Measuring skill intensity of occupations with imperfect substitutability between more and less educated labor.

Barbara Gebicka*
CERGE-EI†

September 2009

Abstract

The literature on wage inequality lacks a consistent measure of skill-intensity of occupations, which could be used to investigate recent evolutions it the labor market. This paper proposes such a measure. Instead of using occupation-specfic average educational achievement or skill wage premia, like many studies do, I work with occupation-specific relative productivities of more to less skilled workers. In the setup where differently skilled workers are not perfect substitutes, measurement of relative productivities requires estimation of substitution elasticities. I propose a strategy to consistently estimate occupation-specific elasticities of substitution using March CPS data from 1983 to 2002. These are further used to calculate occupation-specific relative productivities of college to high school graduates, which are argued to capture the skill-intensity of occupations. As an illustration, this measure is applied to test the modified skill-biased technological change hypothesis as proposed by Autor et al. (2003, 2006, 2009).

*I would like to thank David Card, Enrico Moretti and the graduat students at the University of California at Berkeley for their hospitability and helpful comments. I am also indebted to Stepan Jurajda and Alena Bicakova for their guidance and advice. All remaining errors are mine. This research was supported by the Grant Agency of the 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

Recent literature has achieved a consensus that not only supply but also demand side of the labor market is heterogenous. First, workers can produce different value added performing different jobs; and second, the structure of jobs present in the market changes over time Thus, to understand the evolution of wages and the demand for labor it is useful to investigate the labor market partitioned into individual occupations which capture heterogenous tasks and skill requirements. The necessity of this approach is to have a measure of skill contents of occupations. Ideally, one would like to know which skills and with what intensity are required to correctly perform the job of individual occupations. For a simplified analysis it is enough to have an index representing skill-intensity of occupations that would order them from easy to learn, repetitive and automatic to those requiring fast decision-making and difficult to learn skills.

A natural approach is to take advantage of skills, abilities and work activities associated with occupations as reported by O*NET (replacing the Dictionary of Occupational Titles - DOT). This dataset is widely used in the literature in the context of income inequality (Autor et al. 2008), overeducation (McGoldrick and Robst 1996) and other. It constitutes a very useful, comprehensive source of information about occupations. The disadvantage of this database is its rigidity and nontransparency. Originally, DOT was updated, on average, every 4-5 years. O*NET is updated continuously, however, new occupation descriptions always depart from the previous ones and I am not aware of any clear algorithm governing the assessment of skill requirements, what could lead to some inconsistencies. More tractable alternatives have been offered by researchers. Some studies use occupation-specific average years of schooling as a proxy for their skill content (Goos and Manning 2007, Autor et al. 2006). This approach is transparent and accounts for short-term changes in occupations' characteristics. On the other hand, it uses an implicit assumption that employment structure reflects the level of skill requirements, which does not have to be the case, especially in the fast-changing occupations. Average years of schooling as a measure of education requirement is, for example, used by Pryor and Schaffer's (1997) to analyze of the fraction of college graduates underutilizing their skills, i.e., working in the so called "noncollege" occupations. This line of research, strongly relying on having accurate requirements of occupations, offers one more measure of the skill intensity of occupations. Gottschalk and Hansen (2003) argue that the occupation-specific college-high school wage gap reflects the relative productivity of college and high school graduates and as such could be used to order occupations according to their skill intensity.

As mentioned above, Gottschalk and Hansen (2003) propose to order occupations according to a parameter of their production function - the relative productivity of college to high school graduates. Although these authors use the estimate of relative productivity for a binary classification of occupations into "college" and "noncollege", it could be utilized in its whole continuum as a measure of skill contents of occupations. Relative productivity of more to less skilled workers directly reflects utilization of skills in given occupation's production technology and, as such, it offers a consistent and objective measure of occupation-specific skill-intensity. In this study I propose a methodology to calculate this measure. It is important to note that, as opposed to Gottschalk and Hansen, I do not treat college and high school graduates as perfect substitutes. Although this assumption significantly simplifies the analysis, it is questionable - there exist many studies estimating both the short-run and long-run elasticity of substitution between more and less educated labor in the whole U.S. economy to be around 1.4.[1] This is not possible to obtain in the setup where the majority of occupations interchangeably use both types of workers.[2]

In this study I extend the methodology proposed by Gottschalk and Hansen (2003) to allow for finite within-occupation elasticity of substitution between high school and college graduates. Following the common practice in the literature, I assume that

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

[2] Gottschalk and Hansen allow for occupations which can employ only college graduates (e.g. medical doctors or judges) by setting productivity of the other skill group to zero at these occupations.

occupation-specific production function is of constant elasticity of substitution (CES) type and use a modification of the approach proposed by Card (2001) to estimate the elasticity parameter. Calculation of occupation-specific relative productivity of college to high school graduates is then straightforward. In this way a new measure of skill contents of occupations is produced. It may fit applications in many contexts. The examples presented in this paper include the replication of Gottschalk and Hansen analysis of the fraction of college graduates employed in "noncollege" occupations and testing the modified version of the SBTC as proposed by Autor et al. (2003, 2006, 2008).

