Who Earns Their Keep? An Estimation of the Productivity-Wage Gap in Hungary 1986-2005

Anna Lovász
Institute of Economics, Hungarian Academy of Sciences

Mariann Rigó
Central European University

Draft of July 14, 2009.

Abstract

In this paper we seek to provide new empirical evidence on the relative productivities and wages of various worker groups (by gender, age, and education), based on longitudinal matched employer-employee data from Hungary covering 1986-2005. We estimate the productivity and wage gaps from firm-level production functions and wage equations, using firm-level data on productive inputs and output, wage costs, and the demographic composition of the work force obtained from the linked worker data. This methodology allows us to assess whether productive differences can account for the wage gaps between worker groups, as well as the evolution of these gaps following the transition to a free market. We take firm fixed effects into account to assess the role of selection at the firm level, and estimate the production function via the method of Levinson and Petrin to account for endogeneity of input choice. The results show that while there may be significant differences in productivities and wages between groups in the OLS specification, these mostly become insignificant within firms. We find that much of the fall in the value of skills obtained prior to the transition is due to selection of workers at the firm level.

The authors would like to thank John Earle, Álmos Telegdy, Bruno Van der Linden, János Köllő, Gábor Antal, Shelly Lundberg, Dick Startz, and participants of seminars at the CEU, IE-HAS, and the UCL Doctoral Workshop for valuable comments and ideas. Funding for this research was provided by the GDN CERGE-EI Regional Research Competition 2009 Grant. The harmonized dataset was provided by the Institute of Economics of the Hungarian Academy of Sciences, and further data cleaning was done by the CEU Labor Project.

I. Introduction

Wage differentials between various groups of workers are commonly estimated using Mincerian earning equations, in which the control variables are typical observable worker characteristics such as education, experience, marital and family status, race, and gender. In this context, a significant female dummy coefficient is often interpreted as evidence of wage discrimination, while the coefficients corresponding to various levels of schooling are assumed to represent the returns to education. However, these interpretations are only valid if all productivity-related differences among different individuals are controlled for, or if the observable control variables are not correlated with unobservable worker characteristics.[1] For example, unobserved systematic differences in productivity between men and women could lead to an overstatement of gender wage discrimination based on the female coefficient, as pointed out by Becker's (1985) theory of effort allocation. If women with equal observable characteristics supply a lower effort on the job than men – due to different responsibilities outside the job, attitudes towards careers, etc. – then the estimated coefficient of the female dummy will overstate the level of wage discrimination, since the observable variables do not fully capture productive differences between men and women.[2]

In this paper, we seek to provide new empirical evidence on the returns to worker characteristics, by separating the productivity-related and other components of the wage gap. We do this by using a methodology that controls for the productive differences between various employee types based on relative productivity estimates that we estimate independently of wages. In this way we are able to assess whether worker groups – as defined by gender, age, and education level – are being compensated according to their

[1] Omitted firm information from the earning regression may also bias estimates if workers sort into firms non-randomly. However, with the increasing availability of matched employer-employee datasets, firm-specific controls or firm fixed effects are usually included among the explanatory variables. Data on worker-level productivity, on the other hand, is still rarely available.

[2] Many other papers study why unobserved systematic productive differences may exist. For example, Bowlus (1997) examines the role of differences in job matching quality, while Ichino and Moretti (2006) study biological productive differences. Another source of bias may be measurement error in the observable worker variables, for example, if potential experience is used to proxy for actual years of labor market experience. Since women are usually absent from the labor market when they have children, this measure may overstate their actual work experience and productivity.

productivity, as in the case of no wage discrimination. We use a large matched employer-employee dataset from Hungary covering the years 1986 to 2005, which is well suited to the analysis due to its sample size, detailed firm and worker variables, and the fact that we are able to follow firms longitudinally. Additionally, the period covered by the data includes the transition to a free market, providing new evidence on transition-related labor market phenomena such as skill obsolescence, and whether market forces in the form of increased competitive pressures led to more efficient wage-setting behavior of firms.

The ideal method for assessing the components of wage gaps between groups would be the estimation of earnings equations including individual-level data on each worker's productivity. In such a specification, a significant coefficient on any worker category dummy (such as the female or university dummy) would suggest a deviation from competitive wages that is not explained by productivity of the given group (such as negative or positive wage discrimination). However, datasets with direct measures of individual productivity are rarely available, limited in size, and not representative of the economy as a whole. Some studies that use direct measures of individual productivity include Foster and Rosenzweig (1993), who use data on piece-rate wages to measure productivity in time-rate work, and some other studies based on performance ratings of workers (Korenman and Neumark 1991, Holzer 1990). Another strand of the literature focuses on finding a suitable measure of worker ability, the control variable most often considered to be correlated with observable worker characteristics such as gender and education, but not observed by the researcher. Various proxies are used to account for the missing information from the earning regression: some studies (e.g. Griliches and Mason, 1972; Griliches, 1977; Neal and Johnson, 1996) use IQ or AFQT test scores to apply a productivity-based ability control instead of the input-based education variable. However, these studies refer only to a small subsample of individuals, and do not attempt to capture other unobserved productive differences besides ability.

The availability of matched employer-employee datasets and panel data on workers has allowed some new methods to emerge. Abowd, Kramarz, and Margolis (1999) developed a measure of human capital that incorporates individual observable (experience, education, sex, race) and unobservable (innate ability, educational quality, social capital, effort, etc.) productivity components. They do this by following both individuals and firms over time and identifying unobservable worker and firm fixed effects. Their measure of human capital is used subsequently by numerous studies, e.g. Abowd, Lengerman and McKinney, 2003; Haskel, Hawkes and Pereira, 2005; or, Iranzo, Schivardi, Tosetti, 2006. Unfortunately, the method applied by Abowd, Kramarz, and Margolis (1999) requires a database that is a panel in both workers and firms, and such databases are still relatively rare.

In this paper, we use a matched employer-employee dataset in a different way: we estimate an independent measure of the relative productivities of different worker groups based on firm-level data on production output and inputs, and the demographic composition of plants' labor forces. Estimation of a production function in which the labor term is augmented with the worker composition of the firm provides us with a measure of the relative productivity of each worker group that is independent of wages. This measure incorporates all productivity-related differences between the various demographic groups; hence, both observed and unobserved productivity components are accounted for.[3]

There are two possible ways in the above production function – wage equation framework to test whether wages deviate from productivities significantly. First, as Griliches (1960) notes and Hellerstein and Neumark (2004) explain, assuming that labor markets operate in a competitive way, the traditional labor input variable can be transferred into efficiency units by using wages as weights for the corresponding worker group. Then, a test of the equality of the coefficients on the labor and on the quality index term serves as an indirect test of the competitive market hypothesis[4]. However, imposing

[3] The difference compared to the previous approach by Abowd, Kramarz and Margolis (1999) is that while they compute a truly individual measure of productive human capital, our approach yields only group-level relative productivities. However, these estimates are well-suited to the assessment of the existence of group-level wage gaps, as they allow us to control for group-level systematic differences in productivity.

[4] For example, if workers are grouped only by gender, and w_F and w_G refer to the wages of female (F) and male (G) employees, then the labor quality term can be expressed as:

$$QL = w_G \cdot G + w_F \cdot F = w_G \cdot L \cdot \left[1 + \left(\frac{w_F}{w_G} - 1\right) \cdot \frac{F}{L}\right].$$

Substituting this into the Cobb-Douglas production function gives:

the competitive market assumption in advance will likely yield false test results, as any measurement error of the labor quality index will bias the coefficients of the other variables in an unpredictable way.[5] For this reason, we use an alternative method here, in which we obtain separate estimates of the relative productivities and relative wages without a priori imposing the competitive market hypothesis. This has the added benefit of allowing us to examine directly the estimated relative productivities and wages of the worker groups, which we feel are of interest in themselves.

This approach pioneered by Hellerstein and Neumark (1999), and Hellerstein, Neumark and Troske (1999) has been used by numerous other studies on data from various countries.[6] To test whether different worker groups are paid according to their productivity, we compare the estimated productivity differential to each group's wage differential obtained from an earning regression. A discrepancy between wage and productivity gaps may correspond to various phenomena, e.g. discrimination, efficiency wages, or the existence of compensating wage differentials. While testing the equality of wages and productivity is not necessarily a test for perfect competition (or for the presence of discrimination), this methodology does allow us to assess more accurately the contribution of productive differences between groups to the wage gap. In the context of the Hungarian data, one would expect that the discrepancy becomes smaller as competitive forces get stronger following the transition[7], leading to a decrease of the gap

---

$$\ln Y = \alpha \ln K + \beta \ln M + \gamma w_G + \gamma \ln L + \gamma \ln\left[1 + \left(\frac{w_F}{w_G} - 1\right) \cdot \frac{F}{L}\right].$$

Testing the equality of coefficients on the labor and on the quality index term, one can indirectly test whether relative wages are equal to relative marginal products.

[5] For example, if the mismeasurement of the labor quality term biases its own coefficient and that of the labor term in opposite directions, we may falsely reject the equality of the two. Further, labor quality may differ along multiple dimensions, for example, occupation, which makes the interpretation of the test even more difficult.

[6] See also Zhang and Dong (2009), Kawaguchi (2007), Ilmakunnas and Maliranta (2003), Van Biesebroeck (2007), Dostie (2006), Hellerstein and Neumark (2004), and Deniau and Perez-Duarte (2003). Frazer (2007) advances the method of Hellerstein and Neumark (1999) further by trying to obtain precise estimates of productivity and wage returns separately, not only of the gap between the two measures. He does this by identifying a firm-level ability component from the production function and plugging this ability measure into the firm-level wage equation.

[7] The level of competition facing Hungarian firms increased significantly and rapidly during the time period covered by the data. The number of registered economic organizations increased from 391 thousand to 1.1 million by 1998, and eighty percent of the GDP was produced by the private sector. The mean of the concentration ratio (measured by the Herfindahl Index at the three-digit level) fell from .34 to .16.

between productivities and wages.[8] Previous studies on Hungary have underlined the importance of the process of skill obsolescence, in which skills acquired before the transition lose value relative to new skills.[9] Our aim is to extend these results using the Hellerstein and Neumark methodology for the years 1986-2005, paying special attention to the estimation of the production function for more detailed worker categories.

