Changes in wage differentials during transition: testing alternative explanations based on matched employee-employer data

Márton Csillag
marton.csillag@econ.kuleuven.be

Center for Economic Studies, K. U. Leuven

Abstract

In this paper, we compare the relative marginal productivity and wages of different groups of workers during late socialism (1986-1989) and after transition (1997-2000), using matched employer-employee data from the Hungarian manufacturing sector. We find that relative wages were distorted away from relative productivity during late socialism, with women and more educated being underpaied and older workers overpaied. Then, we test whether the large changes in wage differentials during transition came about due to a shift in the relative productivity of different types of labor, or rather to an adjustment of wages towards marinal productivity. We find that changes in the productivity of more-educated and older workers was behind the increase in the education wage premium and the fall in the returns to experience, while an important decrease in gender discrimination during transition lead to an improvement in women's wages. We further investigate whether it was stronger product market competition that lead to this decrease in the gap between women's relative productivity and wages by contrasting industries with different market structures, and find some support for the above hypothesis as there is a positive relationship between domestic market concentration and gender discrimination.

1 Introduction

Transition has brought about major changes in the wage structure in Eastern European countries. I use data from Hungary to look at these changes, which was a frontrunner in the transition to capitalism and has experienced radical changes in employment and relative wages. Hungary also exemplifies three of the major trends that seem common to Central and Eastern European transition countries: decrease in the gender wage gap, increase in the returns to skill, and the obsolence of skills acquired during socialism (decrease in the returns to experience). Though there exists substantial empirical evidence on these changes, there has been little agreement about their causes[1]. The explanations put forward fall into two classes. On one hand, liberalization and more generally introduction of free markets could have decompressed earnings restrictions imposed by the old socialist wage-setting mechanism and adjusted wages to the true differences in marginal productivity. On the other hand, there might have been a change in the relative productivity of different types of labor due to product demand shift, the introduction of new technologies and organizational changes.[2]

The literature on the evolution of the wage structure during transition in Hungary has pointed out that it might be useful to consider two periods when thinking about these changes. Early transition (roughly from 1989 to 1995) corresponds to changes brought on by the collapse of the socialist system, when there were major sectoral (and occupational) reallocation of labor as a result of the collapse of the demand for products of heavy industries. Late transition was characterized by a technological renewal (and changes in work organization), which came about following privatization, the inflow of foreign capital and the strengthening of market forces. Kertesi-Köllő (2001) and Kézdi (2002) concentrate on returns to education and experience to show that the first period brought about a general increase in the returns to skills, and that the second period was associated with an increase in the relative wages of young skilled workers. Their evidence suggests that this second change was a

[1] See Svejnar (1999) and Boeri-Terrell (2002) for overviews.

[2] See Sabirianova (2003) for a discussion of these issues in the context of changes in returns to education during transtion in Russia.

result of the increased productivity of young and educated labor, mainly due to its association with new technology. The evidence on the changes in the gender wage gap is more mixed. It is clear that most of the fall in the gender gap came about shortly after the collapse of state socialism, afterwards it has only decreased slightly. An important part of this decrease came from women holding occupations and working in industries for which the demand has grown. Still, a decrease in "discrimination" meaning the part of the gender gap that cannot be explained by differences in endowments has been equally important (Campos-Jolliffe (2006)).

My work on changes in relative wages and the relative productivity builds on the methodology proposed by Hellerstein et al. (1999). Their approach is based on the observation that in competitive labor markets, the wages of different types of labor should equal their relative marginal products. Thus by simultaneously estimating firm level wage equations and firm level production functions where workers with different demographic characteristics are allowed to have different marginal products, and testing whether relative wages differ from relative productivities, one can assess deviations in firms' wage setting behavior from what would prevail in competitive markets. This method of evaluating wage differentials across different types of workers alleviates the problem of omitted variables bias that plagues the traditional wage-regression based approach.[3] Having access to large matched employer-employee datasets from the Hungarian manufacturing sector from both before transition and after transition times allows me to look at some important issues of wage setting during transition. First, I am able to examine whether and to what extent wages were distorted away from productivity due to equalizing wage policies in late socialism.[4] Second, I am able to disentangle which of the two competing hypotheses – market adjustment or changes in productivity – can better explain changes in relative wages during transition.

---

[3] Omitted variables bias is a particlarly important question in our case, as the data we use contains very little information on workers' background characteristics (as is common in matched employer-employee data). For a more detailed discussion of the advantages of using matched employer-employee data se Hellerstein-Neumark (2006).

