The Evolution of Productivity-Wage Gaps Following the Transition in Hungary – Evidence from Linked Employer-Employee Data

Anna Lovász
Institute of Economics, Hungarian Academy of Sciences
lovasz@econ.core.hu

Mariann Rigó
Central European University
rigom@ceu.hu

June 30, 2010

Abstract

We explore the relative productivities and wages of worker groups over a 20 year period following the transition in Hungary. Due to the economic transition, firms may have become more efficient in terms of setting wages, relative productivities and wages would converge over time. The linked employer-employee dataset allows us to control for selection bias at the occupation, firm, region, and industry level, and to assess long-term trends. The results do not suggest that firm wage setting became more efficient: we find a persistent gap between the relative wages and productivities of both the high-skilled and older workers. Firms who entered the market after the transition set wages more efficiently than older firms.

The authors would like to thank John Earle, Álmos Telegdy, Bruno Van der Linden, János Köllő, Gábor Antal, Shelly Lundberg, Dick Startz, Vincent Vandenberghe, 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.

1. Introduction

Economic literature highlights the importance of firm behavior and the decisions of employers in determining productivity and economic growth. These play an especially key role in a transitional setting, where the change from the centrally planned system to more efficient market-driven choices is key to successful long-term growth. The recent availability of longitudinal datasets that link employers to data on employee characteristics has enabled researchers to assess not only the contribution of employer's decisions regarding capital, material inputs, and the size of their workforce to firm productivity, but also the role of the skill endowment and the demographic composition of their workers. In this paper, we examine the changes over several time periods in the relative productivity and wage of various worker groups using a dataset from Hungary that covers the pre- and post-transitional period.[1] During communism, firms did not have the incentive to set wages equal to marginal productivity, in fact, wages tended to be compressed on an ideological basis. Theory suggests that as markets become more competitive, firm behavior should become more efficient, employers should choose a better mix of workers, and workers' wages should increasingly be in line with their productivity. Labor market distortions – such as occupational or wage discrimination or the compression of the wage structure – are expected to decrease over time.[2]

To assess this hypothesis, we estimate jointly firm-level wage equations and production functions augmented with the worker composition of the firm. The estimated coefficients provide us with information on how the different groups of workers contribute to productivity and labor costs. Following the methodology pioneered by Hellerstein and Neumark (1999), and accepting certain simplifying assumptions, estimated coefficients can be translated into group-level relative productivities and wages[3]. This procedure allows us to estimate the relative productivities and wages of three worker groups – differentiated by gender, education, and age –

[1] Brown and Earle (2008) study the contribution of inter-firm employment reallocation to aggregate productivity growth following the transition in six economies, including Hungary. They find that reallocation rates were very low during socialism, increased dramatically - contributing to rapid growth - following the transition, and can be seen as a correction of previous distortions in the economy. In this paper, our focus is on the effects of employment allocation at another level: in terms of the demographic composition of workers within firms.

[2] For example, Becker's (1957) model of employer taste discrimination implies that in the long run, product market competition may force discriminating employers out of the product market, leading to a fall in the wage gap.

[3] We will discuss these assumptions and their applicability in greater detail in the methodology section.

at the firm level.[4] We assess the changes over four time periods (between 1986 and 2005) in the relative productivities of each group, in their relative wages, and in the gap between the two. If competitive forces led to increased efficiency, as expected, the relative wage of each group should approach their productivity, and the gap should decrease. On the other hand, there are several reasons why the marginal product and wage of a group may not be equal: for example, older workers generally receive a wage premium that does not reflect an increase in their productivity even in highly competitive markets, and differentials may exist in the case of compensating wage differentials or efficiency wages. In this paper, our goal is not to look for evidence of behavior that is not profit-maximizing, such as wage discrimination, though our methodology provides an appealing alternative to the traditional wage equation methodology for doing so.[5] We assess whether there is an overall trend of convergence in the relative wages and productivities of the groups over time. The methodology used also allows us to determine whether changes in the gaps were due to an underlying change in the productive contributions of the groups, or to changes in the firms' wage setting practices.

The linked employer-employee dataset includes several variables that we use to control for a variety of firm characteristics and segregation effects at several levels. Besides having information on the gender, age, and education of the employees, our detailed worker variables include occupational categories, so we control for the selection of occupational composition of

[4] While we estimate group-level differences at the firm-level, some studies use direct measures of individual productivity: see, for example, Holzer (1990). Another strand of the literature focuses on finding a suitable measure of worker ability, such as Griliches and Mason (1972), Griliches, (1977), Neal and Johnson, (1996). Abowd, Kramarz, and Margolis (1999) develop a measure of human capital that incorporates individual observable and unobservable productivity components, by identifying unobservable worker and firm fixed effects (see also: Abowd, Lengerman and McKinney (2003), Iranzo, Schivardi, Tosetti, (2006)). Unfortunately, these methods require data that are still relatively rare.

[5] Campos and Jolliffe (2005) use traditional wage equation and decomposition to estimate the residual gender wage gap in Hungary for 1986-1998. They find that the unexplained portion of the wage gap decreased following the transition, which is in line with a decrease in discrimination. However, the residual gap may be due to unobserved differences between males and females, so this methodology cannot be used to determine whether the fall in the gap was due to more efficient behavior of firms, or changes in the productive capabilities of women compared to men. Studies that use the Hellerstein and Neumark (1999) methodology usually find significantly lower discrimination against women, since the relative productivity estimates may capture group level unobserved productive differences.

firms, since workers may be selecting into jobs with differing productivity and wage levels. We also estimate the relative wages and productivities separately by broad industrial categories, to see if there are significant differences in composition and wage setting between industries. We use information on employers to control for differences among firms that are due to ownership types (foreign, state, or domestic private), and regional differences. Perhaps most importantly, since the dataset follows firms over time, we are able to estimate within-firm effects by controlling for firm characteristics that are stable over time. This has not been possible in many studies using the Hellerstein and Neumark methodology, as these are usually carried out on a cross-section of the data[6]. These studies could not determine the extent of firm level selection, which may significantly impact the estimated productivities and wages of worker groups. Recently, there are papers using panel databases and following firms over time, however, these databases tend to be less detailed regarding employee information and shorter in time than the database available to us[7]. Due to the fact that our data covers almost twenty years, we are able to estimate the within-firm effects for several separate time periods, allowing us to assess changes over time in the relative productivities and wages of the groups while also controlling for firm level selection. Finally, we also separate the sample of firms into those that existed prior to the transition and those that entered the market later to see if they behave differently.

Our results indicate significant differences in the estimated relative productivities and wages of the groups that remain significant in the last period in the case of high-skilled and older workers, and no significant gap for women. Occupational and firm segregation play an important role in determining the productivity and wage gaps between groups, with high-skilled workers selecting into better (more productive and better paying) firms, women and older workers into worse firms. For women, firm level segregation increased, while within-firms, their relative productivity and wage increased to above 100% of men's. In the OLS specifications, women

[6] There are several studies identifying the wage-productivity gap from between-firm variation, e.g. Hellerstein, Neumark (1999, 2004), Hellerstein, Neumark and Troske (1999) or Van Biesebroeck (2007).

[7] For example, Ours and Stoeldraijer (2010) uses a database of Dutch firms covering 2000-2005, and has information only on the age and gender of the employees. Crepon et al (2002) analyzes a French database of 1994-1997 including information on the gender, age and occupation of the workers. Borowczyk and Vandenberghe (2010) and Vandenberghe and Waltenberg (2010) analyzes Belgian data covering the years of 1998-2006 and includes information on the gender, age and occupation of the employees.

appear to be overpaid, but once we account for firm-level selection, the gap between their relative wage and productivity decreased to 0. Firms that existed prior to the transition have significantly different wage setting practices than new entries. In old firms, the gap between wages and productivity increased over time, even within firms. College graduates remain underpaid relative to their productive contribution, especially in old firms, while workers older than 40 are slightly overpaid, though less so in old firms, where their relative productivity is higher than in new firms.

The remaining sections of the paper will be organized as follows: in Section 2, we discuss our empirical approach and the relevant previous literature on the relationship between worker composition, productivity, and wages, as well as the main estimation issues. In Section 3, we describe the data used in the estimation, including summary statistics and preliminary graphs. Section 4 presents the results of the estimation for each worker group examined: women, college graduates, and workers over the age of 40, comparing them to previous international results, and summarizes the evidence regarding our main hypothesis. Section 5 concludes.

2. Empirical Methodology

The methodology we use relies heavily on the pioneering work of Hellerstein, Neumark (1999) and Hellerstein, Neumark and Troske (1999). Our approach is to estimate production functions and wage equations at the firm level and examine the evolution of the productivity and the labor cost contributions of various worker groups over time. By computing group-level relative productivities and relative wages, we can also study the evolution of the wage – productivity gap. Our empirical hypothesis is based on the comparison of group-level relative productivities and wages: if increased competition after the regime change leads to more efficient behavior of firms, we would expect the gap between the relative productivities and wages of various worker groups to decrease.

We will first turn our attention to the empirical specification of the production function and wage equation separately, then we will discuss the estimation issues and data limitations, and their consequences.

2.1. Production function

We estimate a Cobb-Douglas production function with value added as our measure of output and the input controls are capital and a labor quality variable:

$$\ln VA_{jt} = \alpha_0 + \alpha \cdot \ln K_{jt} + \gamma \cdot QL_{jt} + \delta \cdot Z_{jt} + u_{jt} \qquad (1)$$

Value added is defined as sales minus material costs, and our capital measure is year-average tangible assets. The labor quality variable (*QL*) serves to account for the different productivity contributions of the various worker groups, and it is defined as the productivity-adjusted sum of employees. The matrix *Z* includes additional controls that may determine a firms' productivity. These controls are industry, year, region, ownership variables, and in some cases, firm fixed effects[8].