The rest of this text is organized as follows. In the next section I present a model of workers allocation across occupations characterized with different skill-intensity used further for empirical analysis. Section 3 describes econometric procedures used to identify occupation-specifc elasticities of substitution between college and high school graduates which then allow to estimate skill-intensity of occupations. Next section presents results of these estimations. In section 4 I use the estimated occupation-specific skill-intensities to test the modified version of the SBTC and analyze the evolution of the fraction of college graduates in "noncollege" occupation. The last section concludes.

2 Theoretical specification

I argue that the within occupation relative productivity of college to high school graduates, where college graduates represent highly skilled labor and high school graduates represent less skilled labor, could be used as a proxy for occupation-specific skill-intensity. Let me illustrate this using a relatively general occupation-specific production function, a CES aggregate of college- and high school-educated labor, as specified in Equation (1).

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

where Y_j is the amount of output produced by occupation j, L_{Cj} is the number of college graduates, L_{Nj} is the number of high school graduates employed in occupation j, and γj

is a parameter describing substitutability between these two labor types.[3] In this context, $\frac{\alpha_{Cj}}{\alpha_{Nj}}$ describes the occupation-specific relative productivity of differently educated workers.[4] In occupations where this parameter assumes high values, college graduates are much more productive than high school graduates, what could be attributed to the skill difference between differently educated workers. That is why $\frac{\alpha_{Cj}}{\alpha_{Nj}}$ describes skill-intensity of an occupation. It tells us how crucial college-gained skills[5] are for the tasks performed within a specific occupation.

Under the simplifying assumption made by Gottschalk and Hansen (2003), i.e. when $\gamma_j = 1$, $\frac{\alpha_{Cj}}{\alpha_{Nj}}$ is fully reflected in relative wage of college to high school graduates paid by occupation j. The assumption that differently educated labor can be perfectly substituted within an occupation reduces Equation (1) to $Y_j = F_j\left(\alpha_{Cj}L_{Cj} + \alpha_{Nj}L_{Nj}\right)$. Rearrangement of first order conditions gives $\frac{\alpha_{Cj}}{\alpha_{Nj}} = \frac{w_{Cj}}{w_{Nj}}$. This is why Gottschalk and Hansen categorize occupations according to the college wage premia which they pay. In the setup where college and high school graduates are allowed to be imperfect substitutes, one needs to know the elasticity of substitution to calculate occupation-specific relative productivity.

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

[3] The elasticity of substitution is $\sigma_j = \frac{1}{1-\gamma_j}$.

[4] Let me note that parameters α_{Cj} and α_{Nj} capture both the ralative income shares and productivities of college and high school graduates. What actally matters, is their relative value, $\frac{\alpha_{Cj}}{\alpha_{Nj}}$ (see Card and DiNardo, 2002, for a more detailed discussion), so I could actually write $Y_j = F_j\left(\left(\alpha_j L_{Cj}^{\gamma_j} + L_{Nj}^{\gamma j}\right)^{\frac{1}{\gamma j}}\right)$. I keep the notation as presented in the text to be consistent with Gottschalk and Hansen (2003).

[5] College-gained skills are assumed to be mainly nonroutine.

2.1 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 a labor aggregate L_j being a CES aggregate of college- and high school-educated labor. Occupations differ by their skill-intensity ($\frac{\alpha_{C_j}}{\alpha_{N_j}}$) and the elasticity of substitution between college and high school graduates ($\sigma_j = \frac{1}{1-\gamma_j}$). The production function used by occupation j could be summarized in the following way:

$$L_j = \left(\alpha_{Cj} L_{Cj}^{\gamma_j} + \alpha_{Nj} L_{Nj}^{\gamma j}\right)^{\frac{1}{\gamma j}}, \tag{2}$$

where L_{Cj} and L_{Nj} are the amounts of college- and high school-educated labor employed in occupation j[6].

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

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

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

$$\frac{w_{Cj}}{w_{Nj}} = \frac{\alpha_{Cj}}{\alpha_{Nj}}\left(\frac{L_{Nj}}{L_{Cj}}\right)^{1-\gamma_j}, \tag{3}$$

[6] Under the assumption that each worker works the same amount of hours, L_{Cj} and L_{Nj} represent the count of workers of each education level employed in occupation j.

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

$$\ln\left(\frac{L_{Cj}}{L_{Nj}}\right) = \sigma_j \ln\left(\frac{\alpha_{Cj}}{\alpha_{Nj}}\right) - \sigma_j \ln\left(\frac{w_{Cj}}{w_{Nj}}\right). \tag{4}$$

Equation (4) describes the relative demand for labor in occupation j. It depends on relative wages of the two education groups and their relative productivities within occupation j.