[4] In this sense, I am able to improve upon the literature on wage setting in late socialism that concentrated upon whether across-firm wage differentials reflected productivity differences.

My results confirm that wages were not set competitively in late socialist Hungary. Distortion not only took the form of equalization of wages across firms, but also relative wages within firms did not reflect the relative productivity of different worker types, and this was most pronounced for women, who were paid less than their male co-workers despite their higher productivity. Regarding changes in relative wages, I find no uni-causal explanation. Most of the doubling of the returns to education was caused by a large growth in the difference in productivity (and wages) between firms where better-educated and where less-educated work. By contrast, the most important factor behind the roughly 30 percent reduction in returns to experience was the (within-firm) reduction in older workers' relative productivity. Finally, the one-fourth reduction in the gender wage gap during transition came about as a result of two opposing processes. One the one hand, a fall in women's relative productivity (within firms) coupled with women increasingly working at low-productivity (and low-wage) firms lead to a decrease in women's pay; on the other hand a large reduction in discrimination raised women's relative wages.

I further investigate whether the decrease in the gap between women's relative productivity and wages during transition was due to an increase in competition. In essence, I test one implication of Becker's theory of taste discrimination that product market competition will force by discriminating employers out of the market. I do this by contrasting industries with different market structures. The results are largely inconclusive as I find that indeed discrimination against women was more pronounced in more concentrated industries in late socialism, and that this negative correlation between competition and discrimination still exists after transition, albeit to a much lesser extent. Furthermore, there is some evidence that discrimination against women decreased more in industries that were more concentrated before transition.

The chapter is organized as follows. In Section 2, I review the methodology of testing the competitiveness of labor markets based on the simultaneous estimation of production functions and wage equations. In the next section, I shortly describe the data used. The main results on the evolution of relative

productivity and wages of different types of workers are presented in Section 4. The next section is devoted to various checks of the robustness of these results. In Section 5 I describe the empirical strategy of testing for the effect of product market competition on gender discrimination, and present preliminary results.

2 Methodology

The basic idea of the Hellerstein-Neumark method is to specify a structural production function that incorporates the heterogeneity of labor. I will assume that firms produce output with a technology that is of the Cobb-Douglas form with heterogeneous labor input:

$$\ln(Y) = \alpha \ln(K) + \beta \ln(L^*) \tag{1}$$

where L^* is the quality-adjusted labor input. Assuming that there are k different types of labor that are perfect substitutes with (possibly unequal) productivities λ_k, the labor input of a given firm can be written as:

$$L^* = \sum_0^K \lambda_k L_k = \lambda_0 L + \sum_1^K (\lambda_k - \lambda_0) L_k \tag{2}$$

where λ_0 is the productivity of the base category. Then, the labor input can be rewritten as:

$$\ln(L^*) = \ln(\lambda_0) + \ln(L) + \ln\left(1 + \sum_1^K \left(\frac{\lambda_k}{\lambda_0} - 1\right)\frac{L_k}{L}\right) \tag{3}$$

This leads to a production function of the form:

$$\ln(Y) = \alpha \ln(K) + \beta \ln(\lambda_0) + \beta \ln(L) + \beta \ln\left(1 + \sum_1^K \left(\frac{\lambda_k}{\lambda_0} - 1\right)\frac{L_k}{L}\right) \tag{4}$$

In a similar fashion, (logarithm) of firm-level average wages can be written as:

$$\ln(\overline{W}) = \ln(W_0) + \ln\left(1 + \sum_1^K \left(\frac{W_k}{W_0} - 1\right)\frac{L_k}{L}\right) \tag{5}$$

where W_k are the wages of a worker type k. This equation can be interpreted as an aggregation over workers in the firm of individual-level wage equations.

The basic idea of Hellerstein-Neumark is to assume that the relative productivity and wages of different types of labor are constant across firms, and estimate the production function and the firm-level wage equation simultaneously and to test whether the relative wage of a certain type of labor is equal to its relative productivity (which in their original formulation is done on a cross section of manufacturing firms). They interpret this as a test of whether the labor market for a certain type of labor is a competitive one.