The characteristics of our dataset allow for several improvements over previous studies that use the same approach. The large sample size – the final regression sample includes 67,928 firm-years and 1,245,577 worker-year observations – and the relatively large sample of workers from each firm allows us to include more detailed worker categories, and to estimate the production functions and wage equations based on these groups with fewer restrictions than previous studies.[10] Since we do not observe every worker at each firm, we do not have a true measure of workforce composition, but the random sampling design based on the birthdates of workers does yield a more representative sample of workers than many other studies, where the number of workers sampled from each firm is limited regardless of firm size.[11] The data contains detailed firm and worker variables, allowing for more accurate estimation of the production function and wage equation,[12] and several robustness checks. Most importantly, the dataset enables us to follow firms over time, so we can control for bias due to omitted firm characteristics that may also be correlated with workforce composition by taking firm fixed effects into account. To our knowledge, very few previous studies used a database that followed firms over a relatively long time period to assess the importance of

[8] Lovász (2008) found evidence supporting this in post-transitional Hungary: an increase in competition at the industry level (measured by market concentration, import penetration and export share) led to a fall in the within-firm endowment-adjusted wage gap between men and women.

[9] Köllő and Kertesi (2002) use the same Hungarian data for 1986-1999 to study skill obsolesence following the transition. They estimate the returns to skills of detailed worker groups based on gender, education, and experience, and find that after 1993, the return to skills from the pre-transitional era decreases, while the returns to new skills increase. They also estimate the effect of worker composition on firm-level productivity for unskilled, skilled young, and skilled old workers, and find that the changes in relative wages reflect the relative productivities of the groups.

[10] Hellerstein, Neumark (1999) has information only on the gender composition and the occupational distribution of the workforce, but they do not have the occupational distribution by gender, which necessitates the use of a restricted labor quality aggregate.

[11] For example, Van Biesebroeck (2007) observes a maximum of 10 workers within plants; and Dostie (2006) has worker data on maximum 12 employees within firms.

[12] For example, Dostie (2006) has no information on firms' capital stock, and uses industry average capital stock as a proxy for the individual firm's capital stock.

selection of workers at the firm level[13]. Hence, we feel that our study adds valuable insight to both the existing international literature that uses Hellerstein & Neumark methodology, and to previous empirical studies assessing the changes in the labor market position of different worker groups in Hungary following the transition.

The remainder of the paper is organized as follows: in section II. we describe our model and estimation strategy, including estimation issues; section III describes the sampling design of the dataset and the relevant variables; in section IV we present the estimation results and interpretation for the pooled years, as well as for separate time periods to assess the changes and several robustness checks; section V concludes.

II. Methodology

Our approach is to estimate the relative productivities and wages of worker groups at the firm level, and check whether they differ significantly for any group.[14] Following the idea of Griliches (1957, 1970), the motivation for the estimation procedure can be based on a simple model with two types of perfectly substitutable labor and a general production function:

$$Y = f(L_1 + \varphi \cdot L_2)\,, \qquad (1)$$

[13] Hellerstein and Neumark (1998) provide within-firms estimates, but they use only two years of data. Vanderberghe and Waltenberg (2009) are using similar methodology for Belgian data covering 1998 to 2006 to assess the effect of ageing an workforce on productivity and wages.

[14] A less rigorous approach to estimating the contribution of worker characteristics to firm-level productivity is to augment a labor productivity regression with firm-level worker characteristics. In this case, the following general equation is estimated:

$$\ln\left(\frac{Y}{L}\right)_{jt} = \alpha_0 + \sum_{n=1}^{N} \phi_n \frac{L_n}{L} + \delta \cdot Z_{jt} + u_{jt}$$

Basically, this can be considered a restricted version of our production function specification with *K*, *M* ommitted and $\gamma = 1$. Studies using this method do not aim to estimate relative productivities. Rather, they want to see if firm-level average worker characteristics have a significant impact on the productivity of the firm. For example, Haltiwanger, Lane and Spletzer (1999 and 2007) investigate the connection between labor productivity and the composition of the firm's workforce described by gender, age, education, and the ratio of foreign born employees. A similar concept is applied in the studies of Malmberg et al (2005) and Lallemand and Rycx (2009). These studies use a log-log specification with log labor productivity regressed on the log of firm-level average worker characteristics. In this case, the estimated coefficients of the share variables can be interpreted as elasticities, but no precise estimate of relative productivity can be inferred.

where Y is firm output and L_1 and L_2 are the number of employees in group 1 and 2. The parameter φ reflects the marginal productivity of type-2 employees relative to type-1 employees: $\varphi = MP_2/MP_1$. In a competitive spot labor market where firms are maximizing their profits, the marginal product of each type of labor equals its wage. Introducing the notation λ for relative wages: $\lambda = w_2/w_1$, the following relationship should hold: $\varphi = \lambda$. Hence, if firms are behaving efficiently, there is no wedge between the relative wages and relative productivities of different types of employees: each type receives remuneration according to its productivity. In the case of $\varphi \neq \lambda$, the profit-maximizing firm will arrive to a corner solution: hire only the cheaper type of labor. Empirical evidence showing that $\varphi \neq \lambda$ is inconsistent with the assumption that we are observing profit-maximizing, or cost-minimizing firms in a competitive spot labor market.

Though our aim is not to test the existence of perfect competition in the Hungarian labor market, the theoretical relationship between relative marginal products and relative wages serves as a framework to study the relationship between the two measures. Results indicating that we can reject the equality of φ and λ support alternative hypotheses, such as taste-based discrimination or the existence of efficiency wages, but may also fit into the competitive market hypothesis if certain workers are rewarded by compensating wage differentials or if there are short-term frictions on the market.

II.1. Production function

Production function estimation has a burgeoning literature. Despite the growing number of studies, there is no consensus yet as for the proper approach. Since Marshak and Andrews (1944) a vivid discussion emerged about the treatment of unobserved productivity shocks[15], and, more recently, the impact of worker characteristics on firms' productivity generated a large number of interesting studies. The main methodological issues centering on the production function estimation are (1) correct measurement of the input and output variables, especially of the labor input, (2) treatment of unobservable

[15] Recent studies include e.g. Olley and Pakes (1996), Levinsohn and Petrin (2003), Ackerberg, Caves and Frazer (2005).

productivity shocks, and (3) the functional form assumption of the production function, which is mostly a choice between the Cobb-Douglas and the translog case.[16] Since our aim is to compare the relative productivities of different groups with their relative wages, we devote special attention to the first issue, measuring labor input. We also perform robustness checks using methods that correct for the endogeneity bias inherent in production function estimations, and that allow for the possibility of imperfect substitution among worker types.

The standard, 3-input Cobb-Douglas production function is of the form:

$$\ln Y = \ln A + \alpha \ln K + \beta \ln M + \gamma \ln L \qquad (2)$$

where A is a technology parameter, Y is output, which may be proxied by value added or sales; and the inputs are capital, material costs, and labor (K, M, and L, respectively). In our database we have accounting data on sales, different profit measures (before-tax profit, after-tax profit, operating profit), end-of-year tangible assets, material costs and the average number of employees. As a capital measure we use deflated year-average tangible assets. Regarding the output measure, we have several options. As we do not have data on physical output, we proxy output by deflated sales revenue in our baseline specification, but we check the robustness of our results by using the deflated value-added specification as well.[17] To mitigate the problem of differential market prices, which are incorporated both in the sales and value-added data, we include industry and year dummies among the control variables. This may be especially relevant if certain

---

[16] In our specification, we use the traditional Cobb-Douglas specification of the production function. The translog alternative offers a more flexible specification, as it includes higher-order polynomial terms of the productive inputs, most often a second-order approximation including all possible interactions of the productive inputs besides the Cobb-Douglas controls. Results, as reported e.g. by Hellerstein, Neumark, Troske (1999), Hellerstein, Neumark (2004), or Dong et al (2009), tend to be robust between the two functional form specifications.

[17] Hellerstein, Neumark and Troske (1999) argue that the value-added specification avoids estimating the coefficient of the most endogenous variable, material input. However, in our database, sales data is considered to be more reliable than the profit measures, so this will be our baseline measure.

industries are concentrated and there are differences in market power among firms (Melitz, 2000).[18]

As our aim is to analyze relative productivities, the measurement of the labor input deserves special attention. The standard measure of labor input is the number of employees at the firm. However, Griliches' early work (Griliches 1957) on the specification biases in production function estimation brought attention to the possibility that labor may be more precisely captured by a labor quality variable, which more accurately describes the productive capabilities of the workforce of each firm. To measure this, he suggests the use of a multiplier, which transforms the traditional labor input into "equivalent effective labor units", since neglecting the inclusion of labor quality variable may bias the returns to other inputs.[19]

The most straightforward way to relate worker characteristics to productivity is based on a generalized version of the theoretical model introduced in the previous section. Workers are grouped into $n = 0, 1, \ldots, N$ categories, and instead of L, the total number of employees, a labor quality variable, QL is used in the production function:

$$QL = \sum_{n=0}^{N} \varphi_n L_n \qquad (3)$$

where L_n is the number of employees in group n and the φ_n coefficients are economy-wide productivities of employees in group n[20]. QL can be interpreted as the productivity-adjusted sum of employees.[21] Applying simple algebra, QL can be expressed as the product of L and a worker-quality term with relative productivities:

---

[18] On the other hand, if firms within the same industry produce differentiated products, then firm-level prices will be different from the industry-level price index, implying that the residuals will incorporate to some extent also the pricing policies of the firms (Melitz, 2000).

[19] Recent empirical results (e.g. Fox and Smeets, 2007; Hellerstein and Neumark, 2004) do not show large biases of returns to productive inputs when estimating production functions excluding the labor quality variable, hence, the standard Cobb-Douglas production function is suitable for studying many research questions. However, in our research context, including the labor quality variable into the production function yields the appropriate framework.