2.2 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 market specific to occupation j). These numbers capture all workers who would supply labor to occupation j under favorable labor market conditions, i.e. if the wages offered there were high enough. We observe just some of these people actually working in occupation j because of their heterogenous preferences towards job attributes - workers differ in their reservation wage. The notion of occupation-specific labor markets, introduced by Card (2001), pins down the observation that workers can switch occupations, as a reaction to productivity shocks affecting these occupations, however, they are limited to occupations within their specialization. With that in mind, I define the supply of labor to occupation j to be log-linearly dependent on wages:[7]

$$\begin{aligned} \ln\left(\frac{L_{Cj}}{N_{Cj}}\right) &= \beta_j \ln w_{Cj} \\ \ln\left(\frac{L_{Nj}}{N_{Nj}}\right) &= \beta_j \ln w_{Nj}. \end{aligned} \tag{5}$$

Log-linear aggregate labor supply functions are commonly used when describing supply of workers to different units of production, usually occupations (Card 2001, Gottschalk and Hansen 2003). The occupation-specific elasticity of supply, $\beta_j > 0$, represents workers' aggregate preferences towards occupation j. It is assumed to be the

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

same for each education group within the occupation-specific labor market. This assumption is crucial for the model to have a closed-form solution. Although at first sight questionable, it actually finds support in the literature saying that these are gender and specialization (not education level) that are major determinants of preferences towards job attributes.

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

$$\ln\left(\frac{L_{Cj}}{L_{Nj}}/\frac{N_{Cj}}{N_{Nj}}\right) = \beta_j \ln\left(\frac{w_{Cj}}{w_{Nj}}\right). \tag{6}$$

2.3 Equilibrium

Equations (2) and (4) represent the relative demand and supply of labor for occupation j. Equalizing supply with demand, and rearranging, we arrive at the system describing the equilibrium relative wages and relative employment in each occupation:

$$\begin{cases} \ln\left(\frac{w_{Cj}}{w_{Nj}}\right) = \frac{\sigma_j}{\sigma_j+\beta_j}\ln\left(\frac{\alpha_{Cj}}{\alpha_{Nj}}\right) - \frac{1}{\sigma_j+\beta_j}\ln\left(\frac{N_{Cj}}{N_{Nj}}\right) \\ \ln\left(\frac{L_{Cj}}{L_{Nj}}\right) = \frac{\sigma_j\beta_j}{\sigma_j+\beta_j}\ln\left(\frac{\alpha_{Cj}}{\alpha_{Nj}}\right) + \frac{\sigma_j}{\sigma_j+\beta_j}\ln\left(\frac{N_{Cj}}{N_{Nj}}\right) \end{cases}. \tag{7}$$

Let us note that both relative wages and relative employment depend on the 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 parameters of occupation-specific elasticity of labor supply and elasticity of substitution between the two labor types.

The system derived above relies strongly on the functional forms assumed. Nevertheless, these are the most widely used functional forms, CES production function and log-linear supply function, which constitute a good baseline for this study.

3 Econometric approach

To estimate the parameters of the above presented model, let us analyze the economy, as described in the previous section, in several consecutive periods (t). In each period occupation-specific labor markets, $\frac{N_{Cjt}}{N_{Njt}}$, are different (new graduates enter labor markets, some people leave to retirement, etc.) and the relative productivity of college to high school graduates changes (due to the SBTC and other shocks). Thus, in each year we observe different equilibrium values of occupation-specific relative wages and employment. This variation could be used to identify the system of equations as presented in (7).

To completely specify the model, let us note that the relative productivity of labor varies over time and across occupations. It consists of three unobservable components: occupation-specific (characteristic to a given occupation, constant over time), year-specific (common for all occupations) and occupation-year specific effects, what can be expressed as $\ln\left(\frac{\alpha_{Cjt}}{\alpha_{Njt}}\right) = \varepsilon_j + \varepsilon_t + \varepsilon_{jt}$. Using this notation, the system (7) could be rewritten as the following econometric model:

$$\begin{cases} \ln\left(\frac{w_{Cjt}}{w_{Njt}}\right) = \frac{\sigma_j}{\sigma_j+\beta_j}\varepsilon_j - \frac{1}{\sigma_j+\beta_j}\ln\left(\frac{N_{Cjt}}{N_{Njt}}\right) + \frac{\sigma_j}{\sigma_j+\beta_j}\varepsilon_t + \frac{\sigma_j}{\sigma_j+\beta_j}\varepsilon_{jt} \\ \ln\left(\frac{L_{Cjt}}{L_{Njt}}\right) = \frac{\sigma_j\beta_j}{\sigma_j+\beta_j}\varepsilon_j + \frac{\sigma_j}{\sigma_j+\beta_j}\ln\left(\frac{N_{Cjt}}{N_{Njt}}\right) + \frac{\sigma_j\beta_j}{\sigma_j+\beta_j}\varepsilon_t + \frac{\sigma_j\beta_j}{\sigma_j+\beta_j}\varepsilon_{jt} \end{cases},$$

which, for simplicity, could be written as

$$\begin{cases} \ln\left(\frac{w_{Cjt}}{w_{Njt}}\right) = c_{j0} + c_{j1}\ln\left(\frac{N_{Cjt}}{N_{Njt}}\right) + v_t + v_{jt} \\ \ln\left(\frac{L_{Cjt}}{L_{Njt}}\right) = d_{j0} + d_{j1}\ln\left(\frac{N_{Cjt}}{N_{Njt}}\right) + \mu_t + \mu_{jt} \end{cases}, \tag{8}$$

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

This model describes the simultaneous determination of occupation-time specific relative wages and relative employment as the function of occupation-time specific relative labor markets. Let us note that the occupation-specific elasticity of substitution between college and high school graduates, σ_j, could be expressed as $-\frac{d_{j1}}{c_{j1}}$. Thus, consistent estimation of $\widehat{c_{j1}}$ and $\widehat{d_{j1}}$ allows for identification of $\widehat{\sigma}_j$.