If workers are grouped into $n = 0, 1, \ldots, N$ categories, and L_n, φ_n show the number and the economy-wide productivities of employees in group *n*, respectively, the *QL* term takes the following form:

$$QL = \sum_{n=0}^{N} \varphi_n L_n = \varphi_0 L_0 + \sum_{n=1}^{N} \varphi_n L_n = \varphi_0 L\left[1 + \sum_{n=1}^{N}\left(\frac{\varphi_n}{\varphi_0} - 1\right)\frac{L_n}{L}\right] \qquad (2)$$

Hence, the labor quality augmented production function allows us to identify *N* relative productivity parameters, taking $n = 0$ as the reference group[9].

$$\ln VA_{jt} = \alpha_0 + \alpha \ln K_{jt} + \gamma \ln \varphi_0 + \gamma \ln L_{jt} + \gamma \ln\left[1 + \sum_{n=1}^{N}\left(\frac{\varphi_n}{\varphi_0} - 1\right)\frac{L_{n_{jt}}}{L_{jt}}\right] + \delta \cdot Z_{jt} + u_{jt} \qquad (3)$$

Ideally, workers need to be assigned to groups divided along every 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 based on gender, age (under or over 40 years), education (college or no college) and occupation (7 categories based on 1 digit code). The interactions of these categories gives us a total of 56 detailed worker groups, or 55 relative parameters to be estimated.

[8] We control for 19 industrial categories, 7 regions, and state, domestic, or foreign ownership. Year dummies are included in each specification.

[9] The productivity of the reference group cannot be estimated separately, and is incorporated into the constant term.

Since grouping workers into detailed categories requires estimating a large number of productivity parameters, in most studies two restrictions are applied to the labor quality term. 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 college and no college employees; or, the productivity ratio between workers with and without degree 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[10].

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 55 to 9, and of these, we focus on the coefficient estimates of the three groups – women, college graduates and those over 40 – and include occupational categories only as controls.

By using the simplification on the *QL* term, the production function reduces to:

$$
\begin{aligned}
\ln VA_{jt} &= \alpha_0 + \alpha \cdot \ln K_{jt} + \gamma \cdot \ln L_{jt} + \\
&+ \gamma \cdot \ln\left[1 + (\varphi_F - 1)\frac{L_{F_{jt}}}{L_{jt}}\right] + \gamma \cdot \ln\left[1 + (\varphi_O - 1)\frac{L_{O_{jt}}}{L_{jt}}\right] + \\
&+ \gamma \cdot \ln\left[1 + (\varphi_U - 1)\frac{L_{U_{jt}}}{L_{jt}}\right] + \delta \cdot Z_{jt} + u_{jt}
\end{aligned}
\tag{4}
$$

[10] 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.

In equation (4) the parameters $\varphi_{F,}$ φ_O and φ_U show the relative productivity of female to male, workers above 40 to workers below 40, and workers with university degree to workers without diploma[11].

2.2. *Wage 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 wages 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 earnings 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 allows for the simultaneous estimation of the production function and wage equation. The simultaneous model minimizes the impact of the unobserved shocks on productivity and wages, and allows the error terms to be correlated across equations, leading to more efficient estimates. Joint estimation at the firm level also has the benefit of making the hypothesis test of the equality of relative productivities and relative wages straightforward. We use the latter specification with the firm's annual wage bill as our dependent variable in the wage equations.

The firm-level wage equation can be considered a definitional equation, aggregating individual-level equations over all workers[12]. Grouping workers by gender, age, education, occupation, and applying the restrictions of equal relative wages and proportions across other worker groups, the wage equation is:

[11] Occupational shares are included in the Z matrix, as we will not analyze their coefficient estimates here.

[12] For more detailed discussion of the firm-level wage equation, see for example, Hellerstein, Neumark (1999), page 100.

$$\ln W_{jt} = \eta_0 + \eta \cdot \ln K_{jt} + \theta \cdot \ln L_{jt} + \\ + \theta \cdot \ln\left[1 + (\lambda_F - 1)\frac{L_{F_{jt}}}{L_{jt}}\right] + \theta \cdot \ln\left[1 + (\lambda_O - 1)\frac{L_{O_{jt}}}{L_{jt}}\right] + \\ + \theta \cdot \ln\left[1 + (\lambda_U - 1)\frac{L_{U_{jt}}}{L_{jt}}\right] + \xi \cdot Z_{jt} + \varepsilon_{jt} \qquad (5)$$

The matrix Z includes the same control variables as equation (4), hence, we allow for industry, region, ownership, time, and, in some cases, firm fixed effects to shift the company's wage bill[13].

2.3. Estimation specification and issues

We aim to study the evolution of the relative productivities and wages through the time interval of 1986 – 2005. Hence, we divide our sample into four time periods and run separate regressions using the pre-transitional years of 1986, 1989; the years of transitional recession, 1992-1995; the third period covers 1996-2000; and our most recent sample includes the years from just before the EU accession and the accession itself, 2001-2005.

Relative productivities and relative wages can be estimated directly from equations (4) and (5) via nonlinear least squares. However, the results presented in this paper are estimated based on a linear approximation of this non-linear equation.[14] In this case, the proportions of workers in different groups are treated as separate inputs besides K and L, and the production function we estimate becomes:

---

[13] Capital may or may not be included in the plant-level wage 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 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 capital control in our baseline specification. However, we find that the results do not change significantly if we exclude capital in our specification.

[14] Assuming that $(\varphi_F - 1)\frac{L_F}{L} < 0.1$ holds, the linear approximation is:

$$\ln\left[1 + (\varphi_F - 1)\frac{L_F}{L}\right] \approx (\varphi_F - 1)\frac{L_F}{L}.$$

$$\ln VA_{jt} = \alpha_0 + \alpha \cdot \ln K_{jt} + \gamma \cdot \ln L_{jt} + \phi_F \frac{L_{F_{jt}}}{L_{jt}} + \phi_O \frac{L_{O_{jt}}}{L_{jt}} + \phi_F \frac{L_{U_{jt}}}{L_{jt}} + \delta \cdot Z_{jt} + u_{jt} \qquad (6)$$

After linearization, the group share coefficients cannot be directly interpreted as relative productivities, they only indicate how the different worker groups contribute to the production process. However, the estimated values can be translated into relative productivities by dividing the worker share coefficient with the coefficient of the labor (L) term[15].

Similarly, the linear wage equation:

$$\ln W_{jt} = \eta_0 + \eta \cdot \ln K_{jt} + \theta \cdot \ln L_{jt} + \psi_F \cdot \frac{L_{F_{jt}}}{L_{jt}} + \psi_O \cdot \frac{L_{O_{jt}}}{L_{jt}} + \psi_U \cdot \frac{L_{U_{jt}}}{L_{jt}} + \xi \cdot Z_{jt} + \varepsilon_{jt} \qquad (7)$$

Relative wages can be computed in the same way as relative productivities: by dividing the group share coefficient with the coefficient of the labor term.

The linear approximation of the logarithmic term is widely applied in the literature[16]. Moreover, after estimating the same equations with both the linear and nonlinear method, we found that the change in relative productivities and relative wages over the four time periods follows similar pattern in both cases, hence, our conclusions about the change in the parameter estimates are unaffected by the linearization[17].

---

[15] The following relationship holds between the labor share coefficients of equations (4) and (5): $\gamma(\varphi_F - 1) = \phi_F$.

[16] Several studies following the pioneering work of Hellerstein, Neumark and Troske (1999) and Hellerstein and Neumark (1999) applied the linear approximation. Such studies include, e.g. Dostie (2006), Ilmakunnas and Maliranta (2003), Crepon et al (2002), Borowczyk and Vandenberghe (2010) or Vandenberghe and Waltenberg (2010).

[17] Our results of relative wages are almost unaffected by the choice of linearization, but relative productivities computed with the linear method tend to be smaller than the nonlinear estimates, especially in the case of the college graduates. For example, using nonlinear estimation, the relative productivity of college graduates to employees without diploma is 5.4 in the OLS specification in 2001-2005, while the similar computed value in the linear version is 2.1. The departure between the linear and nonlinear FE estimates is much smaller. However, as relative wages are almost identical using both methods and linearly computed relative productivities tend to be smaller, our wage – productivity gap can be considered as the lower bound of the gap obtained with nonlinear method. Moreover, nonlinear methods are usually less robust to outliers, and the extreme values may also be the result of the sensitivity of the nonlinear estimation method.

In our baseline specification, we estimate equations (6) and (7) jointly for each time period via OLS. In this case, the parameter estimates are identified by between firm differences. However, it is possible that some of the observed productivity and wage differential is due to the selection of workers into better (high productivity, high wage) or worse (low productivity, low wage) firms. To separate observed productive differences into the part that is due to selection of workers into good or bad firms, and productivity differences within firms, we run the same regressions including firm fixed effects. Besides taking care of the systematic selection of workers, including firm fixed effects also accounts for time invariant unobserved productivity shocks. Time variant unobserved productivity shocks are generally tackled by the method developed by Olley and Pakes (1996) and Levinsohn and Petrin (2003), by including a proxy composed of material costs and capital[18]. However, the results using this technique do not differ significantly from the results obtained without the correction, so we will only present one set of results – those without – that are robust to this difference in specification.