[20] As pointed out in the Introduction, the φ_n coefficients could be replaced by the wages of the corresponding worker group, assuming in advance that labor markets are perfectly competitive. However, as we argued, imposing a priori the competitive market assumption may easily lead to false test results. Hence, we are estimating relative productivities independently of relative wages.

[21] This additive formulation assumes that different types of workers are perfect substitutes. As Hellerstein, Neumark and Troske (1999) note, the perfect substitutability among all types of workers may be a strict assumption in certain cases; one example is the case of unskilled and skilled employees. In a later section we will talk about the possible ways to relax this assumption. However, as results are most straightforward

$$QL = \sum_{n=0}^{N} \varphi_n L_n = \varphi_0 L_0 + \sum_{n=1}^{N} \varphi_n L_n = \varphi_0 L \left[1 + \sum_{n=1}^{N} \left(\frac{\varphi_n}{\varphi_0} - 1 \right) \frac{L_n}{L} \right] \tag{4}$$

where L is the sum of employees over all categories: $L = \sum_{n=0}^{N} L_n$. Using the labor quality augmented production function, we can identify N relative productivity parameters, taking $n = 0$ as the reference group. In this case, the following equation (5) can be estimated:

$$\ln Y_{jt} = \alpha_0 + \alpha \ln K_{jt} + \beta \ln M_{jt} + \gamma \ln \varphi_0 + \gamma \ln L_{jt} + \gamma \ln \left[1 + \sum_{n=1}^{N} \left(\frac{\varphi_n}{\varphi_0} - 1 \right) \frac{L_{n_{jt}}}{L_{jt}} \right] + \delta \cdot Z_{jt} + u_{jt}$$

where Z includes additional controls that may determine a firms' productivity. Since relative productivities do not enter linearly, the equation can be estimated via nonlinear least squares.[22]

Workers need to be assigned to groups along any characteristic that could potentially differ in productivity. These groups are typically based on gender, race, education, age or experience, marital status, and occupation. The Hungarian WES does not contain information on race or marital status, so we group workers into categories

to interpret in the perfect substitution framework, we will first introduce the econometric methodology for the case of perfect substitutes.

[22] A less rigorous approach is to include a worker-quality term linearly into the log production function. In this case, the proportions of workers in different groups are treated as separate inputs besides K, M, L and the estimable production function is:

$$\ln Y_{jt} = \alpha_0 + \alpha \ln K_{jt} + \beta \ln M_{jt} + \gamma \ln L_{jt} + \sum_{n=1}^{N} \phi_n \frac{L_n}{L} + \delta \cdot Z_{jt} + u_{jt} .$$

In this case the ϕ_n terms cannot be interpreted as relative productivities, they only show the importance of worker characteristics in the production process. However, if $\sum_{n=1}^{N} \left(\frac{\varphi_n}{\varphi_0} - 1 \right) \frac{L_n}{L} < 0.1$, this can be considered as the linear approximation of (5), with $\gamma \left(\frac{\varphi_n}{\varphi_0} - 1 \right) = \phi_n$. Hence, relative productivities can be inferred by dividing the worker-quality coefficient with the coefficient of the labor (L) term. Though the linear version yields a convenient alternative to (5) to estimate relative productivities, its use is restricted by the sum-condition. The linear version is used to compute relative productivities by Dostie (2006) and Ilmakunnas and Maliranta (2003).

based on gender, age (young if age<31, middle if between 30 and 50, and old if over 50), and education (university, secondary, or elementary). This gives us a total of 18 worker categories. Alternatively, we can also use information on occupation, but the number of workers in each worker group cell within each firm becomes too small in some cases, so we avoid grouping workers both along occupation and education in our main specification.

Restrictions on the quality of labor term

As grouping workers into detailed categories often requires estimating a large number of productivity parameters, in most studies two restrictions are usually applied to the labor quality term.[23] First, the number of coefficients to be estimated can be reduced by assuming that relative productivities are constant across other categories. This means that, for example, the gender productivity gap is the same among young, middle-aged and old employees; or, the productivity ratio between young and middle-aged workers is the same among male and female employees, etc. Though in certain cases this assumption may be too restrictive (e.g. gender gaps are probably different in the various occupational categories; or, the returns to education may be different among the different age groups), the same framework is widely applied in the earning regression context when using standard Mincerian earning regressions without interactions[24].

The second restriction assumes that the proportion of workers is constant across other categories (e.g. the proportion of female employees is the same in each age category). This is mostly necessary if the proportion of workers in each group cannot be estimated accurately due to a low percentage of sampled workers for each firm. With the imposition of both restrictions, the number of parameters to be estimated in our specification decreases from 18 to 5. We then estimate the relative productivity of

[23] For example, if worker groups are defined by gender, 3 education categories (university, secondary, or elementary), 3 occupation categories (unskilled, skilled, professionals), and 3 age categories (Y, M, O), we would have 54 groups of employees, which implies the estimation of 53 relative productivity parameters in the unrestricted equation.
[24] As a robustness check, Hellerstein and Neumark (2004) relax the equal relative productivity assumption regarding marriage, race and gender. They refer to empirical evidence that the marriage wage premium and the race differential is larger for men than for women.

females compared to males (φ_F), middle aged and old compared to young (φ_M and φ_O), and university or secondary compared to elementary schooling (φ_S and φ_U), according to the following restricted equation:

$$\begin{aligned}\ln Y_{jt} &= \alpha_0 + \alpha \ln K_{jt} + \beta \ln M_{jt} + \gamma \ln \varphi_{MaYE} + \gamma \ln L_{jt} + \\ &+ \gamma \ln\left[1 + (\varphi_F - 1)\frac{L_{F_{jt}}}{L_{jt}}\right] + \gamma \ln\left[1 + (\varphi_M - 1)\frac{L_{M_{jt}}}{L_{jt}} + (\varphi_O - 1)\frac{L_{O_{jt}}}{L_{jt}}\right] + \\ &+ \gamma \ln\left[1 + (\varphi_S - 1)\frac{L_{S_{jt}}}{L_{jt}} + (\varphi_U - 1)\frac{L_{U_{jt}}}{L_{jt}}\right] + \delta \cdot Z_{jt} + u_{jt}\end{aligned} \tag{6}$$

Hellerstein, Neumark (1999, 2004), Hellerstein, Neumark and Troske (1999) and Van Biesebroeck (2007) use both assumptions, while Dong et al (2008) use only the assumption of equal relative productivities. In case of $n_1 \cdot n_2 \cdot ... \cdot n_J$ categories, both the one- and two-assumption case reduces the number of relative productivities to be estimated from $n_1 \cdot n_2 \cdot ... \cdot n_J - 1$ to $n_1 - 1 + n_2 - 1 + ... + n_J - 1$. However, applying only the constant relative productivities assumption supposes precise estimation on more detailed worker categories (e.g. proportion of female, old employees, etc.), while having both assumptions requires estimation of rough categories (e.g. proportion of female employees, proportion of old employees, etc.), which may be more precisely estimated from survey-type data.

In most cases, data limitations necessitate the use of restrictions on the *QL* term. Hellerstein, Neumark and Troske (1999) argue that using only a sample of workers matched to the plant, they are not likely to get accurate estimates on the proportion of workers defined by their 192 categories[25], and as a benchmark case they use both the equiproductivity and equiproportion assumptions[26]. However, we believe that our Hungarian database offers the potential to loosen some of these constraints and estimate a more flexible model. Our database surveys on average 9 percent of workers within firms, with 5,282 as the maximum number of workers sampled within a firm. Out of the more than 40,000 firms included in the database, 1,416 firms have information on more than 50 employees. Hence, our data make it possible to include a less restricted form of labor

[25] Hellerstein, Neumark and Troske (1999) distinguish workers by sex, race, marital status, age, education and occupation.
[26] As a robustness check, they relax some of the assumptions, but results are robust in most specifications.

quality aggregate in a subsample of larger firms, letting both the ratio and the relative productivity differ across the worker groups.

Unobserved productivity shocks

Perhaps the most sensitive issue of production function estimation is how unobserved productivity shocks are tackled. In the case of unobservable input variables that are correlated with productivity shocks, simple OLS estimates will be biased. We apply two methods to tackle the endogeneity issue: (1) include firm fixed-effects, and (2) apply the „proxy variable approach" developed by Levinsohn and Petrin (2003).

Specification of the production function with firm fixed-effects controls for any time invariant firm-specific productivity shocks. This tool has the potential to remove a substantial amount of bias, but its application is restricted only to productivity shocks that are constant across time. Though fixed-effect panel estimation is widely used in production function estimation without worker characteristics (e.g. Brown, Earle and Telegdy, 2006), it is rarely applied to this type of estimation. The estimated relative productivity is a within-firm differential in this case, identified by changes in the group shares within firms over time, rather than differences between firms. The firm fixed effects method allows us to separate observed productive differences into the part that is due to selection of workers into high or low productivity firms, and productivity differences within firms.

The second method, which was developed by Olley and Pakes (1996) and further by Levinsohn and Petrin (2003), includes nonparametric proxies for unobservable productivity shocks. Olley and Pakes (1996) use investment to proxy the unobserved productivity component, while Levinsohn and Petrin (2003) suggest using intermediate inputs (material costs, energy) as a proxy. Using intermediate inputs as a proxy may be more advantageous if there are many observations with missing or zero investment. This approach is used by Hellerstein, Neumark (2004) and Dostie (2006). In this paper, we apply the nonparametric technique by using material costs to proxy unobserved productivity shocks. Levinsohn and Petrin (2003) suggest a two-stage approach to obtain consistent estimates of the input coefficients. Separating the original error term u_{jt} into an unobserved productivity component ω_{jt} and a pure noise parameter e_{jt}, consistent estimate of the labor quality terms can already be obtained in the first stage by estimating:

$$\ln Y_{it} = \gamma \cdot \ln QL_{jt} + \sum_{p=0}^{3}\sum_{q=0}^{3-p} \varepsilon_{pq} \cdot (\ln K_{jt})^{p} (\ln M_{jt})^{q} + \delta \cdot Z_{jt} + e_{jt} \qquad (*)$$

where the polynomial term is a third-order Taylor approximation of the expression:

$$\phi_t(\ln K_{jt}, \ln M_{jt}) = \beta_0 + \alpha \ln K_{jt} + \beta \ln M_{jt} + g(\ln K_{jt} \ln M_{jt})$$

The function $g(.)$ is used to proxy the unobserved productivity component.