It is important to note that direct estimates of c_{j1} and d_{j1} are likely to be biased upwards because of the endogenous nature of occupation-specific labor markets. As a result of a positive skill-biased productivity shock affecting occupation j, relative wages and relative employment 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 there. In the existing literature, such problem is commonly dealt with by assuming that the time evolution of relative productivity is log-linear (Katz & Murphy 1992, Card & DiNardo 2002, Autor et al. 2008), i.e. that $\varepsilon_t + \varepsilon_{jt}$ can be approximated by a linear time trend. This does not capture all the unobservable shocks to relative labor productivity, however, it captures the ones that can be expected by workers and thus influence the structure of the occupation-specific labor market.

Finally, to identify the above system, we need to observe a variation in the relative wages and relative employment as a reaction to changes in the structure of occupation-specific labor markets. This reaction might not be immediate and thus year-to-year changes in relative wages and relative employment might not well reflect the adjustments to the structure of occupation-specific labor markets. To assure that this causality is captured, I use 3-year changes in the observed variables, i.e. difference system (8) using 3-year-lags. Hence, I finally estimate:

$$\begin{cases} \triangle \ln\left(\frac{w_{Cjt}}{w_{Njt}}\right) = c_{j1} \triangle \ln\left(\frac{N_{Cjt}}{N_{Njt}}\right) + c_{j2}t + \zeta_t \\ \triangle \ln\left(\frac{L_{Cjt}}{L_{Njt}}\right) = d_{j1} \triangle \ln\left(\frac{N_{Cjt}}{N_{Njt}}\right) + d_{j2}t + \xi_t \end{cases}, \tag{9}$$

where $\triangle$ represents a difference between current and three-year-lagged value of a given variable, the time trend is included to account for occupation-specific regularities in the evolution of relative productivity,[8] and ζ_t and ξ_t capture random variations in changes of relative productivity shifters.

[8] The evolution of relative productivity (or of the extent of occupation-specific skill bias) could be accelerating or deccelerating during the analyzed period. Previous analyses of economy-wide evolution of the extent of skill bias suggest that this measure has been growing at a decreasing rate. If this applies to occupation j, we expect to see negative estimates of c_{j2} and d_{j2}.

Estimation of this system as a SUR, giving us consistent estimates of c_{j1} and d_{j1}, together with calculation of the elasticity of substitution between labor types, $\widehat{\sigma_j} = -\frac{\widehat{d_{j1}}}{\widehat{c_{j1}}}$, is presented in Section 5. This Section also shows final calculation of the proposed measure of occupation-specific relative productivity of college to high school graduates. Having the estimates of college wage premium, relative employment (calculated as the weighted average from the data) and the estimates of occupation-specific substitution elasticity, I calculate the relative productivity as:

$$\widehat{\frac{\alpha_{Cjt}}{\alpha_{Njt}}} = \frac{w_{Cjt}}{w_{Njt}} \left(\frac{L_{Cjt}}{L_{Njt}} \right)^{-\frac{1}{\widehat{\sigma_j}}} \tag{10}$$

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

4 Data and measurement issues

The data used in this study come from 1983-2002 March CPS (covering earnings from 1982 till 2001). This is the longest time span with consistent occupational data, which are crucial for my analysis.[9] Due to a limited number of observations offered by March CPS, I need to merge three consecutive years to obtain sample sizes allowing to do occupation-level analysis. 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 covers the decade of rapid increase in the college-high school wage gap as well as the later slowdown in the rate of growth in this gap. Thus it should be enough to capture the interesting phenomena in the labor market.

In order to make my analysis comparable to Gottschalk and Hansen (2003), I apply the same restrictions to the data as they 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

[9] 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.

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 Lemieaux (2001), I concentrate just on recent school leavers defined as individuals with 10 or less years of potential labor market experience.[10] Both full time and part time workers are included in the sample to assure enough number of observations. However, self-employed individuals are excluded from the sample and so are those with reported working hours per week being zero or above 98. The earnings measure used in this analysis is log of weekly earnings defined as yearly wage and salary income divided by weeks worked last year. I express earnings in 2000 dollars.