In terms of our samples, we first obtain average estimates using the pooled sample of firms. Then, we split the database into four broadly defined industry categories: agriculture, heavy industry, light industry and services and obtain industry-specific estimates, to allow for differences between them. Finally, we run separate regressions using the sample of firms which existed prior to transition, termed "old firms", and on the sample of firms set up after transition, termed "new firms" to see if the two groups behave differently. This last set of regressions was only estimated in the periods after 1992, since new firms do not exist prior to the transition.

An interesting attempt of our paper is to control for firm-specific, time invariant unobservables, and obtain within-firm estimates of relative productivities and wages. This was rarely done in previous studies, mostly due to data limitations. Firm-level selection is likely to

[18] 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:

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

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

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

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

play an important role in determining wage gaps. Although our data allows us to control for firm fixed effects, Haltiwanger, Lane and Spletzer (1999 and 2007) point out its identification difficulties due to the small within-firm variation of the group shares. They draw the attention to the stylized fact that labor productivity, earnings per worker and workforce composition are quite heterogeneous across firms and quite persistent within firms. Using US data from 1986 – 1996, they find that the first order AR coefficient from a regression of the proportion female in 96 on its 86 value is 0.82; for college graduates the value is 0.45, for productivity the coefficient is 0.47 and for earning it is 0.45 even after removing industry-year means. While their results of regressing labor productivity on worker group shares without including firm fixed effects is in line with previous empirical evidence of significantly lower productivity of female than male employees, lower productivity in the worker group above 55, and better productivity of college graduates, coefficients lose their significance in the first differenced specification, and even the signs are occasionally opposite of what is expected. The identification difficulties of the first differenced equation were also emphasized by Hellerstein and Neumark (1998).

In our data we also find considerable persistence in the worker composition of the firms, suggesting that the fixed effects results should be interpreted with caution. The first order AR coefficient of the proportion female in 1996 on its 1992 value is 0.61 after removing industry means. The same coefficient for the college graduates is 0.69 and for workers above 40 it is 0.50. Despite the poor identification perspectives of the within-firm specification, we estimate all the regressions including firm fixed effects. As expected, our estimates of relative productivities and wages are not significantly different from 1 in some cases, but we do get significant estimates with signs that are mostly in line with our expectations and are similar to those obtained in previous studies. The identification problem is especially troublesome in the first period, as the sample size is small. Although the FE results are subject to these problems, we look at them as the strictest test for the existence of a gap between wages and productivity.

3. 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.

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 8th grade, elementary, high school, vocational, university), and occupation (4 digit occupational code). For the purposes of defining the various worker groups, we define two age categories (under 40, over 40), two education categories (college or no college), and based on gender. The firm variables used in the estimation are firm output, capital, material costs, employment, and wage bill from the company's Tax Authority data, and industry, region, size, and ownership as controls.

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 1 gives the summary statistics of the firm-level sample for select years between 1986-2005.

4. Results

We now turn our attention to the between- and within-firm estimation results. First, we discuss the results for the three worker groups investigated in this paper separately, reviewing the trends in their relative productivity and wage estimates over time, and the evolution of the gap between the two for the pooled case, by broad industry, and for firms that existed before the transition and those who entered after, as well as for foreign-owned firms. These are compared to international results from studies employing the same methodology. We then summarize the results in light of the hypothesis regarding the effects of the transition on firm behavior, to see whether we can see a movement towards more efficient employment and wage setting practices as a result of the changes in the market.

Gender

While studies using wage equation methods usually find a large unexplained wage gap between men and women, previous international empirical results based on cross-sectional estimates using the method in this paper usually document a much smaller 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.[19] Panel estimates tend to find smaller discrepancy between the productivity and wage of male and female employees and point to no wage discrimination against female employees[20].

---

[19] 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) use Israeli data and Dong et al (2009) use Chinese data to show a 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.

[20] For example, Borowczyk and Vandenberghe (2010), using Belgian data of 1998-2006, found that women are 5-13 percentage points less productive than men, and women earn 7 percentage points less than men resulting in no wage discrimination against women. Crepon et al (2002) also concluded that there is no wage discrimination

Our results regarding female workers in the OLS specification are shown as the calculated relative productivities, wages, and gaps shown below in Figures 1.a-1.c.

In line with previous empirical results, our pooled sample OLS estimates suggest that firms with a higher ratio of female workers are less productive in most industries. The estimated relative productivities suggest that in the initial period, female workers were about 20 percent less productive than males, and their relative productivity decreased over time. By the last period, 2001-2005, the OLS estimates suggest that women were 50 % as productive as male employees.

Splitting the sample into broad industries reveals considerable heterogeneity in the coefficient estimates.[21] The largest negative effect is found in the service sector, where the relative productivity of female employees decreases to 0.36 and 0.16 in the later periods. As expected, the female-male relative productivity is low in the heavy industry sector as well, with computed values ranging in the interval of 0.3 to 0.6. In the light industry, the relative productivity of female workers shows a U-shape curve over time: drops from 0.93 to 0.59 in the second period (1992-1995) and catches up gradually and becomes insignificantly different from 1 by the last period. In the agriculture sector, none of the coefficient estimates are significant, and computed relative productivities are not significantly different from 1. Hence, from the disaggregated results we can see that the declining negative trend of female productivity in the labor market as a whole is mostly due to the sharp decrease in the service sector; females are less productive than males in both manufacturing sectors, while in the agricultural sector there is no significant difference between the productivity of female and male workers throughout the whole period.

There is much less heterogeneity in the wage profiles by industries: after a sharp drop in the first period, relative wages are quite stable throughout 1992 - 2005. On average, female are paid less, and the computed relative wages show that females receive 10-15 percent lower wages than males. The more detailed picture by industries confirm the negative female-male wage gap

against women, and estimated that women are 11 percent less productive than men and are also paid 14 percent less using French manufacturing data of 1994-1997.

[21] Of course, the estimation carried out by industries is based on a much lower number of observations than the pooled estimates. The results are more prone to noisiness as can be seen from the more extreme values.

with lower relative wages of 0.75 - 0.85 in the heavy and light industries, and relative wages around 0.9 in the agriculture and services.

Comparing productivity and wage profiles gives a more accurate picture of the causes of the wage differences. Also, by testing the equality of the computed relative productivities and wages, we can get an idea about their degree of departure from the competitive state.

On average, the gradual decrease of productivity seen in the OLS estimates was not followed by similar drop in relative wages, implying a positive and increasing wage – productivity gap. Assuming that the simplifying assumptions are valid, this means that women are paid more than their productivity merits. The wage – productivity gap is the largest in the service sector; it is also positive in the heavy industry sector, while agriculture and light industry show the smallest discrepancies, with insignificant gaps by the last period. Regarding our hypothesis of a "learning effect throughout the transition", we can see the pattern of closing gaps in the manufacturing sectors and in the agricultural sector, but no clear learning pattern in the OLS results for the services or in the pooled results.

The within-firm estimates (FE) are depicted in Figures 2.a-2.b. The estimates confirm much smaller productivity and labor cost impacts, underlining the importance of firm level selection in the labor market. On average, the productivity profile is upward sloping over time, with a small negative impact in the 1992-1995 and 1996-2000 periods, and a positive impact in 2001-2005. A comparison of the between-firm and within-firm estimates suggests that women are grouped into less productive and lower-wage firms, and this selection effect explains to a large extent the significantly lower productivity of women in the OLS estimates. This systematic selection is especially pronounced along the productivity dimension. For example, OLS estimates suggest that women are half as productive as men in 2001-2005, while FE estimates reveal that within firms, women are more productive than men in the same period.

In the separate industry estimates, the grouping of women into less productive firms is most pronounced in the heavy industry and services. There is much less heterogeneity by industries in the FE specification, and many of the relative productivity estimates are not significantly different from 1. Women are significantly less productive than men in the light industry and services sectors in 1996-2000, and significantly more productive than males in the light industry in the last period. These results drive the patterns seen in the pooled estimates.

As for the wage profile, a somewhat similar pattern is found in the FE case: women have lower wages than men in the 1992-1995 period, but their relative wages increase over time, and they receive more than men in 2001-2005.

The comparison of the within-firm productivity and wage profiles of women shows that the upward sloping productivity profile is accompanied by an upward sloping wage profile, implying an insignificant gap in most cases. Figure 7 shows that there is no significant gap within firms in the later time periods, in line with previous literature that suggests that this methodology usually points to less wage discrimination than the estimates derived from traditional wage equation coefficients. Examining the gap by industry, also suggests some closing of the gap following the transition. The greatest change is after the initial period, in the early years after the transition, but the trend continues in the later periods as well: the gap disappears by the last period in all cases except in the light industry, while in 1992-1995 the gaps show a diverse pattern with value of -0.3 in the heavy industry and 0.4 in the light industry.

Figures 3-4 show the results when firms are divided into those that existed prior to the transition (old firms) and those that entered the market after 1990 (new firms). As the pooled estimates are very close to those of new entries, while the results for old firms are significantly different, we graph only the wage and productivity schedules of old firms in comparison to the pooled estimates[22].

The relative productivity of female workers in old firms is about 0.2 higher in all periods than it is in new firms in the OLS specification, though their trends over time are similarly decreasing over time. The FE or within-firm results show an important change in the final period: while women's relative productivity increased to above 1 in new entrants, it decreased sharply in the old firms. In terms of relative wage, women are paid higher wages in old firms in both the OLS and FE cases, but within firms, the relative wage of women in new firms catches up to that paid in old firms.