Crépon et al. (2003) extended the above methodology to the case where employees' earnings are directly observed. The essential novelty of their approach is to decompose the quality adjusted labor input in the following way:

$$L^* = \sum_0^K \lambda_k L_k = \sum_0^K \frac{\lambda_k}{W_k} W_k L_k \qquad (6)$$

They call λ_k / W_k the "markdown" coefficient, which tells us the productivity to wage ratio in a certain group is higher than in the reference group, with a higher markdown signaling higher "relative exploitation". Then (logarithm of) the quality adjusted labor input can be written as:

$$\ln(L^*) = \ln(\frac{W}{L}) + \ln(L) + \ln\left(\frac{\lambda_0}{W_0}\right) + \ln\left(1 + \sum_1^K \left(\frac{\lambda_k / W_k}{\lambda_0 / W_0} - 1\right)\frac{W_k L_k}{WL}\right) \qquad (7)$$

The advantage of this methodology is that only the ratio of productivity to wages is assumed to be constant across firms. By contrast, it adds a potentially endogenous term to the production function, the average wage at the firm level.[5]

Hellerstein-Neumark use matched employer-employee data, where they have access only to a sample of workers within each firm (similarly to the data

[5] There are several resaons to believe that average wage at the firm level might be endogenous in a production function. These might be (1) to rent-sharing: more productive firms might be sharing the profits they earn with workers; (2) production and wages might be determined simulateously: firms with a higher product market demand might require workers to put in more effort that leads to higher wages and production; (3) omitted variables bias: if labor quality is not captured adequately by the labor quality term, then unobserved ability at the firm level will be part of the error termi n the production function, and unobserved abilty might be correlated with average wages.

used in this paper). Thus, they are only able to get accurate measures of the shares of different types of labor by limiting the dimensionality of the problem using three kinds of restrictions (over and above the basic assumption that different types of labor are perfect substitutes). First, it is assumed that the relative productivity of two types of workers is the same within one demographic group as in another one (say the relative productivity of females to males is the same among more and less educated). Second, similar assumptions about the relative wages are also made. Finally, it is assumed that the proportion of a given type of labor is constant across demographic groups. In our case, these simplifying assumptions imply that the quality adjusted labor will be made up of three terms instead of having to estimate the marginal productivity of eight different types of labor – given that I form worker groups along three dimensions: gender, schooling and experience (see the next section for details on the definition of worker types). Furthermore, by using the fact that $ln(1+x) \approx x$ when x is small, the linearized and simplified labor term will be:

$$\ln(L^*) = \ln(L) + \left(\frac{\lambda_F}{\lambda_0} - 1\right)\frac{L_F}{L} + \left(\frac{\lambda_E}{\lambda_0} - 1\right)\frac{L_E}{L} + \left(\frac{\lambda_O}{\lambda_0} - 1\right)\frac{L_O}{L} \qquad (8)$$

where the indices F, E and O refer to females, more educated and older workers, respectively; in this case, young males with low education are taken as the reference category. This is the form of the quality adjusted labor input that I will use in the chapter.

Finally, the identification of the above models deserves some discussion. Besides the usual problems associated with the estimation of production functions (see for example Griliches-Mairesse (1998)), we have additional issues due to the presence of different labor inputs and – in the Crépon et al. method – the average wage. In order to assess the identification and endogeneity issues involved in the above models, let me focus on the Crépon et. al method and turn to the panel data version of the model, so index i will signify firms, while index t will be used for time periods. Then, the (simplified and linearized) estimating equation will have the following form:

$$\ln(Y_{it}) = \alpha \ln(K_{it}) + \beta \ln(\overline{W}_{it}) + \beta \ln(L_{it}) +$$
$$+ \beta\delta_o + \beta\left(\frac{\delta_F}{\delta_0} - 1\right)P_{it}^F + \beta\left(\frac{\delta_E}{\delta_0} - 1\right)P_{it}^E + \beta\left(\frac{\delta_O}{\delta_0} - 1\right)P_{it}^O + e_{it} \quad (9)$$

where I denote by δ_k the markdown coefficients, and by P^k the share of different types of labor in the wage bill, and e_{it} stands for a generic error term.[6]