II.2. Earnings Equation

We now turn our attention to the second part of the estimation procedure, the estimation of the relative wages of the worker groups. Relative earnings can be estimated either at the worker level, using Mincer-type earnings equations, or, similarly to the production function, one can take a structural approach using firm-level variables. For example, Hellerstein, Neumark (1999), Hellerstein, Neumark, Troske (1999) and Van Biesebroeck (2007) estimate structural earning equations; while Dostie (2006) analyzes relative wages on individual data. The advantage of individual earning equations is that individual unobserved heterogeneity can be controlled for as well. On the other hand, estimating earning equations in similar fashion to the production function, makes the estimated productivity and wage differentials directly comparable, and the simultaneous model minimizes the impact of the unobserved shocks on productivity and wages. Moreover, at the firm-level, one can estimate the production function and the earning equation jointly, allowing the error terms to be correlated, and obtaining more efficient estimates. Besides, joint estimation has the benefit of making the hypothesis testing of equal relative productivities and relative wages straightforward. An additional advantage of estimating firm-level wage equations in our case is the opportunity to use different wage bill measures as our dependent variable.

We can use two different compensation measures as our dependent variable in the firm-level analysis: the firm's annual wage and salary bill, or the weighted sum of the individual wages of the employees surveyed within the firm. These measures may differ in two ways. The wage bill contains the employer's contribution payments and all remuneration-related non-wage expenses, while the individual wages do not. Second, the wage bill is based on the actual composition of the workforce, while the summed individual wage measure is based on the sample of workers.

The firm-level wage equation can be considered a definitional equation, aggregating individual-level equations over all workers. To see this, let us consider the case when workers are grouped by gender only into two categories. An individual wage equation including only gender controls would be:

$$w_i = w_{Ma} \cdot Ma_i + w_F \cdot F_i \qquad (8)$$

where *Ma* and F are dummy variables for males and females. As we are assuming that wages are identical in each unique worker category, w_{Ma} and w_F are the average wages for employees in the male and female groups[27]. Summing up for all employees, gives the total wage bill of the firm:

$$w = w_{Ma} \cdot L_{Ma} + w_F \cdot L_F \qquad (9)$$

where L_{Ma} and L_F are the number of employees in the corresponding groups. The total wage bill is the weighted sum of the employees' individual wages, where the weights are the number of employees in each cell.

Using simple algebra, choosing male employees as the reference category, and introducing the notation λ for relative wages and including similar controls as in the production function yields the estimable equation:

$$\begin{aligned} \ln w_{jt} &= \alpha_0 + \alpha \ln K_{jt} + \beta \ln M_{jt} + \gamma \ln \lambda_{Ma} + \gamma \ln L_{jt} + \\ &+ \gamma \ln \left[1 + \left(\frac{\lambda_F}{\lambda_{Ma}} - 1 \right) \frac{L_{F_{jt}}}{L_{jt}} \right] + \delta \cdot Z_{jt} + u_{jt} \end{aligned} \qquad (10)$$

In a more general form, with any number of worker types, equation (11) is estimated:

$$\ln w_{jt} = \alpha_0 + \alpha \ln K_{jt} + \beta \ln M_{jt} + \gamma \ln \lambda_0 + \gamma \ln L_{jt} + \gamma \ln \left[1 + \sum_{n=1}^{N} \left(\frac{\lambda_n}{\lambda_0} - 1 \right) \frac{L_{n_{jt}}}{L_{jt}} \right] + \delta \cdot Z_{jt} + u_{jt}$$

[27] More precisely, the plant-level wage equation assumes that after controlling for industry, region, firm size and year, wages of workers in the same category are identical.

Worker groups and restrictions are defined in the same way in the earning equation specification as in the quality of labor term of the production function.

Capital and material costs may or may not be included in the plant-level earning equation. Hellerstein, Neumark (1999, 2004) exclude these productive inputs from the firm-level controls, while Hellerstein, Neumark and Troske (1999) and Van Biesebroeck (2007) use capital in the firm-level earning regression. The inclusion of capital and material costs in the plant-level earning equation may serve to control for unobserved worker ability, as there may be complementary relationship between capital and unobserved skills (Hellerstein, Neumark, Troske, 1999). We include the inputs as controls in our baseline specification. However, we find that the results do not change significantly if we exclude them in our specification.

Estimation strategy

In our baseline specification, we estimate equation (6) and the analogous firm-level wage equation grouping workers along gender, age and education, including year, region and industry dummies, ownership variables, and, in certain specifications, firm fixed-effects in Z as the control variables. We then relax the restrictions on the quality of labor term, and estimate jointly equations (5) and (11) in which separate relative productivities and wages are estimated for all 18 detailed categories. This is the most flexible formulation of the labor quality term, as it allows for differences in productivity and wages among more specific worker group types.

Both the restricted and the unrestricted equations will be first estimated by NLS. In this case, relative productivities/wages are identified using the between-firm variation of the corresponding variables. As a second scenario, firm fixed-effects will be included in the production and earning functions, and the equations will be estimated in a first-differenced form. This step corresponds to the within-firm identification strategy, removes selection into high/low productivity/wage – firms and controls for time-invariant unobserved productivity shocks. To account for time-variant unobserved productivity shocks, we follow the method suggested by Levinsohn and Petrin (2003). We obtain consistent estimate of the labor quality term in the production function by estimating equation (*).

Our preferred specification of the production function takes care of both firm fixed-effects and time-variant unobserved productivity shocks. In this case, we will estimate jointly the production function (*) with the firm-level wage equation in a first-differenced form. The joint estimation allows for correlation of the error terms of the two equations. The joint estimations will be followed by the hypothesis testing of equal relative productivities and wages.

To see if the results are stable across time, we also estimate the equations on a sample divided into three different periods. This relaxes the assumption that relative productivities (and the coefficients of the other inputs) are constant across time, which may be too strict as the Hungarian economy underwent significant changes following the transition. In this specification, the first period covers the pre-transitional and transitional years (1986-1993), the second period is the time of major reforms (1994-1999), while the third period covers a few years before the EU accession and the early union-years (2000-2005). A comparison of the relative productivities and wages (discussed in the next section) over time reveals whether competitive forces led to a decrease in wage differentials that are not due to productive differences.

One critical assumption of the model outlined above is that different types of workers are perfect substitutes. As we have already mentioned, this may be too restrictive in certain cases, for example, among skilled and unskilled workers. One way to relax the assumption of perfect substitution is to include the different types of workers as separate inputs in log form. Relaxing the assumption between two types of workers, and leaving the perfect substitutability within these groups, the following form of the production function can be estimated:

$$\ln Y_{jt} = \alpha_0 + \alpha \ln K_{jt} + \beta \ln M_{jt} + \gamma_1 \cdot \ln QL_{1_{jt}} + \gamma_2 \cdot \ln QL_{2_{jt}} + \delta \cdot Z_{jt} + u_{jt} \qquad (7)$$

Hence, within group 1 and group 2, workers are perfect substitutes, but this does not hold between group 1 and group 2. The *QL* terms are of the same form as earlier. Checking if relative productivities within the groups have changed gives a test for the reliability of the perfect substitutability assumption.[28] As a robustness check, we allow for imperfect substitution between workers based on education level: in the above specification, group 1 refers to workers with a university-level education, and group 2 refers to lower education levels.

As a last set of robustness checks, we will examine if the estimated coefficients are stable across different categories of the firms. On the one hand, there may be differences in the productivities of worker groups in different industries. To check for this, we split the sample into four broad industry categories (agriculture, heavy industry, light industry, services), and estimate relative productivities for each separately. Another issue pertains to the estimation of worker composition of firms based on the sampled workers: since the sampling of workers is based on birth dates, the ratios are estimated more accurately for larger firms who will have a larger number of workers sampled consistently. Thus we run spate regressions to see if results are robust on a sample of large firms (defined as having more than 90 employees).

III. Data and Sample

The Hungarian Wage and Employment Survey is available from the National Employment Office for the years 1986, 1989, and 1992-2005. The sample frame includes all full time workers from tax-paying legal entities with double-sided balance sheets that employed at least 20 employees in 1986, extended to firms with at least 10 workers in 1995, and from 1999 on to micro-firms as well. In 1986 and 1989, workers were selected into the sample using a random design based on fixed intervals of selection, with every seventh production worker and every fifth non-production worker selected in 1986, and every tenth worker regardless of type selected in 1989. Starting from 1992, workers were selected into the sample based on their date of birth: production workers were included if their birth date fell on either the 5th or the 15th of any month, and non-production workers if it fell on the 5th, 10th, or 15th of a month.

[28] Hellerstein, Neumark and Troske (1999) check the robustness of their estimates dropping the perfect substitutes assumption between production and non-production workers. Van Biesebroeck (2007) allows for imperfect substitution between workers with low and high levels of experience.

The WES includes demographic information for this random sample of workers, matched to detailed characteristics and balance sheet information of the firms where they are employed. Worker variables include the gender, age, highest education level (five categories: less than 8^{th} grade, elementary, high school, vocational, university), and occupation (4 digit occupational code). For the purposes of defining the various worker groups, we define three age categories (under 30, 31-50, over 50), three education categories (university, secondary, or elementary), and use gender. The firm variables used in the estimation are firm output, capital, material costs, employment, wage bill, industry, region, size, and ownership.

The sample used in the production and wage differential estimation is restricted in a few ways. Only firms from the private sector are included. For all years of the data, we include only firms with at least 20 employees, to preserve consistency. To be able to estimate the ratios of employees within each demographic group, and to ensure a representative sample, we include only firms in which at least 5 percent of the total workers employed are included in the WES worker data. The resulting sample includes observations on 67,928 firm-years and 1,245,577 worker-years. Table III.1. gives the summary statistics of the firm-level sample.