I deal with earnings censoring by adjusting the top-coded earnings by the factor of 1.4 times the top-code. Re-coding of occupations due to switch from the 1980 to 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 of education instead of the highest degree obtained, I keep in the sample the individuals with 12-17 years of education. Those with 16 or 17 years of education are assumed to be college graduates. Occupations are defined on 3-digit level. However, some of the 3-digit categories had to be merged with other 3-digit categories to assure enough sample sizes. Merging was done according to Gottschalk and Hansen (Table A1).[11]

March CPS is used to obtain the variables present in the system 9: occupation-time specific relative wages of college to high school graduates, relative quantities of these two labor types and the occupation-specific labor markets. College to high school relative wages are estimated using the log-wage regression widely used to estimate returns to college. Relative employment is calculated as the ratio of weighted amounts of college and high school graduates observed in a given occupation in a given year.

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

[11] See Gottschalk and Hansen (2003) for more detailed description of occupations coding and aggregation.

Occupation-specific labor markets N_{Cjt} and N_{Njt}, need to be defined carefully. They are composed of all workers who would supply labor to occupation j within period t if the labor market conditions were favorable enough. In determining this measure, I draw on Card (2001), who proposes to consider an individual's occupation as a probabilistic outcome (π_{ij}) that depends on her underlying characteristics. Under this assumption, 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 relevant population. I modify Card's approach to fit the analysis with many detaily defined occupations over time (as opposed to 6 broad occupation groups across the US cities).

The first issue is definition of the relevant population. I argue, that potential suppliers of labor to occupation j can be found within current employees of all occupations from which we observe workers switching to occupation j. To find these occupations, I look at occupation-switchers observed in the matched panel subsamples of March CPS.[12] As these samples are too small to capture a reasonable number of switches, I analyze several consecutive years. I divide the time period analyzed in this study with into 3 intervals: 1984-1989, 1990-1995 and 1996-2001, and use switches observed over the whole interval to define relevant populations for each year within that interval.

Within the relevant population, defined separately for each occupation and each education level, I estimate a linear probability model for working in occupation j. I use 6 consecutive March CPS samples (within the time intervals defined above) to estimate the following equation:

$$prob(occ_{it} = j) = X_{it}\beta + \pi t + \nu_{it},$$

where the dependent variable equals one if an individual i works in occupation j at time t, X_{it} contains individual demographic characteristics such as gender, age and race, as well as the region where she lives, t is the time trend, and ν_{it} captures individual unobservable effects. Then, I estimate the fitted values, $\widehat{\pi_{ij}}$, for each person within

[12] See Peracchi and Welch (1995) for a description of the matching procedure.

the relevant population. The year-specific sum of these fitted values over the relevant population represents the occupation j's specific labor market in the given year. This measure could be thought of as the number of people that would work in occupation j in year t if the productivity shocks experienced by this occupation exactly followed the linear trend. As such, this measure is independent of yearly deviations from the linear trend.

5 Empirical implementation

In this Section, I present step-by-step results leading towards the estimation of occupation-specific skill-intensity. As explained in Section 3, the main challenge of this analysis, and the main contribution of this study, is the estimation of occupation-specific elasticity of substitution between more and less educated labor.

The first step towards estimation of skill-intensity of occupations involves identification of occupation-specific labor markets. I follow the strategy outlined in Section 4 to measure the number of college and high school graduates who would work in each occupation in each analyzed year under perfect conditions (i.e. if wages were high enough). A complete list of these estimates is available on request. Here, I present a sample of occupation-specific labor markets.

Having all the necessary measures, I proceed to the estimation of system 9 for each of 90 occupational categories and record the estimates of c_{j1} and d_{j1}. For many occupations these were found to be not statistically different from zero. These are plausible values. 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, also d_{j1} should be zero, while in the former, d_{j1} is expected to be one. This property can be used to distinguish between these two cases. Additionally, d_{j1} is expected to be zero 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 positive and finite. The first column of table **??** reports the estimated of elasticities of substitution between college and high school graduates ($\widehat{\sigma_j} = -\frac{\widehat{d_{j1}}}{\widehat{c_{j1}}}$).

The estimated elasticities of substitution can be further used to calculate occupation-time specific relative productivities of college to high school graduates – the measure of skill-intensity of occupations – using equation (10). The point estimates of this measure for years 1984 and 2001 are presented in columns 4 and 5 of table **??**.

6 Applications of the measure of the skill-intensity of occupations

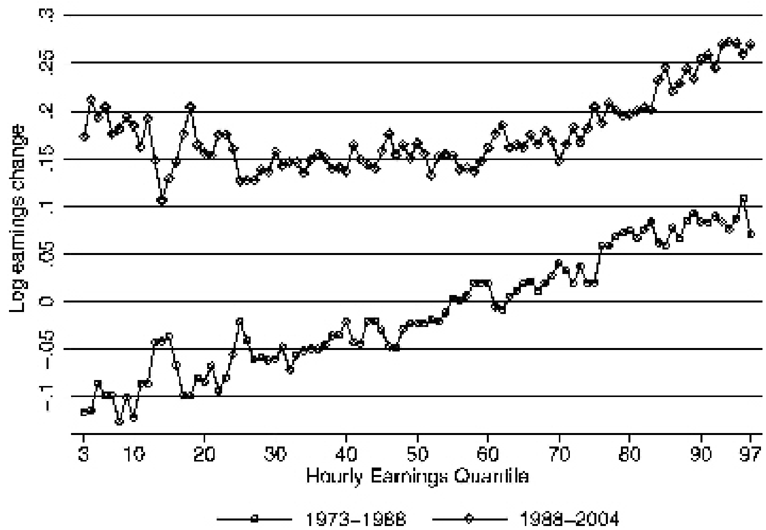
The measure of skill-intensity of occupations derived in this study has multiple applications. This section will presents two of them: testing the modified version of the skill-biased technological change and measuring the fraction of college graduates employed in "noncollege" occupations.