Then, one needs to think about the structure of the error term in the above equation. We can categorize the error components based on whether they are thought to vary through time, whether they are known to the firms' managers when making their input decisions (or in other words, whether they are correlated with the right-hand side variables in the model). First, there are differences in managerial ability (γ_i), that might be considered as fixed for the period under observation, since we will be working with short panels. Second, there might be differences in the average quality of the workforce of the firms (that are not captured by proportion of women, better-educated and older workers), due to the unobserved ability of employees (A_{it}). Third, there might be productivity shocks (coming from shocks to demand for the given firm's products), part of which might be observed by the manager (ω_{it}), and we can assume that these evolve according to some first-order stochastic process. Finally, there are pure measurement errors (or productivity shocks unknown to the manager), that might be considered as time-varying and uncorrelated with the right-hand side variables (ε_{it}). So the generic error structure of the model can be written as:

$$e_{it} = \gamma_i + A_{it} + \omega_{it} + \varepsilon_{it} \quad (10)$$

Let me first consider identification in the case where we do not rely on the time dimension of the data (nor on instrumental variables). We know that these estimates will biased unless the unobserved differences across firms are captured by the control variables[7], given that if firm managers have additional information about demand conditions etc. over and above what is observed by

[6] Note also that I have used the assumptions that (1) the markdown coefficients are constant across firms, and (2) that the coefficients are time-homogenous, or in other words, the parameters do not vary across *i* and *t*.

[7] Which in our case will be industry, region, size and ownership type of the firm.

the econometrician and are able to act upon this when choosing inputs, then the right-hand variables in the production function will be correlated with the error term. We have the additional problem, that as we are unable to control for unobserved labor quality differences, and these are probably correlated with our key variables of interest (the proportion of better educated in particular) the coefficients on these will be biased. The exception to this is the scenario where unobserved ability is uncorrelated with other characteristics of the workforce and under the null hypothesis of perfectly competitive markets for labor, and more specifically, for unobserved ability.

Crépon et al. (2003) propose the alternative identifying assumption that conditional on the average wage, the error terms are uncorrelated with the right hand side variables, thus in essence they do not constrain the coefficients on labor and average wages to be equal (I will call this the unconstrained CR model). This assumption might work under two different scenarios. First, if (i) time-constant unobserved productivity differences (those that do not stem from differences in labor quality) are captured by other control variables and that (ii) the quantity and composition of labor, and capital are fixed in the short run (they do not adjust to current productivity shocks). In this case, average wage will serve as a proxy for unobserved labor quality differences. Alternatively, we could work with the assumption that (i) time-constant productivity differences are captured by the control variables, (ii) no differences in the quality of labor input (or they are constant through time and are captured by the control variables). Then average wage is a proxy for the current productivity shock, and so inputs can be assumed to be variable.

Turning to the panel data case, it is obvious that the identification of the between effects model is essentially the same as in the cross-section case. The advantage of the within effects model is that we do not need to rely on control variables to capture time-invariant productivity differences. This means that the fixed effects estimator is consistent under the assumption that (i) firm-level average unobserved ability does not vary over time (no matter whether abiliy is priced competitively) and (ii) there are no time-varying productivity shocks that can be acted upon by managers. It is also consistent when we relax the assumption of no productivity shocks, in the case where (i) inputs are

fixed in the short run, and (ii) no correlation trough time in the productivity shocks, as managers cannot adjust inputs to this period's shock, and the previous period's shock conveys no information on this period's shock. The fixed effects model when introducing the firm-level average wages as a control variable (the unconstrained CR model) is even more flexible, as it allows one to relax either the assumption that unobserved ability does not vary over time, or that other inputs are fixed in the short run. In the first case, wages act as a proxy for ability, and in the second, they are a proxy for productivity shocks.[8]

3 Data, sample, and variable construction

The data come from the Wage Surveys of the Hungarian National Labor Center, which are repeated cross-section representative surveys of workers, and have been carried out in 1986, 1989, and every year since 1992. Although the sampling frame and the sampling procedure have changed to a minor extent, the basic data-collection strategy remained the same throughout the period. This means that in each year (more precisely in the month of May), a (stratified) random sample of firms is selected, and then a random sample of these firms' workers is drawn. Then, basic demographic and job-related information (including wages) on these workers is recorded directly from the employers' records. Furthermore, through a unique firm identification number, the employing firms' balance sheet information can be matched to the sampled workers' data.

The data structure has the consequence that we are unable to follow workers over time. By contrast, panels of firms can be constructed, but these are limited in terms of time-span, as firm identification numbers have changed each time a firm changed its legal status. Thus, in essence two panel of firms can be created: one for the before-transition period, and second one for after-transition, after the period of major privatization and reorganization.