IV. Results

We now present the results of the relative productivity and wage estimation. First, we discuss the restricted results (Tables 2 through 5), in which separate estimates are obtained for females relative to males, skilled relative to unskilled, and middle-aged and older aged compared to young workers for both the pooled years and time period cases. We then turn our attention to the unrestricted results (Tables 6 to 8), which allow for differences within different groups. We compare our results to both existing international estimates that use the Hellerstein and Neumark methodology, as well as with an eye on previous studies on Hungary, which mostly focus on the phenomenon of skill obsolescence following the transition.

Restricted results

A quick look at Tables 2 and 3 shows that relative productivity estimates obtained with and without Levinsohn and Petrin (2003)'s polynomial term are qualitatively the same, indeed they mostly differ only in the second or third digit in the within-firm specification. This implies that time-variant unobserved productivity shocks are not a major issue in the production function estimation. However, as the comparison of between-firm and within-firm specification reveals, selection of workers along firm-level unobservables plays a much larger role. Hence, when interpreting the results, we will mostly focus on our starting NLS specification and on our most preferred within-firm specification including Levinsohn and Petrin (2003)'s polynomial term (FDLP). In our most preferred FDLP specification, the production function and the wage equation is estimated jointly and the hypothesis testing of equal relative wages and productivities is presented.

Gender

Previous international empirical results based on cross-sectional estimates usually document a negative female – male wage gap, and also a negative association between firm-level productivity and the proportion female within the firm. The female-male relative productivity is mostly estimated in the range of 0.7 – 0.9 and female wages are usually 15 – 40 percent less than the wages of male employees.[29]

Table 2 shows the results of the restricted specifications on the pooled sample encompassing all the years of our data. The base specification, estimated with nonlinear least squares (NLS) which reflects between-firm effects, shows that the relative

[29] For example, Hellerstein, Neumark and Troske (1999) and Hellerstein, Neumark (2004) using a US database from 1990 found that the productivity of women is 0.85 – 0.87 that of the men, however, relative wages are even lower: women receive 32-38 percent less than men. Their results point to a negative wage – productivity gap, which may be interpreted as wage discrimination. Hellerstein, Neumark (1999) on Israeli data and Dong et al (2009) on Chinese data also shows negative association between wages, productivity and the proportion female, however, they do not document a significant wage – productivity gap. Both studies find that relative wages and relative productivities of female employees are 75 – 80 percent that of the men. Dostie (2006) uses Canadian data covering 1999-2002 and estimates a female – male wage gap of 0.85, and a female relative productivity of 0.8 – 0.9 depending on specification. Haltiwanger, Lane and Spletzer (1999) also found a negative relationship between labor productivity and the fraction of female workers, however, they do not aim to compare directly relative productivities and relative wages.

productivity of women compared to men is 1.54, while the relative wage is 1.08 based on the firm level wage bill, and only 0.89 based on the sum of individual wages.[30] This means that an increase of the ratio of women tends to increase the productivity of firms, and is contrary to the findings of international studies. Testing the equality of relative productivity and wages, the p-values of 0 suggests that women are significantly underpaid compared to what their productivity would suggest.

The estimates based on the first differenced data (shown in Table 3) which take firm fixed effects and thus selection at the firm level into account, suggest that this difference is due to the selection of women into more productive, and slightly better paying firms. The preferred specification, in which the Levinsohn and Petrin method is carried out on the differenced data (FDLP), estimates the relative productivity of females at 1.03 with a standard error of 0.03, suggesting that there is no significant difference in productivity of men and women within firms (Table 3.a). The relative wage estimate is 1.03 based on the firms' total wage bill, which means that women are paid in line with their productivity. However, the estimated relative wage is very different based on the sum of individual wages (Table 3.b): the relative wage estimate of 0.7 suggests that there is still a significant wage gap between men and women that is not in line with the non-existent productivity gap.

One main concern regarding the pooled estimates is the fact that the Hungarian labor market (and economy) underwent significant changes between 1986 and 2005, and the assumption of the same structure in all years is too strict. Estimation of the production function and wage equations on separate time periods is preferable for this reason, and may shed some light on the true underlying changes in the relative situation of the worker groups. We divided our sample into periods based on the transition process and previous literature (Kollo and Kertesi 2002): 1986-1992 covers the "transformational recession", 1993-1999 the period in which the market economy began to evolve, and 2000-2005 the final period. Table 4.a shows the estimates of relative productivity for the three time periods in the restricted specifications. The between-firm estimates of relative productivity suggest that women were three times more productive than men in the initial

---

[30] Individual wage regressions give results that are very similar to the firm level regressions based on the sum of individual wages, as would be expected, since the firm level equation is a weighted sum of the individual level equations.

period, and their relative productivity fell to 1.12 in the last period. Tables 4.b. and 4.c. show that their relative wage was below their relative productivity using either wage measure, though the firm level wage bill suggests they were paid more than men. The relative wage fell more slowly than the relative productivity in the subsequent periods, approaching their relative productivity by the last time period.

To assess the extent of selection at the firm level, the second panels of Tables 4.a-4.c show the time period results of the restricted specification in the FDLP within-firm case. The relative productivity of women is 0.75 within firms in the first period, suggesting that in the socialist era, women tended to be grouped into more productive firms, but were actually less productive within firms. In the second and third periods, their relative productivity increases, and is not significantly different from one, suggesting that this positive selection at the firm level decreased following the transition, while the productivity of women within firms is the same as that of men. One reason for the increasing productivity of women after the transitional period may be the growing share of the service sector in the economy where women became dominantly employed. Checking the robustness of our results by industries may help in sorting out the underlying processes. The two wage measures again give very different results regarding the relative wage of women: the firm level wage bill suggests that women were initially overpaid, but later paid in line with their productivity, while the sum of individual wages suggests a significant gap between relative productivity and wages of about 0.3 that remains through the final period.

Schooling

In general, previous empirical results point to a positive association between wages, productivity and the ratio of workers with diploma within the firm. However, results regarding the wage – productivity premium (relative wage paid in excess of relative productivity) are mixed. We would expect a positive wage – productivity premium as predicted by efficiency wage theories, or relative wages in line with relative productivities if competitive forces dominate.[31]

The pooled results in Table 2 suggest that between firms, skilled workers are much more productive than unskilled workers, with a relative productivity estimate of 3.3. The relative wage – using either the wage bill of sum of individual wages – is around 1.45, showing a significant gap between the relative productivity and wage of skilled workers. Turning our attention to the FDLP specification (Table 3.a), we see that the relative productivity of skilled workers falls to 1.05 within firms, while the relative wage falls to 1.01, which is in line with positive selection of skilled workers into more productive and better-paying firms.

The time period results in Table 4 paint a more accurate picture of the underlying processes. The NLS results show that the relative productivity of skilled workers increased from 2.3 to 3.9 by the last period, and the relative wages increased slightly from 1.4 to 1.6. The within firm results suggest that within firms, the relative productivity actually decreased from 1.36 to .99 (not significantly different from one), while relative wages increased from 0.94 to about one. These results mean that skilled workers are increasingly selected into better firms, and seem to be paid in line with their productivity within firms. This selection effect appears to be much stronger than seen in other international results, and contradicts the existence of a wage-productivity premium for skilled workers as seen in Dostie (2006).

Age

Empirical results are rather diverse regarding the relationship between wages, productivity and the age composition of the firm. However, the majority of studies (e.g.

[31] For example, Hellerstein and Neumark (2004) estimate a 56 percent productivity premium for diploma, which exceeds the 36 percent wage premium of college graduates. This result is somewhat opposite to our expectations, as not predicted by any standard theories. Dostie (2006) estimates a positive wage – productivity premium in an OLS specification with relative productivity of 1.18 and a graduate – no graduate relative wage of 1.27, while in a Levinsohn-Petrin framework for production function and estimating wage equation with individual unobserved heterogeneity, the converse is true, graduate relative productivity (1.22) is higher than relative wages (1.19). Haltiwanger, Lane and Spletzer (1999) also estimate a positive relationship between firm-level productivity and the proportion of workers with college education.

Hellerstein, Neumark, 2004; Dostie, 2006) find that prime-age workers increase productivity the most, and that higher proportion of old employees is associated with lower productivity. Wages are usually found to be rising and concave with age, and the comparison of relative productivities and wages usually imply that the older employee group receives a wage premium.[32] Studies examining the relationship between labor productivity and the age composition of the firm usually find that older workers decrease productivity and the ratio of young and prime-age workers is positively associated with firm-level productivity (e.g. Haltiwanger, Lane and Spletzer, 1999; Thierry and Rycx, 2009).[33]

Our pooled results (Table 2) show that between firms, the ratio of middle aged and older workers decreases the productivity of firms, the relative productivity estimates are 0.6 for middle-aged workers, and 0.4 for older workers. Relative wages do not reflect their productivities: middle aged and older workers are paid more than young workers. Relative wages increase with age, which is line with previous international results. The within-firm estimates (Tables 3.a and 3.b) show that much of the productivity gap is due to selection of young workers into better firms: the relative productivity of middle aged workers is just below one at 0.96, while older workers are equally productive as the young at 1.04 (not significantly different from one). Relative wages based on the firm-level wage bill are near one for every worker group, while the relative wage of older workers exceeds their relative productivity based on the sum of individual wages, especially for the middle-aged group.