The gap between relative wage and productivity of women increases in the between-firm case, for both types of firms, though it is lower for old firms. The within-firm wage gap, which is

[22] The similar estimates of the new firms' sample and the pooled database are due to the relatively small number of old firms. For example, in the regressions of 1996-2000, we had 1,919 observations in the old firm sample, while 13,840 observations of new firms. Nevertheless, the markedly different behavior of old firms and new entries deserves special attention.

insignificant in new firms following the transition, increases sharply in the case of old firms. This is due to the fact that while they pay the same relative wage as new firms, the productivity of women in old firms fell sharply in the last period. These results suggest that, contrary to what was previously found in non-transitional economies, where new firms learned the better practices of successful older firms over time, in Hungary new firms behave more efficiently than firms that were already established prior to the transition. The fact that the relative productivity of females is much higher in new firms in the last period suggests better sorting and matching of female workers in new firms, and the fact that their relative wages reflect their relative productivity is in line with more efficient wage setting in new entrants. It seems that old firms pay a relative wage to women that is similar to that in new firms, suggesting a convergence in wage setting among the firm types, but they are not benefiting from a productive contribution that is in line with the relative wage or with that seen in new firms.

Education

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.[23] The relative wages and productivities computed from OLS estimates are depicted by Figures 5.a-5.c.

[23] 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. Panel databases including educational information on employees are rarely available, instead, the occupation of the employee is used. For example, Crepon et al (2002)

OLS estimates suggest that on average, workers with a diploma are twice as productive as workers without higher education. The estimates show an upward sloping productivity profile in each industry, with the highest jump right after the transition. Before 1989, on average, the computed relative productivity of workers with a diploma was 1.17, while in the following period it increased to 1.93. Hence, it seems that a higher share of graduated employees became more and more an essential tool for achieving higher productivity. Looking at the separate industries, we find some heterogeneity in terms of the strength of this positive relationship, but all the estimated relative productivities are high and significantly different from 1. The positive relationship is largest in the manufacturing sectors, and it is slightly smaller in agriculture and services.

OLS wage profiles also show an upward sloping pattern, however, the increase of relative wages over time was not enough to compensate for the jump in productivity, resulting in a growing and negative wage – productivity gap in the pooled estimates. Interestingly, by industry, wages are most in line with productivity in the services sector and depart the most in the light industry. However, by the last period, the discrepancy between relative wages and productivity is very similar in each sector except for services where the gap is smaller. Hence, comparing relative wages and productivities we do not find evidence of a closing gap, instead, we observe that remunerations are less and less in line with productivity over time. However, the widening gap is not due to wages not following productivities at all, but is the result of wages not compensating quickly enough for the jump in productivity.

The above results are relatively stable in the within-firm specification: we tend to obtain positive productivity and labor cost effects and negative wage – productivity gaps, though the magnitudes are much smaller and the estimates are not significant in all cases. Relative productivities and wages computed with the FE method are shown by Figures 6.a-6.c.

The drop in the magnitudes of the coefficients compared to the OLS specification suggests that college graduates tend to be employed by more productive and better-paying firms. For example, the double relative productivity of workers with diploma drops to 1.25 in the FE specification in 2001-2005, but it is still significantly different from 1.

group workers into unskilled, skilled and highly skilled categories, and find that highly skilled are underpaid, and unskilled are overpaid relative to the skilled category.

Splitting the sample by industries shows that the positive productivity impact is larger in agriculture, suggesting a lower level of selection, and is close to 1 in the services. However, the insignificant parameter estimates in the service sector are probably due to the poor identification possibilities and are not the sign of graduated workers being as productive as workers without diploma. The only odd result was found in the heavy industry sector, where the computed relative productivity of workers with a diploma is 60 percent of those with elementary education in the period of 1996-2000[24].

The within-firm labor cost impacts differ more by industries and again point to the poor identification possibilities of the fixed effect estimation. While computed relative wages are mostly greater than 1, we obtained relative wages of 0.6 – 0.8 in the light industry and agriculture sectors in the periods of 1992-1995 and 1996-2000.

As relative wages do not catch up with relative productivities, the FE estimates also confirm negative wage – productivity gap in most cases. The negative gap is significant in the last period on average and in the sample of agricultural firms, and the sign is negative, but statistically not different from zero in the other sectors. The closing of the gap is most pronounced in the light industry, while a widening gap can be observed in the agriculture.

Figures 7-8 show the OLS and FE estimation results dividing the pooled sample into old and new firms. In terms of firm age, we again see a significant difference between pre-transition firms and new entrants. The relative productivity of college graduates is lower in old firms in the OLS case, but we see an increase in relative productivity to well above that in new firms in the FE specification. The productivity estimate rises to almost 2 in old firms in the last period, compared to 1.25 in new firms. There is a drop in relative wages of college graduates in both the between- and within-firm cases during the same time, leading to a sharp rise in the gap between wages and productivity in old firms. While college graduates are underpaid in the new firms as well, the gap is much larger in old firms, again pointing to wage setting practices that are not in line with efficient behavior relative to what we see in new firms.

[24] Another odd result is the negative relative productivity of college graduates in the years of 1986, 1989 in the agriculture. However, this is clearly the sign of identification difficulties in the first period: small number of observations (783) and low share of college graduates.

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. 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 implies that the older employee group receives a wage premium.[25] 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; Lallemand and Rycx, 2009).[26]

Our results are in line with estimates obtained in the international literature: we estimate that the ratio of workers above 40 is negatively associated with firm-level productivity, especially in the OLS specification. Comparing relative wages and relative productivities, we find a positive wage – productivity premium for more experienced employees in most specifications. Similar results were found by Crepon et al (2002) who estimated a positive wage – productivity gap for workers above 35. Our OLS estimates are shown by Figures 9.a-9.c.

[25] 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 finds that prime-age workers receive 5 percentage point less than their productivity of 1.21. Vandenberghe and Waltenberg (2009) using Belgian data estimates in a within-firm specification a positive wage premium for older workers with relative productivity of around 60 percent compared to prime age workers and relative wages not significantly different from one. 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. Crepon et al (2002) also identifies a positive wage – productivity gap for workers above 35 in a within-firm specification.

[26] 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.

Except for the first period, coefficient estimates are significant, large in magnitude, and, on average, translate into a relative productivity of 0.5. In 1986 – 1989, workers above 40 did not have significantly different productivity from workers below 40. The sharp drop right after the transition may be the sign of skill obsolescence[27], or the rapid segregation of younger workers into more productive firms, or may be simply the sign of the poor identification possibilities (due to the lower number of observations) in the first period.

In terms of the heterogeneity of estimates by industry, the computed relative productivities are diverse in magnitude, and lie in the range of 0.23 to 0.76 in the periods from 1992. However, the trend of a sharp decrease after the first period and staying significantly below 1 after 1992 is common in all industries. By 2001-2005, workers above 40 are the most productive in the heavy industry, with a computed value of 0.75, and they are the least productive in the agricultural sector with a relative productivity of 0.3.

The labor cost effects mostly have a positive sign, and the estimated coefficients translate into a relative wage of 0.95 – 1.09. The result of relative wages exceeding productivity is in line with many previous empirical results finding that more experienced employees receive a wage premium. The wage profiles are rather flat over time in all industries, and there is no sign that the drop is productivities was followed by a similar change in relative wages. The gap profiles do not show considerable heterogeneity by industries: the gap widens in each case from being insignificant to a positive level.

The result of smaller relative productivities that are smaller than one, somewhat higher relative wages and a small positive wage – productivity gap is confirmed by the within-firm estimates. FE estimation results are depicted by Figures 10.a-10.c.

As expected, the small relative productivities obtained in the OLS regressions are mostly due to selection of workers into less productive firms. Though relative productivities are still significantly smaller than 1 on average, computed values lie in the range of 0.9 – 0.93 in the pooled sample estimates. Industry-specific values are less diverse: they range from 0.8 to insignificantly different from 1.

Looking at the within-firm wage results, we find evidence of a slight selection of more experienced workers into better paying firms. Relative wages are somewhat smaller than in the

[27] For a detailed study on this topic, see Kertesi-Köllő (2002), who examine relative wages and the underlying changes in productivities of workers grouped by skill level and age following the transition.

OLS specification. For example, the pooled sample relative wage of 1.085 in 2001-2005 dropped to 0.974 in the within-firm specification. The pooled sample relative wages are all significantly smaller than 1, however, the industry-specific values are mostly insignificant, but below 1 in magnitude.

Comparing within-firm relative productivities and wages, the gap is mostly positive, and smaller in magnitude than the between-firm values. We find a significant positive gap in our pooled sample in the last period, in the light industry in 1996-2000, and in services between 1996 and 2005. We do not see evidence pointing towards the closing of the gap, most values are close to zero but positive throughout the whole period.

The comparison of old and new firms again confirms significant differences between the two groups. The OLS and FE results are shown in Figures 11-12.

The relative productivity of workers over 40 shows an interesting rise after the second period, followed by a drop by the last period. The pattern of differing relative productivities between new and old firms before 2000, and much closer values after 2000 shows that over time, old and new firms became increasingly homogeneous in terms of valuing experience, and skill obsolescence is less and less of an issue in the later years. Older workers seemed to be more productive in older firms following the transition, leading to a sharp drop in the gap between their wage and productivity. In this period, older workers in older firms were actually underpaid compared to their productivity contribution to the firms. However, their productivity fell in the final period to a level close to that in new firms, so their relative wage is more in line with their productivity, with a gap close to 0 in the within-firm specification, and, as in new firms, positive and significant in the between-firm case.

Summary of Evidence on Competitive Implications

The main goal of our analysis is to find evidence regarding the effects of the transition on the efficiency of firm behavior regarding the employment and wages of the three worker groups differentiated by gender, education and age. Increased efficiency should translate into a closing of the wage-productivity gaps, if firm wage setting is increasingly based on the productivity of workers. Non-efficient behavior, such as discrimination, results in a gap between wages and productivity, and in a deviation from the profit-maximizing composition of the workforce. At the same time, it is interesting to see if firms that entered the market after the transition are significantly different from older, pre-transitional firms, and if their behavior is closer to more efficient outcomes. If there is learning over time, then in the transitional context, we may expect older firms to eventually adopt more market-based practices that are in line with what we see in new firms that were not hindered by already existing distortions during socialism.