6.1 Testing the modified version of the SBTC

The modified version of the SBTC hypothesis was proposed by Autor et al. (2003, 2006, 2008) to explain the observation of polarization of the labor market in the last decade of the 20th century. This phenomenon is presented in Figure 1. One can see that between 1984 and 1994 real wages grew significantly in the high end of wage distribution, while they almost stayed unchanged in the low end. Between 1991 and 2001, real wages grew in both high and low ends of the distribution, while they remained unchanged in the middle.

Note: Adopted from Autor et al. (2006)

Figure 1: Changes in occupational wages by occupational wage percentile.

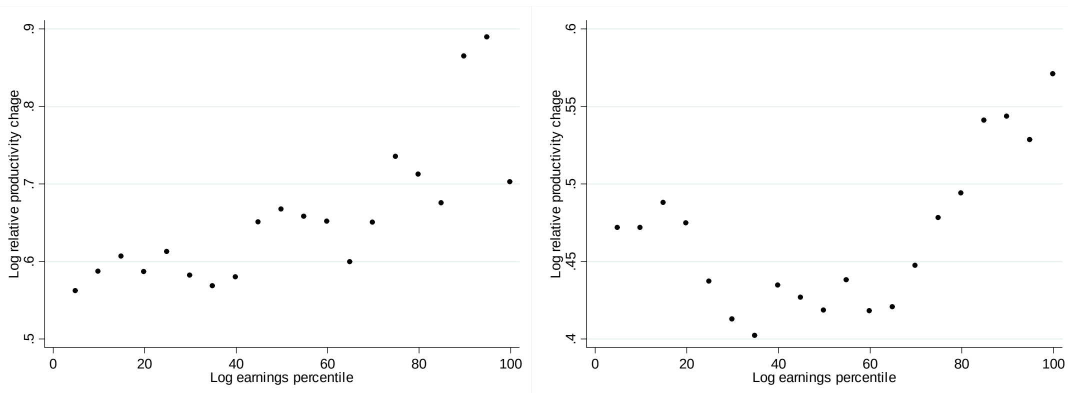


Autor et al. (2003, 2006, 2009) suggest that the observed pattern could be explained by a modified version of the SBTC. It assumes that new technologies complement workers in high-skilled tasks, substitute for them in middle-skilled tasks and are neutral to workers performing low-skilled tasks. This explanation could be tested by measuring the average skill-intensity of occupations characterized by different rates of wage growth. Under the modified SBTC hypothesis, the highly skill-intensive occupations should be growing, the "frozen" occupations should be characterized by average skill-intensity, while occupations with low skill-intensity would be the ones which also experience a wage growth. I use the estimates of skill-intensity of occupations obtained in the previous section to check this hypothesis, what is presented in Figure 2.

Note: Calculations were performed using the sample as described in Section 4.

One could see that Figure 2 mimics Figure 1. This suggests that the skill bias of technology in the 1990's followed the pattern as suggested by Autor et al. (2003, 2006, 2009).

Figure 2: 1984-1994 and 1991-2001 changes in occupational skill-intensity by earnings percentile.



6.2 College graduates in noncollege occupations

Skill-intensity, as defined in this study, provides an alternative measure on the basis of which occupations could be classified into "college" and "noncollege". Using skill-intensity, as compared to the wage premium, as Gottschalk and Hansen (2003) do, allows to recognize whether an occupation values college-gained skills independently from workers ' interest in that occupation. Thus, in turn, avoids misclassifying occupations very popular or very unpopular among college graduates[13] leading to underestimation of the fraction of college graduates working in "noncollege" occupations.