The time period results in Table 4 seem to be in line with the theory of skill obsolescence described in previous transitional studies. The NLS results suggest that old

[32] For example, Hellerstein and Neumark (2004) documents a positive wage – productivity gap for prime-age workers with relative productivity of 1.12 and relative wage of 1.21 compared to young employees. Old employees are found to be less productive with relative productivity of 0.79, while their relative wages are 1.12 compared to young employees. Dostie (2006) also documents a positive wage premium for old workers with relative productivity of 0.95 and relative wage of 1.09, but she finds that prime-age workers receive 5 percentage point less than their productivity of 1.21. Vandenberghe and Waltenberg (2009) using Belgian data also estimates in a within-firm specification a positive wage premium for old workers with relative productivity of 50 percent and relative wages of 74 percent. However, young employees are found to be equally productive as prime-age workers and receive somewhat less than would be expected according to their productivity

[33] The empirical result is somewhat different in Sweden: Malmberg et al (2005) finds that the lower productivity of older workers reflect only the plant-specific lower productivity, as older workers tend to be employed in firms with old and less efficient technologies. After controlling for firm fixed-effects, they find a positive relationship between labor productivity and the ratio of old employees.

workers were more productive than the young in the initial period (relative productivity of 1.09), and were paid higher (relative wage of 1.17). Their productivity fell sharply in the later periods to 0.43, while their relative wages continued to rise to 1.56. The within-firm results reflect the same pattern: the relative productivity of middle-aged and older workers was 1.26 and 2.65 in the initial period, then fell to 0.95 and 1.01 in the latter periods. Older workers appear to have been underpaid within firms in the initial period, but their relative wages approach relative productivities by the last period, around 1 for both groups. Middle-aged workers seem to be overpaid and less productive than the young and the old in each specification, which is not in line with previous results that suggest that prime age workers contribute the most to firm productivity. This may reflect the deterioration of the value of skills gained before the transition, and reflects the selection of older workers to worse firms following the transition.

Unrestricted Results

Our goal in estimating the unrestricted specifications is two-fold. First, we seek to add new empirical evidence to the burgeoning international literature based on the Hellerstein and Neumark methodology. Due to the characteristics of our data - matched employers and employees, long time series panel in terms of firms, higher sampling ratio of workers - we are able to lift some of the restrictions of previous studies in estimating relative productivities and wages by allowing worker groups to differ along more dimensions, and by taking firm fixed effects and endogeneity biases into account. Our second goal is to re-examine post-transitional changes in the value of various worker characteristics described in previous empirical studies on Hungary using the new, longer time-series data and the econometric methodology used in the international studies. The specification of the unrestricted labor term may reveal trends that are hidden by the restricted results. For example, it is possible that the relative productivities and wages of women differ by skill level, or the return to education may vary by age (cohort), as in the presence of skill obsolescence. On the other hand, the estimation becomes imprecise if our measure of the ratio of workers in detailed groups is too noisy, so the specification of the unrestricted equations requires great care.

Our first set of unrestricted results allows workers to differ in productivity and wages along the three dimensions described in the methodology section. This means we lift the previously imposed restrictions on the quality of labor term, and allow for differences between worker groups defined as combinations of gender, age, and education level. Tables 6, 7.a, and 7.b present the estimated relative productivities and wages of the detailed worker groups, with the male, young, elementary education group as the reference category. It is clear that we are not able to obtain significant estimates based on this specification, and even less when we control for firm fixed effects. This means that there is not enough variation in the detailed worker groups to identify an effect. The worker group measure is biased by measurement error as we are calculating the ratios of workers in each category based on a sample of the workers. We thus turn our attention to an alternative specification, which proves more fruitful as it corrects for some of these problems.

Tables 8.a and 8.b provide the between and within-firm estimates of this unrestricted specification. We define a total of 8 worker groups as combinations of gender, "age" (those with pre-transitional work experience and those without), and akill level (those with a university degree and those without). The sample is divided into two time periods: the early period from 1992-1999, and the later from 2000-2005. This specification lends itself to the analysis of the relative value of pre- and post-transitional skills, or rather to differentiate two cohorts of workers based on their labor market participation.

Previous studies on Hungary have mostly found evidence of the devaluation of pre-transitional skills. Kertesi and Köllő (2002) found that the returns to old skills fell, while the return to new skills rose following the recovery from the transition, and that this was due to an increased productivity gap between the old-skilled and the young-skilled. Galasi (2004) found that wage premiums for the young and better educated workers increased increased, esp. after 1996, while the work experience of older workers was devalued in the market. Kézdi (2004) found that returns to education increased substantially between 1989 and 2002, and this increase was steeper for the younger generation. Inter-generational differences decreased, especially for young cohorts, reflecting a decrease in the value of experience from pre-transitional period. On the other hand, Campos and Jolliffe (2008) found that the return to education decreased in the late 90s for the young who received schooling post-transition, while the returns for older workers continued to rise. They suggest that the planned economy had under-valued education, and the market corrected for this over time. In terms of gender, Dong et al (2009) study the gender productivity-wage differential in a transitional context, and find that the state sector paid substantial wage subsidies to the most fragile unskilled female group who would be hardest hit by the market reforms, while skilled male and female workers are paid according to their marginal productivities.

The between-firm results in Table 8.a show that unskilled males withpre-transitional work experience are less productive and overpaid compared to the young. Over time, their productivity decreased, while their wages did not, suggesting that the experience gained prior to the transition did not increase workers' productivity. Table 8.b shows that this productivity gap is mostly due to selection to worse firms, within firms, the productivity of older workers is close to one. Similarly, skilled young workers are more productive and underpaid in the between-firm estimates, but within-firm, older workers are more productive than the young and underpaid. In the case of women, we see increasing selection to worse firms: the between-firm productivity of unskilled workers falls, while within firm, their productivity increases, and is higher than men's. Their wages do not reflect this change, and remain close to one. We also see the negative selection of old workers to worse firms in both the unskilled and skilled female workers. Overall, the results suggest a strong negative selection of older workers to worse firms, especially in the case of men. This means that much of the observed fall in the value of pre-transitional labor experience may be due to selection at the firm level: better firms demand younger workers more than older workers.

V. Conclusion

Our goal in this paper is to estimate the relative productivities and relative wages of different worker groups based on firm-level data from Hungary for the years 1986-2005. We then compare these estimates to see whether certain groups are over- or underpaid relative to their productivity, as would be the case if labor markets are not perfectly competitive, or in the case of wage discrimination. We assess the role of firm-level selection in determining productivity and wage gaps by including firm fixed effects in some of our specifications, and attempt to correct for the endogeneity of our production function estimates by applying the Levinsohn and Petrin method. To assess changes in the productive and wage returns to characteristics over time, we divide our sample into three time periods: that of the early transitional recession in 1986-1992, that of the recovery in 1993-1999, and the last period from 2000-2005.

Our restricted results – in which we estimate the relative productivity of women, middle aged and older workers, and workers with higher education separately- suggest that women selected into more productive firms initially, but this positive selection decreased following the transition, and within firms, their productivity increased over time to slightly above men's. When estimating the relative wage using the firm-level wage bill, there does not appear to be a significant gap by the last period, but when the estimate is based on individual wages, a significant gap of .3 persists through all the periods.

In terms of worker age, older workers were more productive than the young pre-transition, but their productivity fell following the transition, although some of the drop is due to negative selection to firms, as within firms they are equally productive as the young. This is in line with a loss of the value of skills gained before the transition. The middle-aged group appears to be overpaid in the last period, while older workers are compensated according to their relative productivity. In terms of education, the productivity gap between those with a higher education and those without grew following the transition, although it decreased within firms. Skilled workers were undervalued pre-transition, but their relative wage approaches their relative productivity in the later periods as the market corrected for the value of their skills.

The unrestricted specifications seek to shed some light on the processes behind these changes. When we define our worker groups in a way that allows us to assess the differences between skills gained pre- or post-transition, we find that much of the fall in the value of pre-transitional work experience is due to a negative selection at the firm-level: younger workers are selected into more productive firms. However, it is clear that the unrestricted estimation strategy requires further development in order for us to better explore these processes. The specification of worker groups and time periods needs to be set in a way that allows for meaningful estimates that are significant, and the difference between relative wage results using the firm-level wage bill and the weighted sum of individual wages needs to be addressed. We need to estimate separately for different industries to allow for structural differences, and also lift the restricting assumption of perfect substitution between worker types.

Table 1.: Summary Statistics of Firm-Level Sample 1986-2005

Year	**1986**	**1989**	**1992**	**1995**	**1998**	**2001**	**2005**
Number of firms	730	2,253	3,044	4,911	4,922	5,845	4,357
Productive inputs							
ln(Output)	7.49	7.28	6.63	6.42	6.47	6.44	6.92
ln(Capital)	6.11	5.80	5.31	5.08	4.94	4.82	5.49
ln(Materials)	6.33	6.16	5.30	5.11	5.17	5.38	5.89
Employment	366	310	293	227	208	167	205
Wagecost	6.10	6.11	5.50	5.05	4.94	4.78	5.30
Demographic composition							
female	0.41	0.42	0.41	0.38	0.38	0.38	0.41
high school	0.48	0.52	0.62	0.65	0.67	0.69	0.68
university	0.11	0.10	0.11	0.11	0.12	0.14	0.15
middle-aged	0.57	0.61	0.62	0.58	0.56	0.51	0.47
old	0.16	0.17	0.13	0.15	0.16	0.20	0.27

Source: Hungarian Wage and Employment Survey 1986-2005. Sample restricted to private sector firms with at least 20 employees, and 5% of their workforce included in the employee dataset. Values represent means.