---

[8] In principle, one could explicitly consider the mechanism that might relate productivity shocks to average wages, for example through an effort function: a positive demand shock might lead to workers having to put in extra effort, which is translated into higher average wages through compensating differentials. See Ackerberg et al. (2007) for a summary of the proxy approaches to control for unobserved productivity in the estimation of production functions.

The sample I use is restricted to those firms in the manufacturing sector that employed at least 100 persons, and where the Wage Survey included at least 5% of the workers. As for the variables used, output (value added), capital and aggregate labor (average number of employees), as well as basic firm characteristics come from the balance sheet datasets. I created two measures of compensation from the data. First, the firm's total wage and salary bill as reported in the balance sheet divided by the number of employees. Second, the average wages of sampled workers, where (gross) wages comprise all work-related payments made by the enterprise in May of the given year and 1/12 of the premia, bonuses and rewards paid in the preceding year. Finally, in the basic analysis, I delineate worker groups across three dimensions: gender, education (at least 12 years of school vs. less) and experience (more than 20 years of experience vs. less). The proportion of workers in each category is estimated based on the sampled workers.

4 Basic results

In the basic analysis I estimated four alternative models. First, using the HN method (simultaneous estimation of the production function and the wage equation using SUR), with average firm-level wages estimated using individual workers' earnings data. Second, the HN model, but this time average wages were calculated from the balance sheet data. Third, using basic CR methodology, by constraining the coefficient on labor to be equal to the coefficient of average wages. Fourth, I used the CR method, but I do not impose the above constraint. Naturally, in both of the model using the CR methodology, average wages and different groups' share of the wage bill were estimated from workers' data. Then, I went on to calculate for each model the implied "markdown coefficients" (relative productivity over relative wages) and tested whether they were different from 1. Finally, I tested whether the markdowns estimated using the four models differed; more specifically I compared the two HN models, the first HN model with the first CR model and the two CR models.

Earnings differentials

Before presenting the results of the different models, I first provide some evidence on the evolution of the relative wages of these groups based on cross-section individual-level wage equations of the workers from the selected firms. I present three specifications: first, I only use dummy variables for gender, education and experience; then, I add controls for firm characteristics; finally I add firm fixed effects. The objective of this exercise is not only to look at changes in wage differentials, but also to see how well the data correspond to the model where it is assumed that firm-level wage equations are aggregates of individual wage equations, thus the wage differentials come from within-firm wage differences. If that was not the case, so that an important proportion of the wage differentials across groups came from different types of workers being sorted into different types of firms with differing wage levels, our cross-sectional estimates will be biased as the wage differentials are identified from across-firm variation. For example, if women were sorted into low-paying firms, this will bias our estimates of the gender gap downwards (towards being more negative than what would be if relied on within-firm variation).

I present the results of these wage regressions in Figure 1. We can see that before transition, within-firm wage differentials corresponded well to gross wage differentials (as the estimates with and without controlling for firms do not differ significantly). After transition, an important portion of the gender wage gap (Panel (a)) came from women being sorted into low-paying firms, and that a big part of the wage advantage of more educated workers (Panel (b)) came from sorting into higher paying firms, while sorting does not contribute to wage advantage of older workers. These graphs also tell us that by controlling for firm characteristics, we are able to capture a large portion of the sorting.

What is also important is that there have been large changes in these wage differentials in our sample during transition, and they echo the results from other studies. For the gender gap, there has been an improvement of the relative wages of women by roughly 8 log points, with two-thirds of this happening between 1989 and 1993 and then further change after 1996. The rise in the relative wages of more educated workers has been even larger (roughly 15 log points), and this increase in the education wage differential

already started during late socialism, but it levelled off around 1994. There was also a drop in the wage advantage of older workers, with some of it happening directly after transition, but more importantly after 1997.[9]

Estimated markdowns from cross-section regressions

Turning to the results on the relationship between relative productivity and relative wages of different groups of workers, I first discuss cross-sectional estimates of the markdown coefficients, which I present in Figure 2. One has to remember three points in the interpretation of these results. First, that estimated coefficients presented might be biased due to endogeneity problems. Second, these results only partially take into account the sorting of different workers into different firms.Third, that a markdown coefficient equal to one means that the ratio of relative wages to relative productivity of a given type of labor is equal to that of the base category, and that a coefficient above 1 means that the given group is "underpaid", while a coefficient below 1 signifies being "overpaid".

