}, L_{C2}) \end{cases} \\
L_{C2} &= L_C - L_{C1} \\
L_{N1} &= \begin{cases} L_N & \text{if } w_{N1}(L_{N1}, L_{C1}) > w_{N2}(L_{N2}, L_{C2}) \\ L_{N1}: \ w_{N1}(L_{N1}, L_{C1}) = w_{N2}(L_{N2}, L_{C2}) \\ 0 & \text{if } w_{N1}(L_{N1}, L_{C1}) < w_{N2}(L_{N2}, L_{C2}) \end{cases} \\
L_{N2} &= L_N - L_{N1}.
\end{aligned}$$

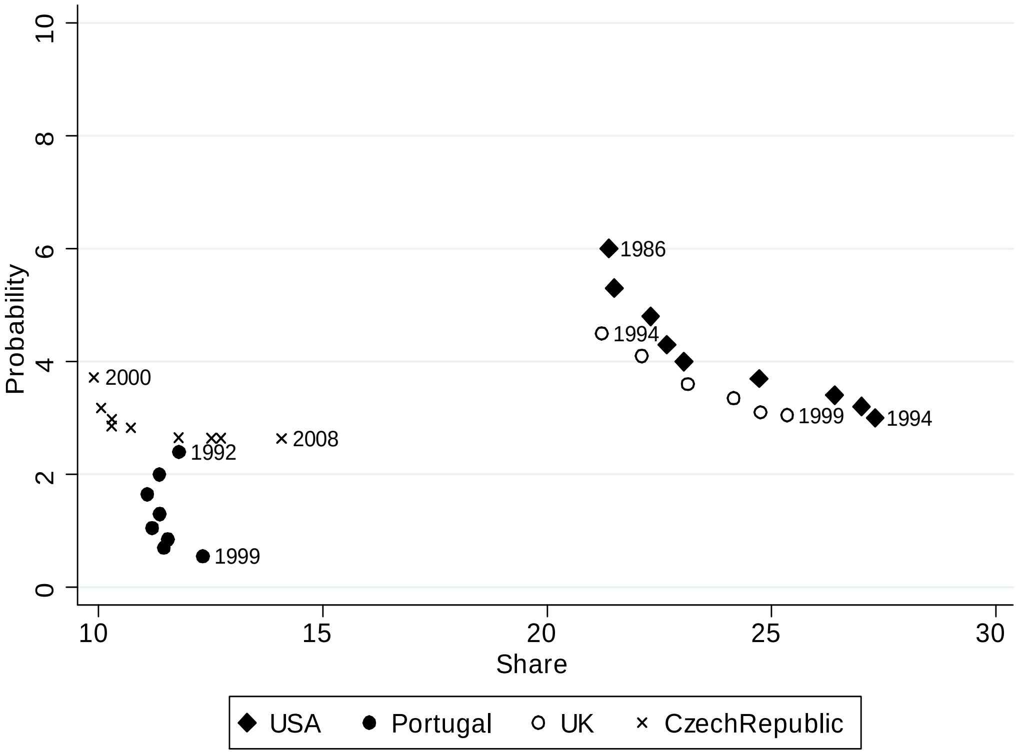
where $w_{C1}(L_{N1}, L_{C1}), w_{N1}(L_{N1}, L_{C1}), w_{C2}(L_{N2}, L_{C2})$ and $w_{N2}(L_{N2}, L_{C2})$ are given by Equations (A8) and (A9), respectively. Recall that by assumption $\frac{\alpha_{C1}}{\alpha_{N1}} > \frac{\alpha_{C2}}{\alpha_{N2}}$, which leads to $\frac{w_{C1}}{w_{N1}} > \frac{w_{C2}}{w_{N2}}$ and thus we must have either $L_{C1} = \min\{L_C, K_1\}$ or $L_{N1} = 0$ (or both). In either case the property allowing for classification of occupations into "college" and "noncollege" is the same as in Gottschalk and Hansen (2003).