Our results are mostly not in line with an increase in efficiency in firm practices. For women, within firms, the gap between relative wage and productivity decreased to 0, suggesting that firms do not discriminate against women within firms, though there is evidence of a negative selection of women at the firm level that may indicate a bias in hiring practices. In firms that existed prior to the transition, the within firm gap increased in the last period. This increase is mainly due to a sharp drop in female productivity in old firms, while the wage setting practices converged to that of new firms.

In all of our specifications, college graduates are underpaid relative to their productive contribution, especially in old firms. The wage-productivity gap did not disappear over time for this group, but it is significantly higher in the case of old firms, again suggesting that new firms behave more efficiently in the market.

Workers older than 40 are slightly overpaid, though less so in old firms, where their relative productivity is higher than in new firms. The wage premium of older workers is in line with results in previous studies, and tends to persist in competitive markets as well. Though older firms may appear more efficient than newer ones in this respect, the lower gap in old firms is due to the higher relative productivity of older workers in these firms, not an adjustment of the relative wage, which is actually higher (better) in old firms. Overall, new firms do appear to have significantly more efficient practices than pre-transition firms in terms of setting wages closer to productivities. Except for workers differentiated by age, old firms do not tend to approach the more efficient wage setting practices of new firms. While the above 40 – below 40 wage – productivity gap of old and new firms got very close to each other by the last period, the female – male, degree – no degree gaps still depart from each other even in the later years.

5. Conclusion

In this paper, we use a linked employer-employee dataset from Hungary covering the years 1986-2005 to assess the evolution of relative productivities and wages of various worker groups over time. During this period, Hungary underwent a rapid economic transition, and joined the European Union, changes that should lead to more efficient wage-setting and hiring practices following the distortions of the socialist system. We examine the situation of women relative to men, college graduates, and workers over the age of 40, and also control for occupation as well as various firm characteristics. Following the methodology of Hellerstein and Neumark (1999), we estimate the group-level relative productivities and wages of these groups based on firm-level production functions and wage equations. We are able to estimate these using OLS as well as a firm fixed effects specification, which allows us to examine the role of firm-level selection as well for four time periods and by industrial categories and firm age.

Our results provide little evidence suggesting that firm behavior became more efficient following the transition. For women, we find a significantly lower wage gap than previous literature based on traditional wage equation estimation and decomposition methods. Within firms, the gap between relative productivity and wage of women decreased to zero, and women's productivity is not lower than men's.

College graduates are underpaid compared to their productive contribution, though less so in new firms, again suggesting that the new firms that entered the market are more efficient than older firms that were possibly constrained by their already existing practices.

In terms of workers over the age of 40, we find that there is a wage premium in all types of firms, which is in line with findings in previous literature. The wages of this group reflect their productivity more closely in older firms due to the fact that older workers have a higher relative productivity, suggesting that their acquired skills and experience were better suited for older firms, and less productive in newer firms.

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.

Becker, Gary (1957). The Economics of Discrimination, Chicago: University of Chicago Press.

Borowczyk Martins. D. and V. Vandenberghe (2010). Using Firm-Level Data to Assess Gender Wage Discrimination in the Belgian Labour Market. *IRES WP*, No 2010-7, Economics School of Louvain, UCL, Louvain-la-Neuve.

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.

Brown, David J., and Earle, John S. (2008). Understanding the Contributions of Reallocation to Productivity Growth: Lessons from a Comparative Firm-Level Analysis, Upjohn Institute Staff Working Paper 08-141.

Campos, Nauro and Jolliffe, Dean (2005). Does Market Liberalization Reduce Gender Discrimination? Econometric Evidence from Hungary 1986-1998, Labour Economics, Vol. 12, No.1. pp.1-22.

Crépon, Bruno, Deniau, Nicolas and Sébastien Pérez-Duarte (2002). Wages, Productivity, and Worker Characteristics: A French Perspective, CREST WP 2003-4.

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.

Griliches, Zvi (1977). Estimating the Returns to Schooling: Some Econometric Problems, Econometrica, Vol. 45, Number 1.

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.

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

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.

Kertesi, Gábor and Köllő, János (2002). Economics Transformation and the Revaluation of Human Capital – Hungary 1986-1999, Research in Labor Economics, Vol. 21. Pp.235-273.

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.

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.

Ours, J.C and L. Stoeldraijer (2010). Age, Wage and Productivity, IZA DP No. 4765.

Van Biesebroeck, Johannes (2007). Wages Equal Productivity. Fact or Fiction?, University of Toronto, Dept of Economics, WP 294.

Vandenberghe, V. and F.D. Waltenberg (2010). Ageing Workforce, Productivity and Labour costs of Belgian Firms, *IRES WP*, No 2010-3, Economics School of Louvain, UCL, Louvain-la-Neuve.

Tables and Figures

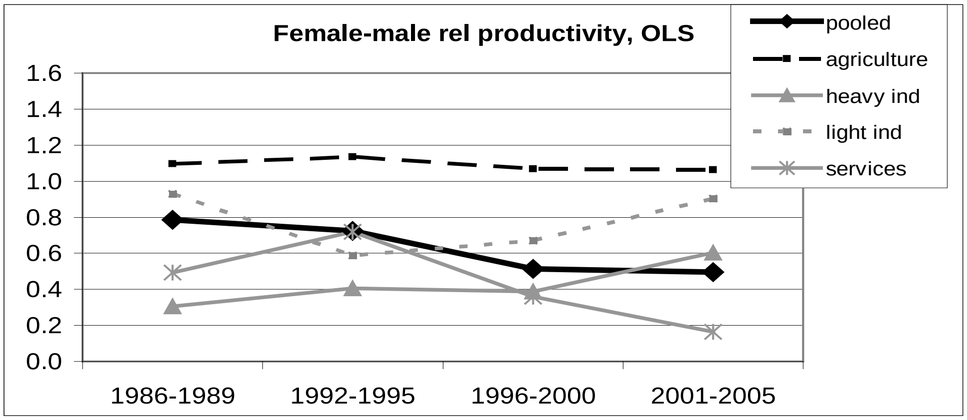
Table 1: Summary Statistics of Firm-Level Sample, 1986-2005

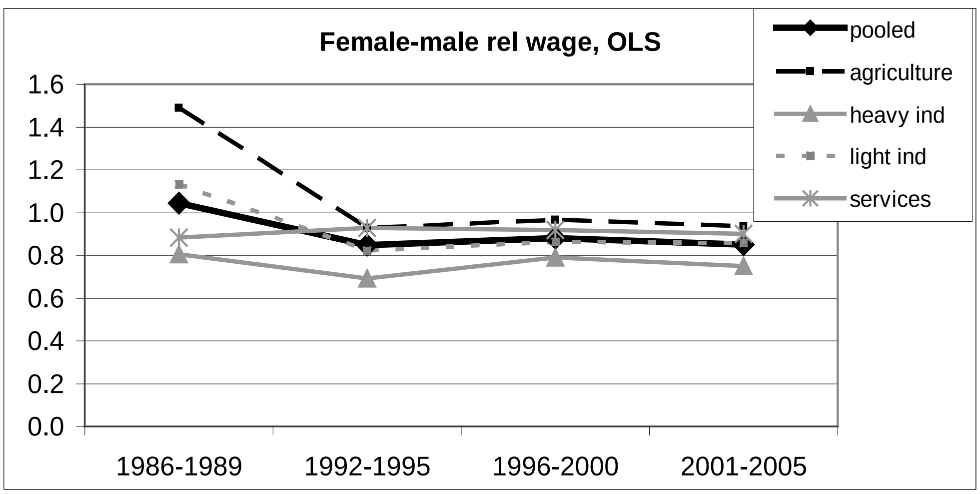
Year	**1986**	**1989**	**1992**	**1995**	**1998**	**2001**	**2005**
Number of firms	730	2,253	3,044	4,911	4,922	5,845	4,357
Productive inputs							
ln(Output)	7.49	7.28	6.63	6.42	6.47	6.44	6.92
ln(Capital)	6.11	5.80	5.31	5.08	4.94	4.82	5.49
ln(Materials)	6.33	6.16	5.30	5.11	5.17	5.38	5.89
Employment	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

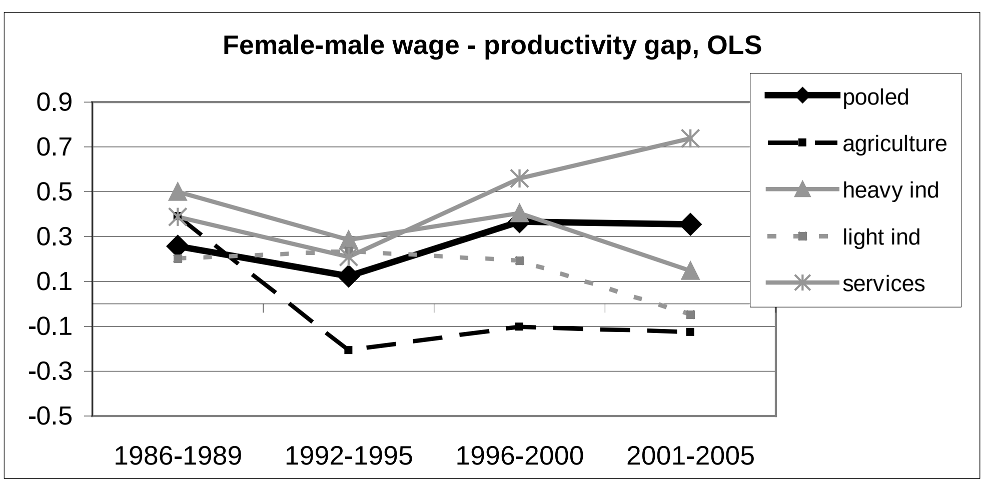
Source: Hungarian Wage and Employment Survey, 1986-2005.