First let me discuss the results relating to female labor. The most important finding of these regressions is that while in late socialism, women were underpaid (relative to their productivity) with respect to men (with this result being significant at 5% level), after transition their relative wages were largely on par with their relative productivity. This indicates that the improvement in the relative wages of women was probably due to a decrease in discrimination. This basic result is confirmed in all of the specifications that use earnings measures form workers' data, while when using aggregate wage data the markdown coefficient of women did not differ significantly from one even before transition. Turning to differences between the four models, we can note a few important findings. First, a comparison of the two estimates based on the HN methodology shows that the estimated markdown coefficients were

[9] In Panel (d) of Figure 1, I display results from wage regression with more detailed individual characteristics: in these regression I included a dummy variable for females, years of schooling, years of (potential) experience and its square, as well as firm fixed effects. Then, I graphed the (adjusted) gender wage gap, returns to eight additional years of schooling, and the (adjusted) wage differential between workers with 30 and 10 years of experience. These results are in line with the ones based on the more simplified specification, the only exception being that we can see a small further increase in the returns to schooling after 1996.

significantly different from each other during socialism, but they did not differ much after transition. Second, comparing the estimates of the HN model based on individual earnings data with the constrained version of the CR model, one cannot find significant differences. Finally, the estimates from the unconstrained CR model are mostly above the ones form the constrained CR model, but these differences are only significant in the period 1993-1997.

The findings for the markdown of more educated workers are less clear-cut, as the estimates from the alternative models differ markedly. For late socialism, they indicate that while in 1986 more educated were being according to their relative productivity, they were underpaid in 1989. After transition, the markdown coefficients based on the HN methodology indicate that the relative wages of more educated were below their relative productivity. The estimates based on the CR methodology are mostly below the ones from the HN methodology (meaning closer to 1), and do not show underpayment of more educated workers (relative to less educated ones). Furthermore, when we use average wages as a control variable in the CR methodology, the markdown coefficients fall even further. These differences between the different estimates are especially pronounced for the 1993-1996 period, when they are significantly different at the 5% level. All in all, the basic message of these cross-section estimates seems clear: the increase in the relative wages of more educated workers during transition is mostly attributable to an increase in their relative productivity.

Finally, the estimates for older workers – though they differ across specifications somewhat – tell the same basic story: while the ratio of relative wages to relative productivity of older workers was equal to that of younger workers, they were significantly overpaid starting from 1989, and the introduction of free markets did not change this. Thus, these results suggest that a decrease in the relative productivity of older workers (vis-à-vis younger workers) is behind the fall in returns to experience during transition.

Panel data results

Now I discuss the results of estimations where I exploited the fact that firms can be matched across years. I limited this panel data analysis to two

periods: late socialism (1986-1989) and late transition (1997-2000) for several reasons. First, these two periods are useful to contrast a situation before the introduction of market forces to one where all major liberalization steps have been taken. Second, the cross-section estimates showed that some of the changes in wage and productivity differentials were taking place between 1993 and 1997, so by analyzing 1986-1989 and 1997-2000 we can make at a before-after comparisons. Third, the differences between the estimation results for the alternative model specifications based on cross-section data was also the largest in the left-out period (1993-1996). The final reason is largely data-driven, as discussed earlier, it is difficult to follow firms over years during the early transition period.

The major advantage of using the panel dimension of the data is that one is able to disentangle two alternative explanations of the gap between the productivity and wages of different groups. First, it might be that a certain type of labor is over-represented in firms where the relationship between productivity and wages differ from the market equilibrium for all workers. Thus sorting might give an explanation to potential "underpayment", and this mechanism will be reflected in the between effects model specification. Second, we can look at whether average wages and productivity move on par in a firm when it changes to proportion of a certain type of labor in its pool of employees. So by looking at the within effects estimates one can more directly test for "wage discrimination". [10]

[10] It has to be noted that there are two problems associated with the panel data analysis. First, the assumption that coefficients are time-homogenous within the two periods (1986-89 and 1997-2000); as we saw in using the cross-sectional data, some of the changes were already under way in 1989. Second there is the traditional concern about the magnification of the effect of measurement errors.