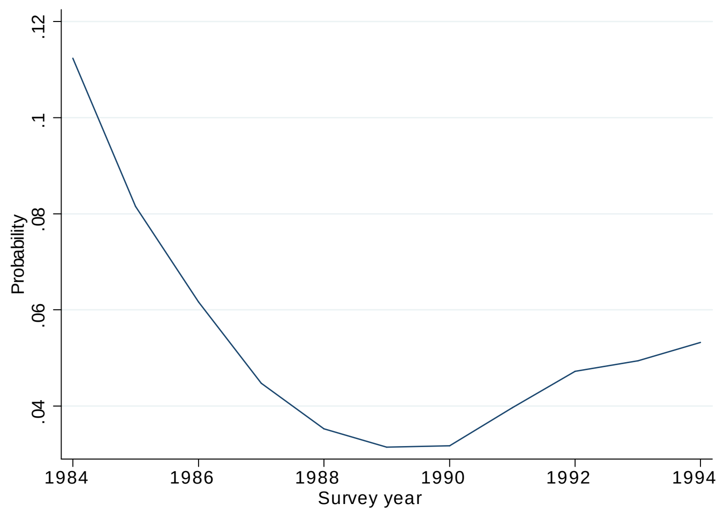
To illustrate how usage of skill-intensity of occupations changes conclusions concerning employment of college graduates in "noncollege" occupations, I replicate the analysis of Gottschalk and Hansen (2003) using their and an mine method for classifying occupations. For the purpose of assuring comparability, I define as "college" all those occupations where college graduates are at least 10% more productive than high school graduates. As explained above, my classification is expected to differ from Gottschalk and Hansen's classification in those occupations which are skewed towards employing one labor type (i.e. they employ many college graduates or many high school graduates).

[13]The very popular occupations could me missclassified as "noncollege", while very inpopular occupations could be missclassified as "college".

This could lead to different conclusions concerning the evolution of the fraction of college graduates working in "noncollege" occupations.

Figures 3 and 4 present a comparison of the two approaches. One can see that basing the analysis on skill-intensity of occupation reveals a different pattern of the evolution of college graduates' propensity to work in "noncollege" occupations.

Figure 3: Probability that a college graduate works in a "noncollege" occupation (1984-1995)



Note:This is a replication of Gottschalk and Hansen (2003) analysis. Slight differences between this and the original figure might result from different sample selection.

Note:This graph was obtained using a classification of occupations into "college" and "non-college" based on occupation-specific skill-intensity.

7 Conclusion

In this study I propose a methodology for determining skill-intensity of occupations. I argue that a good proxy for occupation-specific skill-intensity is the relative productivity

Figure 4: Probability that a college graduate works in a "noncollege" occupation (1984-1995)

of college and high school graduates. When estimating this measure, it is important to take into account occupation-specific elasticities of substitution between the two types of workers, which in many studies is *ex ante* assumed to be infinite. Interestingly, I find that many occupations are characterized by imperfect substitutability between college and high school graduates.

The proposed measure of skill-intensity of occupations has multiple applications. This paper presents two of them. I show that this measure could be used to test the modified version of the SBTC and to analyze the evolution of the fraction of college graduates working in "noncollege" occupations.

References

[1] 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.

[2] Autor D., Katz L., Kearney M. (2006) The Polarization of the U.S. Labor Market, *American Economic Review,* 96(2): 189–194.

[3] 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.

[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] Card D. (2001) Immigrant Inflows, Native Outflows and the Local Labor Market Impacts of Higher Immigration. *Journal of Labor Economics,* 19(1): 22-64.

[6] Card D., DiNardo J. (2002) Skill-biased technological change and wage inequality: some problems and puzzles, *Journal of Labor Economic*s, 20(4): 733-783.

[7] Card D., Lemieux T. (2001) Can falling supply explain the rising return to college for younger men? A cohort-based analysis, T*he Quarterly Journal of Economics*, 116(2): 705-746.

[8] 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.

[9] Gebicka B. (2009) Skill upgrading of occupations and college-degree supply: the case of the Czech Republic, CERGE-EI, *mimeo*

[10] 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

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

[12] Handel M. (2000) Trends in Direct Measures of Job Skill Requirements *Levy Economics Institute Working Paper* No. 301.

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

[14] McGoldrick KM , Robst J. (1996) Gender Differences in Overeducation: A Test of the Theory of Differential Overqualification. *American Economic Review*, 86(2): 280-84.

[15] Meyer P., Osborne A. (2005) Proposed Category System for 1960-2000 Census Occupations, *U.S. Bureau of Labor Statistics Working Paper* No. 383.

[16] Peracchi F., Welch F. (1995) How Representative are Matched Cross-Sections? Evidence from the Current Population Survey. *Journal of Econometrics* 68: 153-179.

[17] 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	$\widehat{\sigma}_j$	$\widehat{\frac{\alpha_{Cjt}}{\alpha_{Njt}}}_{1984}$	$\widehat{\frac{\alpha_{Cjt}}{\alpha_{Njt}}}_{2001}$
Public administration	inf	0.0688	0.1132
Financial managers	?	99	99
Managers, marketing and advertising	inf	0.4101	0.2890
Real estate managers	inf	0.5279	0.4445
Miscellaneous managers and administrators	4.3808	0.0548	0.2267
Accountants and auditors	inf	0.2724	0.3786
Miscellaneous financial officers	?	99	99
Personnel, training, and labor relations specialists	inf	0.2552	0.3056
Purchasing agents and buyers	1.8931	-0.1859	-0.1974
Miscellaneous management-related occupations	0	99	99
Architects	0	99	99
Miscellaneous professional specialty occupations	?	99	99
Engineers, n.e.c.	0	99	99
Electrical and electronic engineers	0	99	99
Mathematical and computer scientists	0	99	99
Biological and life scientists	0	99	99
Health diagnosing occupations	0	99	99
Registered nurses	0	99	99
Health assessment and treating occupations, n.e.c.	0	99	99
Therapists, n.e.c.	0	99	99
Speech therapists	?	99	99
Postsecondary teachers	0	99	99
Prekindergarten and kindergarten teachers	2.5942	0.1442	0.3453
Elementary school teachers	0	99	99
Secondary school teachers	0	0.0881	0.3947
Special education teachers	?	0.8300	0.5612