Note: Sample restricted to private sector firms with at least 20 employees, and 5% of their workforce included in the employee dataset. Values represent means.

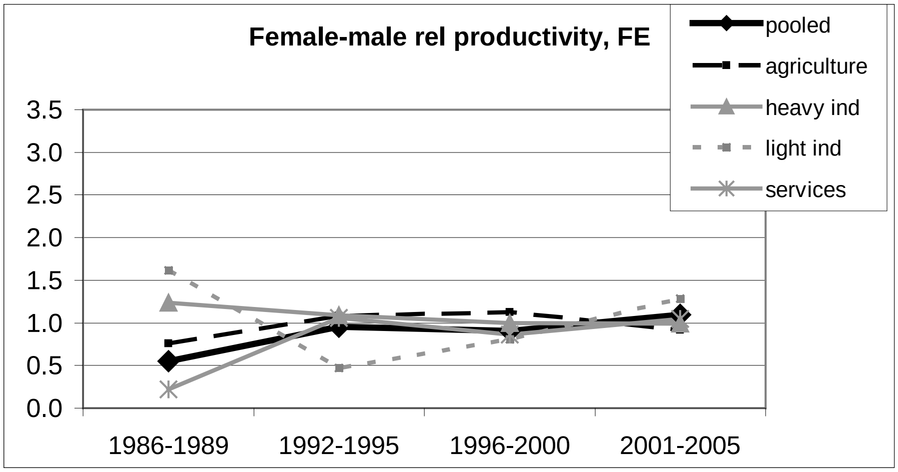
Figures 1.a-1.c: Female-male relative productivities, wages, and gaps - OLS, pooled sample and industry subsamples

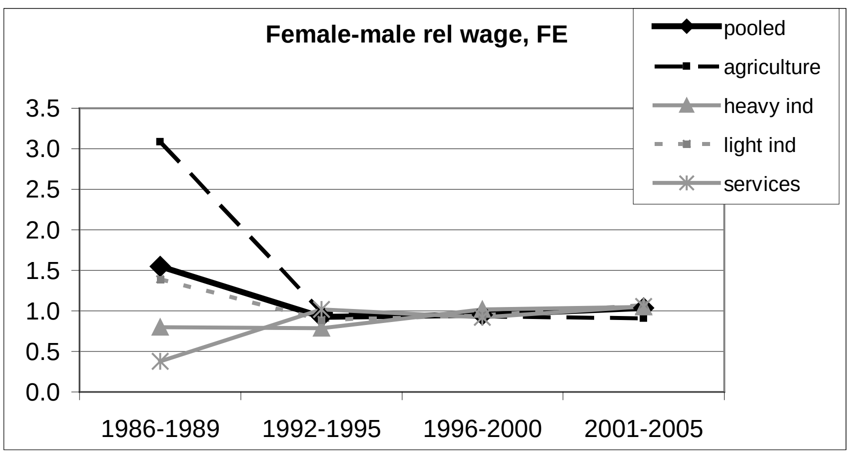






Figures 2.a-2.c: Female-male relative productivities, relative wages, and gaps - FE, pooled sample and industry subsamples





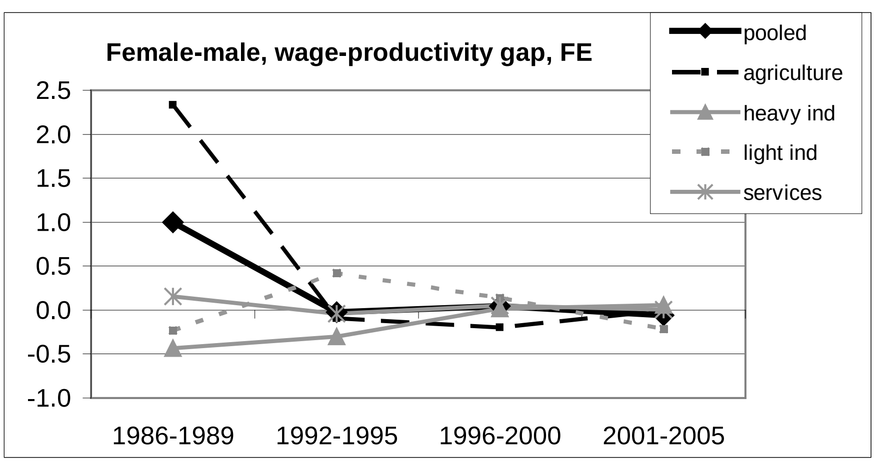
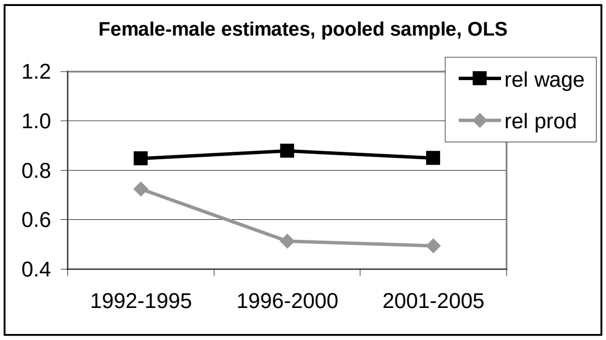


Figure 3: Female-male relative productivities and relative wages - OLS, pooled sample and old firms.



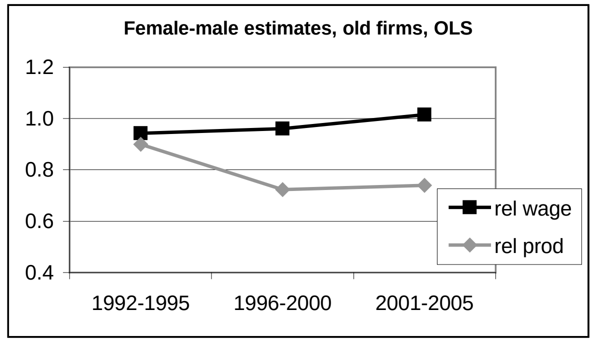
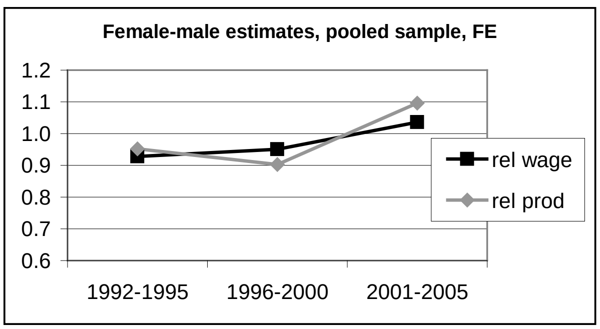
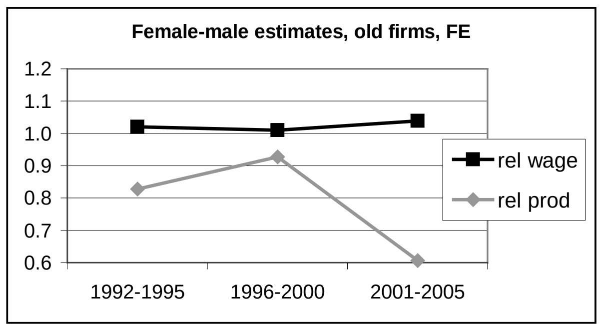
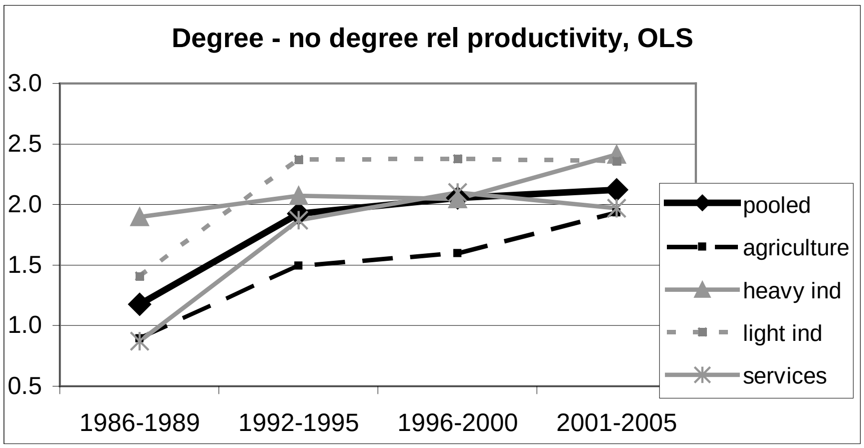


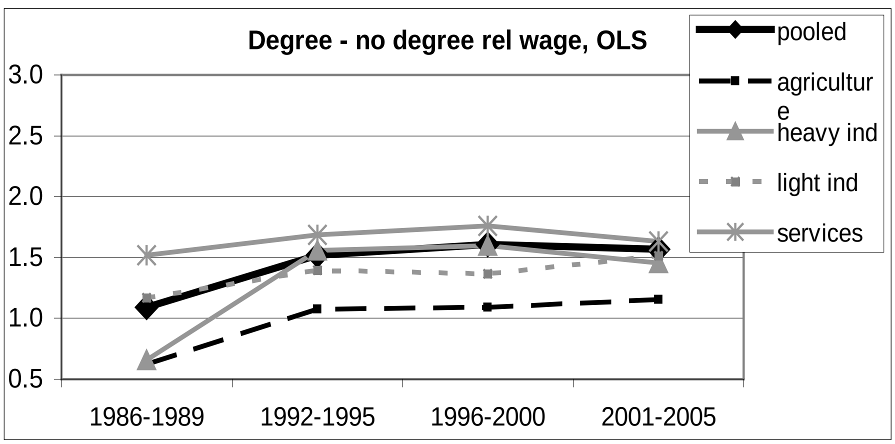
Figure 4: Female-male relative productivities and relative wages - FE, pooled sample and old firms.