Table 2: Pooled Restricted Between-Firm Results

Between-firm regressions, 1986 - 2005				
	NLS	**LP**		
	log sales	**log sales**	**log wagecost**	**log sumwage**
log capital	0.095		0.054	0.058
	0.002		0.001	0.002
log materials	0.586		0.153	0.121
	0.003		0.002	0.002
log labor	0.318	0.309	0.829	0.861
	0.004	0.004	0.002	0.003
female / male	1.535	1.452	1.081	0.886
	0.051	0.050	0.008	0.009
skilled / unskilled	3.288	3.394	1.457	1.497
	0.174	0.187	0.014	0.019
middle-aged / young	0.604	0.599	1.205	1.274
	0.022	0.022	0.012	0.017
old / young	0.436	0.426	1.455	1.549
	0.024	0.024	0.017	0.023
No. obs	67,928			

Table 3.a: Pooled Restricted Within-Firm Results, Firm-Level Wagecost

Within-firm regressions, 1986 - 2005						
	FD		FD+LP			
	log sales	log wagecost	log sales	log wagecost	gap (wage - prod)	p-value (wage = prod)
log capital	0.075	0.027		0.027		
	(0.005)	(0.003)		(0.003)		
log materials	0.425	0.115		0.115		
	(0.004)	(0.002)		(0.002)		
log labor	0.366	0.677	0.370	0.677		
	(0.006)	(0.004)	(0.006)	(0.004)		
female / male	1.029	1.026	1.025	1.026	0.001	0.981
	(0.029)	(0.009)	(0.028)	(0.009)		
skilled / unskilled	1.050	1.012	1.052	1.012	-0.040	0.108
	(0.026)	(0.008)	(0.026)	(0.008)		
middle-aged / young	0.956	0.998	0.957	0.998	0.041	0.059
	(0.022)	(0.008)	(0.022)	(0.008)		
old / young	1.039	1.003	1.036	1.003	-0.033	0.271
	(0.032)	(0.010)	(0.031)	(0.010)		
No. obs	40,868					

Table 3.b: Pooled Restricted Within-Firm Results, Sum of Individual Wages

Within-firm regressions, 1986 - 2005						
	FD		FD+LP			
	log sales	log sumwage	log sales	log sumwage	gap (wage - prod)	p-value (wage = prod)
log capital	0.075	0.034		0.034		
	(0.005)	(0.006)		(0.006)		
log materials	0.425	0.045		0.045		
	(0.004)	(0.005)		(0.005)		
log labor	0.366	0.704	0.369	0.704		
	(0.006)	(0.008)	(0.006)	(0.008)		
female / male	1.026	0.703	1.023	0.704	-0.319	0.000
	(0.028)	(0.015)	(0.028)	(0.015)		
skilled / unskilled	1.050	1.032	1.052	1.032	-0.020	0.523
	(0.026)	(0.018)	(0.026)	(0.018)		
middle-aged / young	0.957	1.122	0.958	1.122	0.164	0.000
	(0.022)	(0.019)	(0.022)	(0.019)		
old / young	1.039	1.069	1.037	1.069	0.032	0.408
	(0.032)	(0.024)	(0.031)	(0.024)		
No. obs			40,868			

Table 4.a: Time Period Restricted Results, Productivity

	Between-firm regressions, NLS			Within-firm regressions, FD+LP		
	log sales			log sales		
	1986-1992	1993-1999	2000-2005	1986-1992	1993-1999	2000-2005
log capital	0.056	0.078	0.107			
	(0.008)	(0.003)	(0.003)			
log materials	0.598	0.574	0.592			
	(0.010)	(0.004)	(0.004)			
log labor	0.338	0.337	0.315	0.213	0.342	0.409
	(0.014)	(0.006)	(0.006)	(0.030)	(0.009)	(0.009)
female / male	2.959	1.715	1.123	0.753	1.016	1.084
	(0.351)	(0.075)	(0.059)	(0.214)	(0.042)	(0.039)
skilled / unskilled	2.259	2.921	3.866	1.359	1.114	0.988
	(0.328)	(0.189)	(0.359)	(0.376)	(0.042)	(0.031)
middle-aged / young	0.791	0.584	0.642	1.261	0.975	0.946
	(0.110)	(0.027)	(0.039)	(0.538)	(0.033)	(0.029)
old / young	1.093	0.402	0.425	2.648	1.009	1.013
	(0.188)	(0.032)	(0.035)	(1.118)	(0.046)	(0.039)
No. obs	5,919	32,886	29,123	1,193	20,438	19,237

Table 4.b: Time Period Restricted Results, Firm-Level Wagecost

	Between-firm regressions, NLS			Within-firm regressions, FD+LP		
	log wagecost			log wagecost		
	1986-1992	1993-1999	2000-2005	1986-1992	1993-1999	2000-2005
log capital	0.045	0.063	0.049	0.005	0.030	0.030
	(0.004)	(0.002)	(0.002)	(0.018)	(0.004)	(0.004)
log materials	0.175	0.147	0.161	0.217	0.130	0.088
	(0.006)	(0.003)	(0.003)	(0.016)	(0.003)	(0.003)
log labor	0.685	0.822	0.852	0.486	0.657	0.747
	(0.008)	(0.003)	(0.004)	(0.021)	(0.005)	(0.005)
female / male	1.218	1.078	1.059	1.293	0.989	1.033
	(0.037)	(0.012)	(0.013)	(0.106)	(0.013)	(0.012)
skilled / unskilled	1.413	1.403	1.539	0.942	1.040	0.999
	(0.047)	(0.018)	(0.025)	(0.076)	(0.012)	(0.010)
middle-aged / young	1.082	1.169	1.259	1.183	0.980	1.008
	(0.043)	(0.016)	(0.022)	(0.139)	(0.010)	(0.010)
old / young	1.172	1.379	1.566	1.333	0.988	0.997
	(0.061)	(0.023)	(0.027)	(0.180)	(0.014)	(0.012)
No. obs	5,919	32,886	29,123	1,193	20,438	19,237

Table 4.c: Time Period Restricted Results, Sum of Individual Wages

	Between-firm regressions, NLS			Within-firm regressions, FD+LP		
	log sumwage			log sumwage		
	1986-1992	1993-1999	2000-2005	1986-1992	1993-1999	2000-2005
log capital	0.023	0.066	0.058	0.077	0.031	0.020
	(0.007)	(0.003)	(0.003)	(0.030)	(0.009)	(0.009)
log materials	0.131	0.111	0.132	0.117	0.048	0.033
	(0.009)	(0.004)	(0.004)	(0.028)	(0.007)	(0.007)
log labor	0.777	0.846	0.887	0.653	0.598	0.813
	(0.012)	(0.005)	(0.005)	(0.036)	(0.013)	(0.012)
female / male	0.706	0.911	0.869	0.573	0.637	0.779
	(0.030)	(0.013)	(0.013)	(0.070)	(0.023)	(0.020)
skilled / unskilled	1.395	1.460	1.563	0.947	1.105	0.987
	(0.065)	(0.026)	(0.031)	(0.098)	(0.034)	(0.021)
middle-aged / young	1.146	1.282	1.269	1.040	1.157	1.097
	(0.066)	(0.024)	(0.027)	(0.143)	(0.033)	(0.023)
old / young	1.388	1.585	1.534	0.970	1.157	1.008
	(0.098)	(0.036)	(0.033)	(0.171)	(0.044)	(0.028)
No. obs	5,919	32,886	29,123	1,193	20,438	19,237

Table 5.a: Time Period Restricted Results, Wage Gaps, Firm-Level Wagecost

Gap (wagecost - productivity), FD+LP specification						
	1986-1992		1993-1999		2000-2005	
	gap	p-value (wage=prod)	gap	p-value (wage=prod)	gap	p-value (wage=prod)
female / male	0.540	0.012	-0.027	0.513	-0.051	0.185
skilled / unskilled	-0.416	0.252	-0.074	0.068	0.011	0.708
middle-aged / young	-0.078	0.880	0.005	0.881	0.062	0.028
old / young	-1.315	0.227	-0.020	0.649	-0.015	0.690

Table 5.b: Time Period Restricted Results, Wage Gaps, Sum of Individual Wages

Gap (sumwage - productivity), FD+LP specification						
	1986-1992		1993-1999		2000-2005	
	gap	p-value (wage=prod)	gap	p-value (wage=prod)	gap	p-value (wage=prod)
female / male	-0.198	0.364	-0.375	0.000	-0.302	0.000
skilled / unskilled	-0.395	0.279	-0.008	0.880	-0.001	0.989
middle-aged / young	-0.203	0.696	0.181	0.000	0.150	0.000
old / young	-1.619	0.127	0.148	0.019	-0.005	0.922

Table 6: Pooled Unsrestricted Between-Firm Estimates

Between-firm regressions (NLS), 1986 - 2005			
	log sales	**log wagecost**	**log sumwage**
male, elem, middle	0.606	0.948	0.953
	(0.100)	(0.028)	(0.036)
male, elem, old	0.507	1.161	1.032
	(0.116)	(0.039)	(0.046)
male, high, young	2.733	1.106	1.015
	(0.281)	(0.026)	(0.031)
male, high, middle	1.669	1.346	1.390
	(0.169)	(0.029)	(0.037)
male, high, old	1.319	1.600	1.751
	(0.154)	(0.037)	(0.051)
female, elem, young	1.242	0.797	0.721
	(0.239)	(0.040)	(0.046)
female, elem, middle	0.767	0.953	0.776
	(0.114)	(0.028)	(0.030)
female, elem, old	0.349	1.077	0.787
	(0.125)	(0.043)	(0.046)
female, high, young	5.217	1.172	0.960
	(0.543)	(0.030)	(0.032)
female, high, middle	2.631	1.479	1.244
	(0.269)	(0.033)	(0.035)
female, high, old	2.143	1.884	1.546
	(0.269)	(0.050)	(0.054)
No. obs		67,928	

Table 7.a: Pooled Unsrestricted Within-Firm Estimates, Firm-Level Wagecost

Within-firm regressions (FD+LP), 1986 - 2005				
	log sales	**log wagecost**	**gap (wage-prod)**	**p-value (wage=prod)**
male, elem, middle	0.879	0.976	0.097	0.030
	(0.046)	(0.016)		
male, elem, old	0.965	0.991	0.026	0.681
	(0.065)	(0.022)		
male, high, young	0.992	0.999	0.006	0.876
	(0.042)	(0.014)		
male, high, middle	0.975	0.998	0.024	0.493
	(0.036)	(0.012)		
male, high, old	1.054	1.008	-0.046	0.301
	(0.046)	(0.015)		
female, elem, young	0.973	1.011	0.038	0.645
	(0.084)	(0.029)		
female, elem, middle	0.964	1.013	0.049	0.338
	(0.052)	(0.018)		
female, elem, old	0.986	1.018	0.033	0.658
	(0.076)	(0.026)		
female, high, young	1.022	1.017	-0.005	0.915
	(0.052)	(0.017)		
female, high, middle	0.957	1.014	0.056	0.144
	(0.040)	(0.014)		
female, high, old	1.056	1.001	-0.055	0.338
	(0.059)	(0.019)		
No. obs		40,868		