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

$$\pi_C^* \equiv 1 - \frac{L_{C1}^*}{L_C} = 1 - \frac{K}{L_C} = f\left(L_C, L_N, K_1, \alpha_{C1}, \alpha_{N1}, \alpha_{C2}, \alpha_{N2}\right). \tag{A10}$$

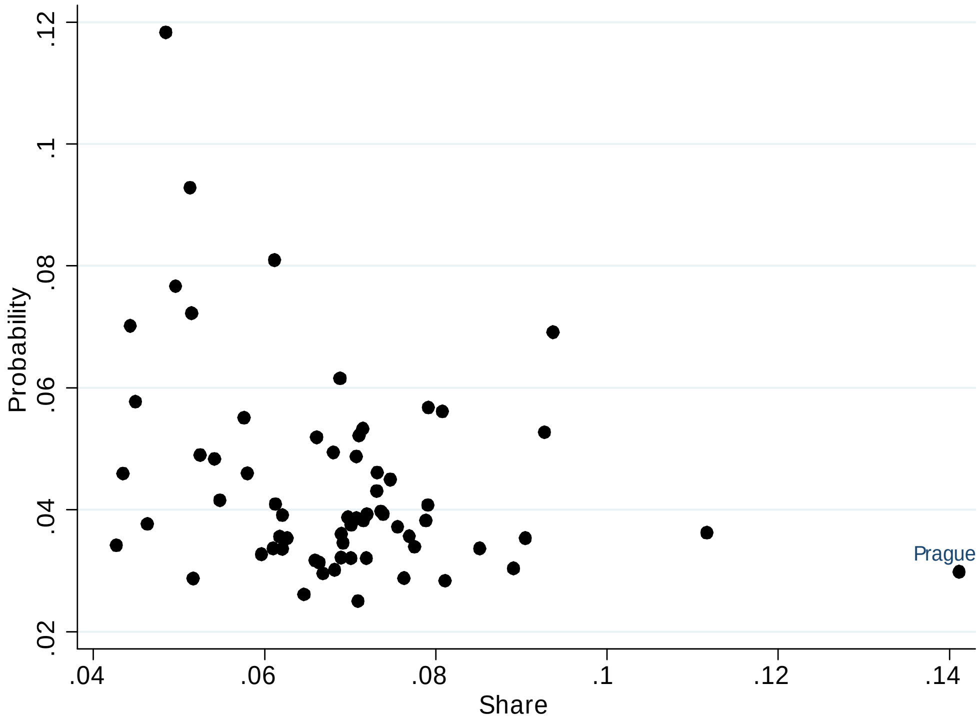
A2. Figures

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.



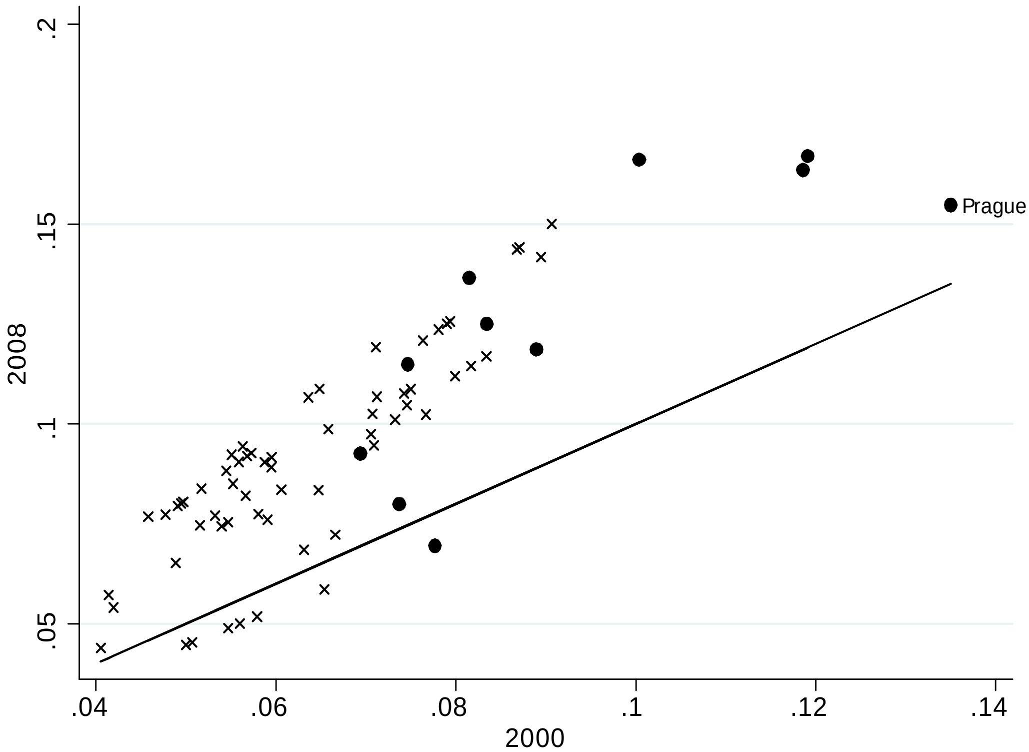
Source: Author's compilation using Gottschalk and Hansen (2003), Cardoso (2007), Grazier et al. (2008), Eurostat, and U.S. Census Bureau as well as the ISPV data.