Occupation group	$\widehat{\sigma_j}$	$\widehat{\frac{\alpha_{Cjt}}{\alpha_{Njt}}}_{1984}$	$\widehat{\frac{\alpha_{Cjt}}{\alpha_{Njt}}}_{2001}$
Counselors, librarians, archivists, and curators	inf	0.2452	0.3751
Social workers	inf	0.1142	0.3267
Recreation and religious workers	1.6399	-0.8768	-0.1551
Lawyers	0	99	99
Judges	0	99	99
Designers	10.1734	0.1953	0.3206
Editors and reporters	0	99	99
Writers, artists, and related workers, n.e.c.	0.4319	-1.8839	-0.5064
Painters, sculptors, and photographers	inf	0.0614	0.3375
Public relations specialists, announcers, and athletes	?	99	99
Clinical laboratory technologists and technicians	inf	0.2632	0.3379
Health technologists and technicians	1.3102	-1.5537	-1.2643
Engineering technologists and technicians	inf	0.1016	0.2087
Drafting occupations, surveying and mapping technicians	inf	0.1364	0.1918
Science technicians	0	99	99
Technicians, n.e.c.	inf	-0.0002	0.2750
Computer programmers	0	99	99
Legal assistants	0.4683	-1.8341	-1.5472
Supervisors and proprietors, sales occupations	inf	0.2758	0.4007
Insurance sales occupations	0.5904	-1.3661	-0.9246
Real estate sales occupations	inf	0.0295	0.1492
Securities and financial services sales occupations	0	99	99
Sales occupations, advertising and other business services	0.7181	-0.8158	-0.1829
Sales representatives, commodities except retail	0.1414	-5.3719	-1.8346
Sales workers, retail	0.3100	-6.6586	-4.8179
Cashiers	0	0	0

Occupation group	$\widehat{\sigma_j}$	$\widehat{\frac{\alpha_{Cjt}}{\alpha_{Njt}}}_{1984}$	$\widehat{\frac{\alpha_{Cjt}}{\alpha_{Njt}}}_{2001}$
Sales-related occupations	0	0	0
Supervisors, administrative support occupations	inf	0.4030	0.1699
Computer equipment operators	inf	0.1023	0.3617
Secretaries	inf	0.0858	0.1136
Stenographers and typists	inf	0.1442	0.1901
Information clerks	0	0	0
Records processing occupations, except financial	2.5496	-0.6468	-0.4734
Financial records processing occupations	0.7010	-3.0858	-2.5941
Administrative support occupations, n.e.c.	37.2995	0.1758	0.2036
Mail and message distributing occupations	0	0	0
Material recording, scheduling, and distrib- uting clerks, n.e.c.	inf	0.2394	0.2812
Insurance adjusters, examiners, and investigators	0.9347	-0.7328	-0.7807
Miscellaneous adjusters and investigators	0	0	0
General office clerks	0	0	0
Service occupations, n.e.c.	inf	0.2160	0.0640
Child-care workers	inf	?	0.2556
Protective service occupations	inf	0.1639	0.3702
Police and detectives	1.4033	-1.0504	-0.7550
Food preparation and service occupations	0	0	0
Waiters and waitresses	inf	0.0192	0.2292
Cooks	0.6865	-4.4781	-3.9778
Dental assistants and health aides	0	0	0
Nursing aides	inf	0.3209	0.3587
Cleaning and building service occupations	inf	0.2910	0.1241
Farm occupations	0	0	0
Agricultural, forestry, fishing, and hunting occupations	inf	0.2465	0.4451

Occupation group	$\widehat{\sigma_j}$	$\widehat{\frac{\alpha_{Cjt}}{\alpha_{Njt}}}_{1984}$	$\widehat{\frac{\alpha_{Cjt}}{\alpha_{Njt}}}_{2001}$
Mechanics and repairers, vehicle and industrial machinery	0.4115	0.0914	0.2728
Other mechanics and repairers	0.2024	0.0596	0.2926
Construction trades, n.e.c.	0	0	0
Carpenters, electricians, and painters	0	0	0
Extractive and precision production occupations	0	0	0
Supervisors, production occupations	0	0	0
Machine operators	0	0	0
Fabricators and assemblers, miscellaneous production	inf	0.1108	0.2103
Transportation and material moving occupations	0	0	0
Handlers and laborers, n.e.c.	0	0	0
Freight, stock, material handlers, and service station	0	0	0

Note: An estimate of zero elasticity of substitution could be a result of having very few observations of one type of workers within occupation.