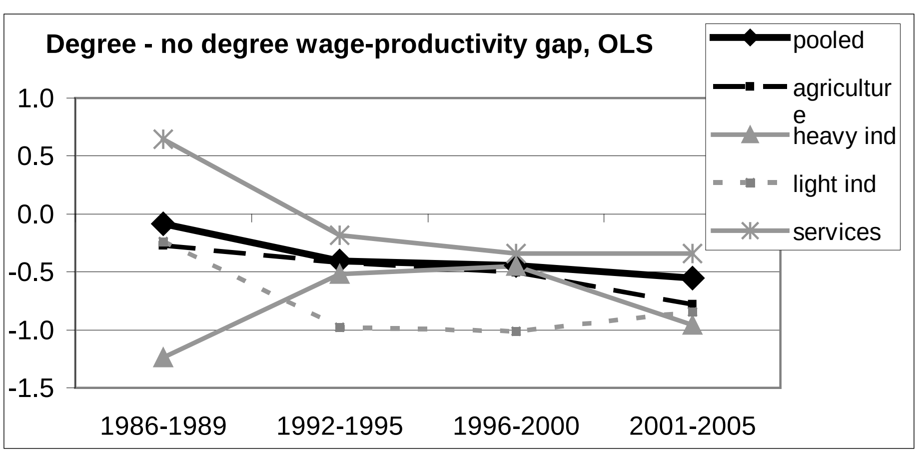




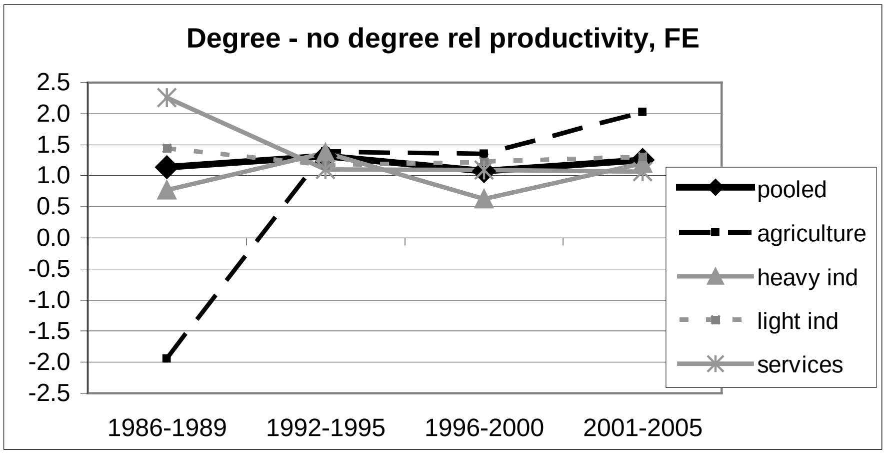
Figures 5.a-5.c: Degree – no degree relative productivities, relative wages, and the gaps - OLS, pooled sample and industry subsamples

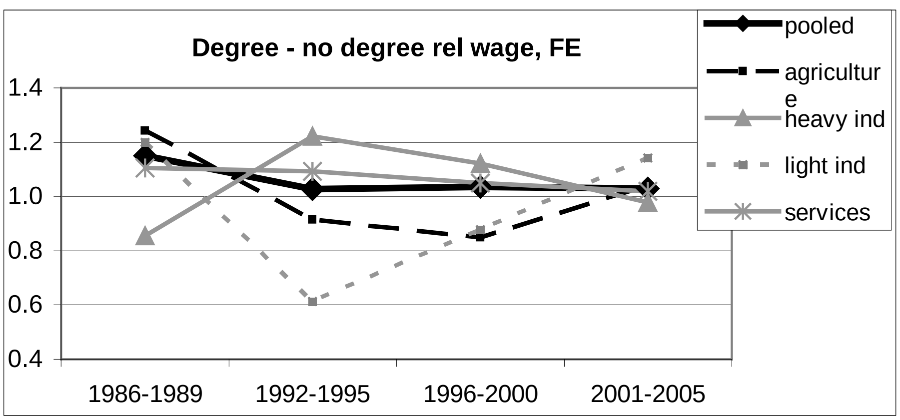






Figures 6.a-6.c: Degree – no degree relative productivities, relative wages, and the gaps - FE, pooled sample and industry subsamples





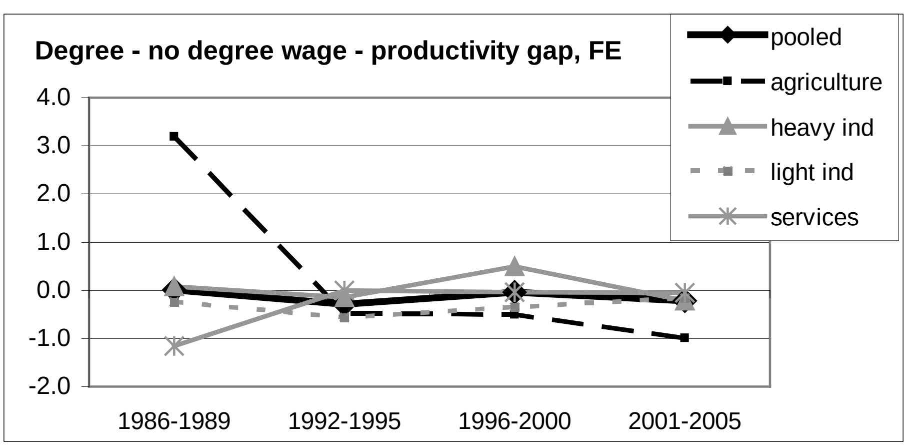
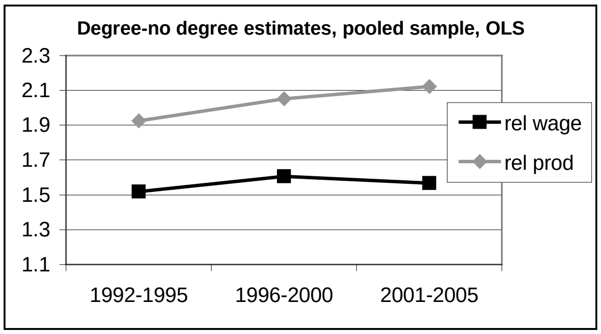


Figure 7: Degree – no degree relative productivities and relative wages - OLS, pooled sample and old firms



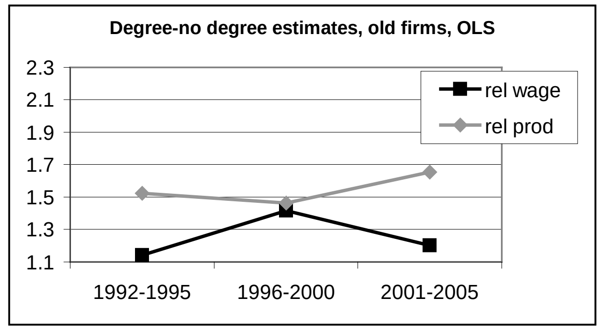
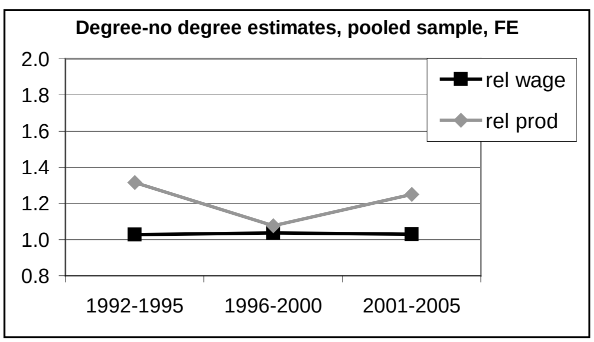
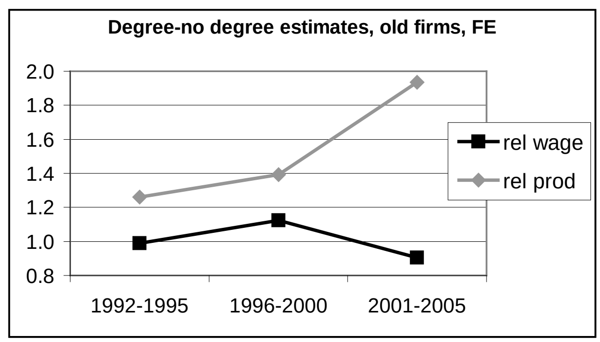
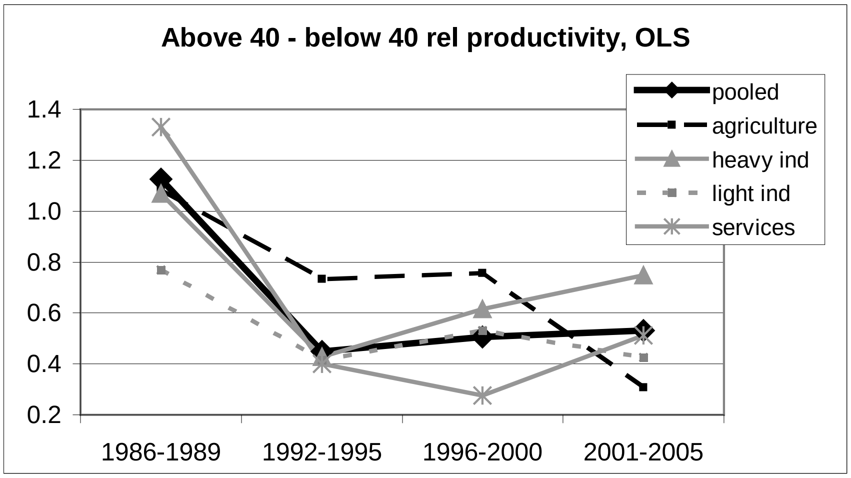