Table 7.b: Pooled Unsrestricted Within-Firm Estimates, Sum of Individual Wages

Within-firm regressions (FD+LP), 1986 - 2005				
	log sales	**log sumwage**	**gap (wage-prod)**	**p-value (wage=prod)**
male, elem, middle	0.879	1.144	0.265	0.000
	(0.046)	(0.041)		
male, elem, old	0.966	0.960	-0.006	0.942
	(0.065)	(0.049)		
male, high, young	0.991	1.057	0.066	0.209
	(0.042)	(0.033)		
male, high, middle	0.976	1.230	0.254	0.000
	(0.036)	(0.032)		
male, high, old	1.054	1.220	0.165	0.005
	(0.046)	(0.038)		
female, elem, young	0.972	0.858	-0.114	0.254
	(0.084)	(0.057)		
female, elem, middle	0.964	0.905	-0.058	0.348
	(0.052)	(0.037)		
female, elem, old	0.988	0.868	-0.120	0.188
	(0.076)	(0.053)		
female, high, young	1.020	0.755	-0.265	0.000
	(0.052)	(0.031)		
female, high, middle	0.956	0.860	-0.096	0.041
	(0.040)	(0.027)		
female, high, old	1.055	0.848	-0.207	0.003
	(0.059)	(0.039)		
No. obs		40,868		

Table 8.a: Alternative Unrestricted Two Period Between-Firm Results

Between-firm regressions, NLS				
	log sales		log wagecost	
	1992-1999	**2000-2005**	**1992-1999**	**2000-2005**
Male, old, unskilled	0.531	0.401	1.492	1.481
	(0.044)	(0.033)	(0.063)	(0.050)
Male, young, skilled	8.195	4.210	3.792	4.009
	(0.902)	(0.387)	(0.235)	(0.165)
Male, old, skilled	2.459	2.050	2.684	2.868
	(0.210)	(0.171)	(0.118)	(0.104)
Female, young, unskilled	1.812	0.905	1.133	1.005
	(0.191)	(0.095)	(0.071)	(0.050)
Female, old, unskilled	0.893	0.529	1.555	1.496
	(0.073)	(0.040)	(0.066)	(0.049)
Female, young, skilled	10.235	2.816	4.237	3.186
	(1.245)	(0.339)	(0.286)	(0.155)
Female, old, skilled	2.555	1.345	3.087	3.658
	(0.262)	(0.187)	(0.151)	(0.151)
No. obs	35,822	29,123	35,822	29,123

Table 8.b: Alternative Unrestricted Two Period Within-Firm Results

	Within-firm regressions, FD			
	log sales		log wagecost	
	1992-1999	2000-2005	1992-1999	2000-2005
Male, old, unskilled	0.933	0.978	0.993	0.984
	(0.064)	(0.050)	(0.024)	(0.015)
Male, young, skilled	0.853	1.098	0.975	0.994
	(0.172)	(0.105)	(0.062)	(0.029)
Male, old, skilled	1.107	1.183	0.973	1.023
	(0.099)	(0.089)	(0.031)	(0.024)
Female, young, unskilled	0.981	1.098	1.007	1.031
	(0.106)	(0.077)	(0.037)	(0.022)
Female, old, unskilled	0.943	1.077	0.999	1.021
	(0.070)	(0.061)	(0.026)	(0.017)
Female, young, skilled	1.046	1.223	1.101	1.003
	(0.219)	(0.124)	(0.076)	(0.033)
Female, old, skilled	1.117	1.169	0.969	1.059
	(0.127)	(0.118)	(0.040)	(0.033)
No. obs	21,045	19,237	21,045	19,237

References:

- Abowd, John M., Paul Lengermann, and Kevin L. McKinney (2003), The Measurement of Human Capital in the US Economy, Unpublished manuscript, Cornell University.
- Abowd, John, M., Francis Kramarz, and David N. Margolis (1999), High Wage Workers and High Wage Firms, Econometrica, Vol. 67, No. 2, pp. 251-333.
- Ackerberg, D. A., Kevin Caves, and Garth Frazer (2005), Structural Identification of Production Functions
- Bowlus, A.J. (1997), A Search Interpretation of Male-Female Wage Differentials, Journal of Labor Economics, Vol. 15, Part 4, pp. 625-657.
- Brown, David J., John S, Earle, and Álmos Telegdy (2006), The Productivity Effects of Privatization: Longitudinal Estimates from Hungary, Romania, Russia, and Ukraine, Journal of Political Economy, Vol. 114, No. 1.
- Campos, Nauro and Jolliffe, Dean (2007), Earnings, Schooling and Economic Reform: Econometric Evidence from Hungary (1986-2004), IZA Discussion Paper No. 2678.
- Dong, Xiao-yuan and Liqin Zhang (2009), Economic Transition and Gender Differentials in Wages and Productivity: Evidence from Chinese Manufacturing Enterprises, Journal of Development Economics, 88, 144-156.
- Dostie, Benoit (2006), Wages, Productivity and Aging, IZA DP. 2496.
- Foster, L., Haltiwanger, J., and Krizan C.J. (2001), Aggregate productivity growth: lessons from microeconomic evidence. In; Dean, Hulten, Harper (Eds.), New Developments in Productivity Analysis. NBER/University of Chicago Press.
- Fox, Jeremy T. And Valerie Smeets (2007), Do Input Quality and Structural Productivity Estimates Drive Measured Differences in Firm Productivity?, WP.
- Frazer, Garth (2007), Firm Productivity, Worker Ability, and Returns to Education, University of Toronto, Working Paper.

- Galasi, Péter (2004), Have Higher-Education Diplomas Been Devalued? Skill-Biased Technological Change and the Reallocation of College, Budapest Working Papers on the Labor Market BWP 2004/3.
- Griliches, Z., and W. Mason (1972), Education, Income and Ability, Journal of Political Economy, Vol. 80, Part 2, pp. 74-103.
- Griliches, Zvi (1977), Estimating the Returns to Schooling: Some Econometric Problems, Econometrica, Vol. 45, Number 1.
- Griliches, Zvi (1960), Measuring Inputs in Agriculture: A Critical Survey, *Journal of Farm Economics*, Vol. 42, No. 5, pp. 1411-33.
- Griliches, Zvi (1957), Bias in Estimates of Production Functions, Journal of Farm Economics, Vol. 39, No. 1, pp. 8-20.
- Griliches, Zvi (1970), Notes on the Role of Education in Production Functions and Growth Accounting, In: L. Hansen, ed., Education, Income, and Human Capital. NBER Studies in Income and Wealth, Vol. 35 (New York: Columbia University Press), pp. 71-115.
- Haltiwanger, John C., Julia I. Lane, and James R. Spletzer (1999), Productivity Differences Across Employers: The Roles of Employer Size, Age, and Human Capital, American Economic Review Papers and Proceedings, Vol 89, No 2, pp94-98.
- Haltiwanger, John C., Julia I. Lane, and James R. Spletzer (2007), Wages, Productivity, and the Dynamic Interaction of Businesses and Workers, Labour Economics, 14, pp 575-602.
- Haskel, Jonathan, Denise Hawkes, and Sonia Pereira (2005), Skills, Human Capital and the Plant Productivity Gap: UK Evidence from Matched Plant, Worker and Workforce Data, CEPR Discussion Paper no. 5334.
- Hellerstein, J. K, and D. Neumark (1999), Sex, Wages and Productivity: An Empirical Analysis of Israeli Firm-Level Data, International Economic Review, Vol. 40, No. 1., pp. 95-123.
- Hellerstein, J. K, D. Neumark and K. R. Troske (1999), Wages, Productivity, and Worker Characteristics: Evidence from Plant-Level Production Functions and Wage Equation. Journal of Labor Economics 17(3), July, 409-46.

- Hellerstein, Neumark (2004), Production Function and Wage Equation Estimation with Heterogeneous Labor: Evidence from a New Matched Employer-Employee Dataset
- Ichino, A. and Moretti, E. (2006), Biological Gender Differences, Absenteeism, and the Earning Gap, NBER Working Paper No. 12369.
- Ilmakunnas, Pekka, and Mika Maliranta (2003), Technology, Labor Characteristics and Wage-Productivity Gaps, ETLA DP. 860.
- Iranzo, Susana, Fabiano Schivardi and Elisa Tosetti (2006), Skill Dispersion and Firm Productivity: An Analysis with Employer-Employee Matched Data, Working Paper CRENoS 200617.
- Kawaguchi, D. (2007), A Market Test of Sex Discrimination: Evidence from Japanese Firm-Level Data", forthcoming in International Journal of Industrial Organization.
- Kertesi, Gábor and Köllő, János (2002), Economics Transformation and the Revaluation of Human Capitak – Hungary 1986-1999, Research in Labor Economics, Vol. 21. Pp.235-273.
- Kézdi, Gábor (2005), "Education and Earnings." in: K. Fazekas and J. Varga (eds.), The Hungarian Labour Market 2005. In Focus: Education and the Labour Market, Institute of Economics HAS.
- Lallemand, Thierry, and Francois Rycx (2009), Are Young and Old Workers Harmful for Firm Productivity?, IZA DP. No. 3938.
- Levinsohn, J. and Amil Petrin (2003), Estimating Production Functions Using Inputs to Control for Unobservables, Review of Economic Studies 70, 317-341.
- Malmberg, Bo, Lindh, Thomas and Max Halvarsson (2005), Productivity consequences of workforce ageing – Stagnation or a Horndal effect?, Institute for Future Studies, ISBN 91-89655-75-3
- Melitz, Marc J. (2000), Estimating Firm-Level Productivity in Differentiated Product Industries, Draft Paper, Harvard University.
- Neal, Derek A., William R. Johnson (1996), The Role of Premarket Factors in Black-White Wage Differences, The Journal of Political Economy, Vol. 104, No. 5, pp. 869-895.

- Olley, G. and Pakes, A. (1996), The dynamics of productivity in the telecommunications equipment industry, Econometrica, Vol. 64, pp. 1263-1297.
- Van Biesebroeck, Johannes (2007), Wages Equal Productivity. Fact or Fiction?, University of Toronto, Dept of Economics, WP 294.
- Vandenberghe, V. and F. Waltenberg (2009), Ageing Workforce, Productivity and Labour costs of Belgian Firms, Conference paper, presented at the Belgian Day for Labour Economists, 16 June