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.



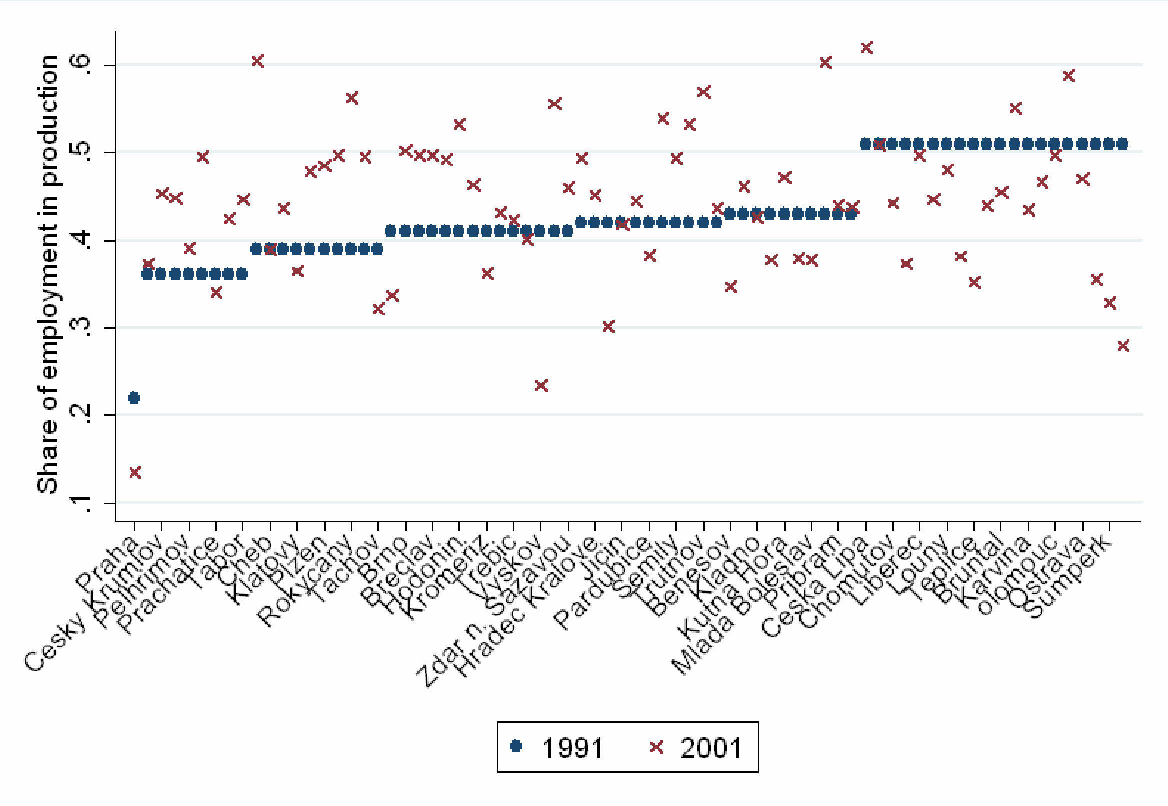
Note: Young population consists of people below the age of 35. Source: Author's calculations using 2001 Census and the ISPV data.

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: Author's calculations using 2001 Census and the 2000-2008 Czech Labor Force Survey.

Figure 4: Share of employment in production industry across districts in 1991 and 2001.



Note: In 1991 the industrial structure is defined on the regional (NUTS-3) level.

Source: Author's compilation using 1991 and 2001 Censuses.

A3. Tables

Table 1: 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 2: Changes in cohort-specific sizes of college-educated population between 1991 and 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%

	Born in 1961-1965			Born in 1956-1960		
	1991	2001	Change	1991	2001	Change
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%
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%

	Born in 1961-1965			Born in 1956-1960		
	1991	2001	Change	1991	2001	Change
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

Note: 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 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 the 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 3: 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 an 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.

Table 4: Determinants of the share of college graduates in "noncollege" occupations across 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 an 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.

Table 5: Determinants of the share of college graduates in "noncollege" occupations across Czech districts in 2001

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	OLS	IV	IV	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 an 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.

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

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	OLS	FE	FE	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 an 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.

Table 7: Determinants of the share of college graduates in "noncollege" occupations across 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 an 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 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.