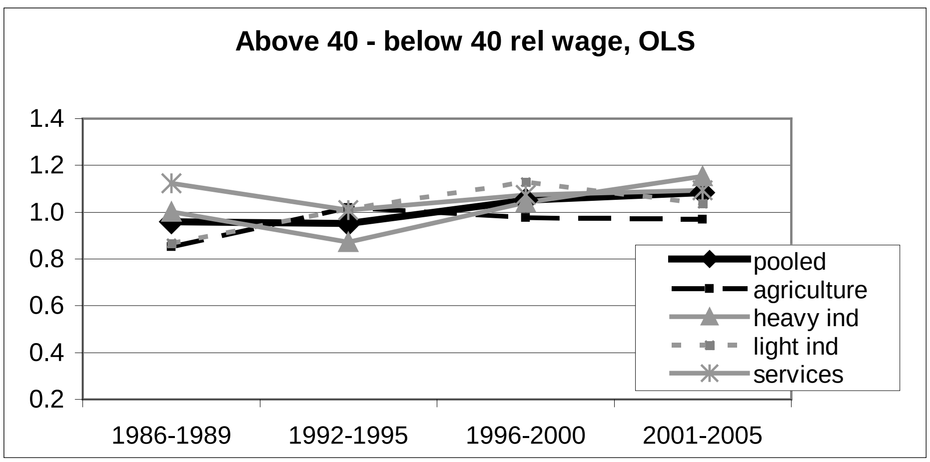
Figure 8: Degree – no degree relative productivities and relative wages - FE, pooled sample and old firms

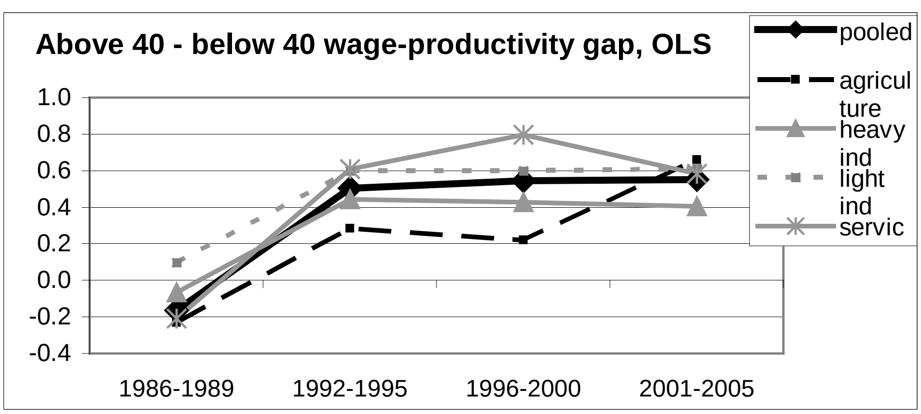




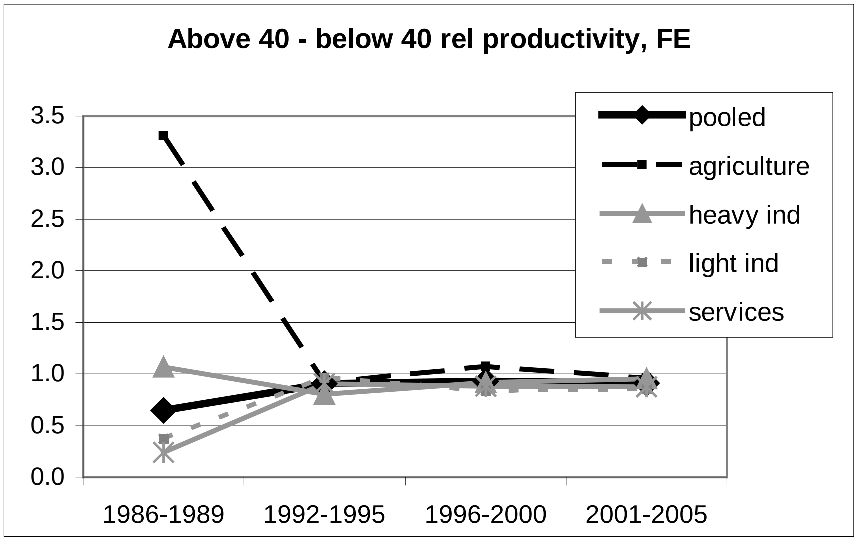
Figures 9.a-9.b: Above 40 – below 40 relative productivities, relative wages, and the gaps - OLS, pooled sample, industry subsamples

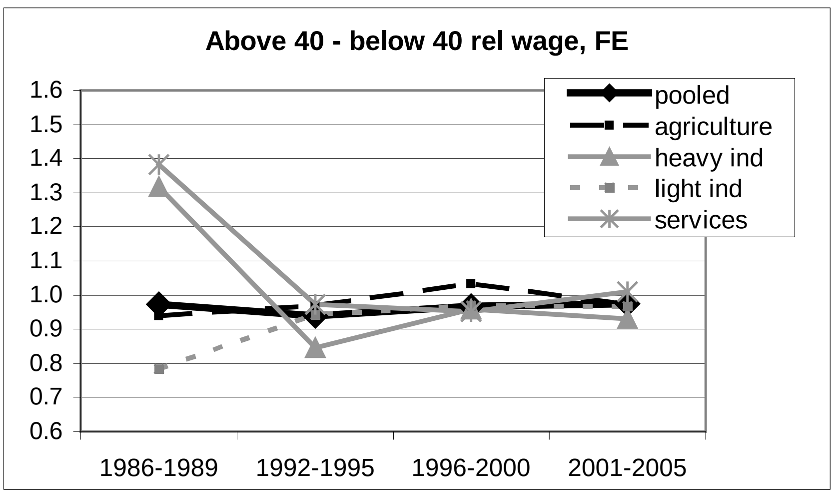






Figures 10.a-10.c: Above 40 – below 40 relative productivities, relative wages, and the gaps - FE, pooled sample, industry subsamples





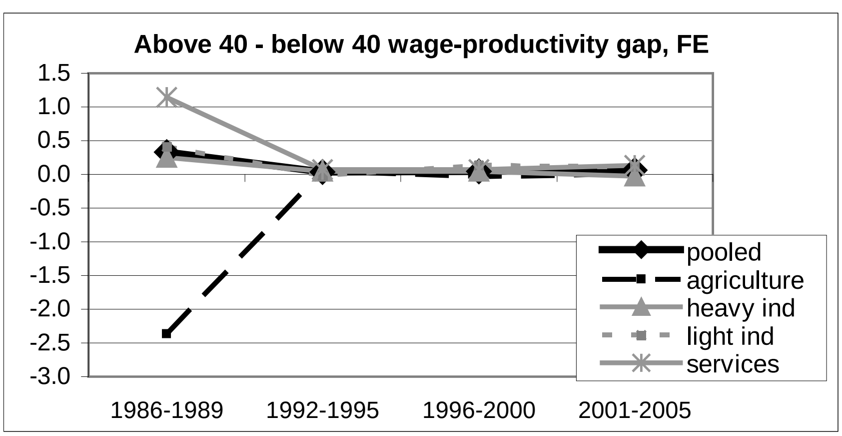


Figure 11: Above 40 – below 40 relative productivities and relative wages - OLS, pooled sample and old firms

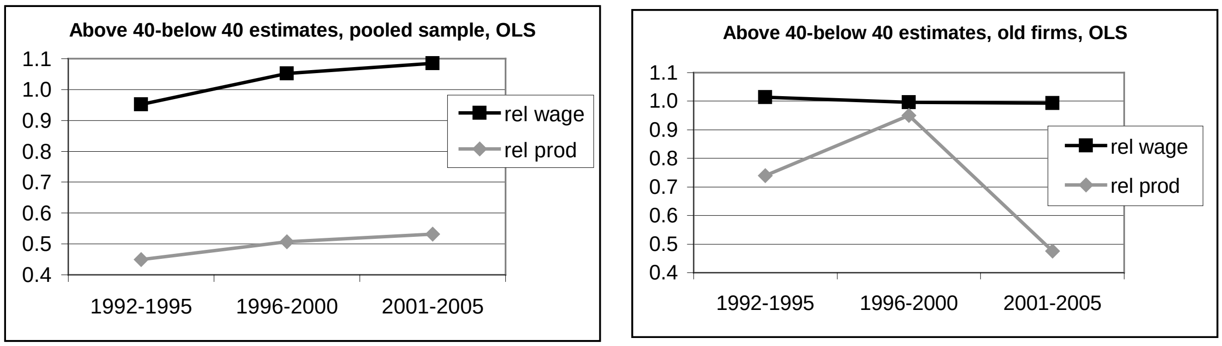


Figure 12: Above 40 – below 40 relative productivities and relative wages - FE, pooled sample and old firms