Table 1: Panel data estimates, between effects model, 1986-1989 (n=1021)

	Female		*Educated*		*Experienced*		*All*
	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value	p-value
HN ind. wages	1.574	0.02	1.421	0.01	0.523	0.01	0.00
HN total wages	1.086	0.53	1.333	0.01	0.726	0.09	0.01
CR constrained	1.455	0.01	1.459	0.07	0.576	0.02	0.00
CR unconstrained	1.695	0.00	1.211	0.41	0.441	0.00	0.00
HN1 vs. HN2		0.00		0.89		0.02	0.00
HN1 vs. CR1		0.61		0.67		0.78	0.92
CR1 vs. CR2		0.00		0.00		0.01	0.01

Note: The p-values in the top panel are from two-sided tests of whether the estimated markdown coefficient is equal to 1. The p-values in the bottom panel test the equality of estimated markdown coefficients from different models.

The results for the between effects model for the late socialist period (Table 1) suggest that women were working in firms that were underpaying their workers. More precisely, firms with a high share of women among their workers were no less productive than male-dominated firms, but they were paying significantly lower wages.[11] Similarly, more educated workers were also working in firms that under-rewarded their employees, as these firms, though they were paying higher wages than those with mostly low-educated workers, this did not match their productivity advantage. Finally, there was an important gap between the productivity and the wages of firms with a high proportion of older workers, as these firms were significantly less productive but were paying roughly the same as firms with predominantly younger employees.

The results for the fixed-effects model (Table 2) are similar to the ones for the between-effects model with respect to gender and experience (though the precision of the estimates is much lower). This means that an increase in the proportion of female workers (or the wage bill paid to women) in a given firm was associated to an increase in productivity, while the changes in wages did not match this productivity gain. By contrast, increasing the proportion of older workers leads to higher wages, without any productivity gain. Finally, though the productivity improvement stemming from an increase in the

[11] See Appendix Table 2, for the productivity and wage profiles.

proportion of highly educated workers was not as high as the associated increase in wages, we cannot reject the hypothesis that the within-firm productivity differential between less and more educated workers was equal to the within-firm wage differential.[12] Two additional remarkable results appear from the comparison of the different models: first, the results based on aggregate wages differ significantly from the ones based on individual wages; second, that we do not find strong evidence of endogeneity (as the two models based on the Crépon et al. methodology do not differ significantly from each other).

Table 2: Panel data estimates, fixed effects model, 1986-1989 (n=1711)

	Female		*Educated*		*Experienced*		*All*
	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value	p-value
HN ind. wages	3.135	0.00	1.050	0.90	0.337	0.06	0.01
HN total wages	1.364	0.01	1.014	0.92	0.744	0.03	0.01
CR constrained	2.617	0.03	0.996	0.99	0.303	0.19	0.09
CR unconstrained	3.148	0.07	0.564	0.54	-0.026	0.22	0.30
HN1 vs. HN2		0.00		0.14		0.99	0.00
HN1 vs. CR1		0.70		0.96		0.86	0.98
CR1 vs. CR2		0.43		0.35		0.44	0.81

Note: The p-values in the top panel are from two-sided tests of whether the estimated markdown coefficient is equal to 1. The p-values in the bottom panel test the equality of estimated markdown coefficients from different models.

I now turn to the panel data estimates from the post-transition period (1997-2000). These show that there are several important differences with respect to late socialism both in terms of the selection of different types of workers to different firms, and the relative productivity and wages of different types of labor.

First I will discuss across-firm variation in productivity and wages (the between-effects model, Table 3).[13] Though the different estimated models give results that differ quantitatively, they all lead to the same qualitative conclusions. First, there are no differences between men and women with respect to the relationship between the relative productivity and relative wages of firms that employ them. More specifically, a higher percentage of women

[12] The estimated wage and productivity profiles can be found in Appendix Table 3.

[13] See Appendix Table 4 for the estimates of relative productivities and relative wages.

among the work force is associated with both lower productivity and lower wages, and this association is roughly of the same magnitude for both variables. Second, more educated workers tend to be employed in firms that "exploit" their workers, meaning that productivity in firms with a high proportion of well-educated workers tends to be much higher than in ones that use predominantly low-educated labor, and though they pay higher wages too, this does not fully match the productivity advantage. Finally, older workers are clustered at firms that "overpay" their workers, in these firms productivity is significantly lower, but wages are roughly the same as in firms with high proportion younger workers.

Table 3: Panel data estimates, between effects model, 1997-2000 (n=1356)

	Female		*Educated*		*Experienced*		*All*
	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value	p-value
HN ind. wages	0.806	0.23	1.423	0.00	0.368	0.00	0.00
HN total wages	0.780	0.09	1.389	0.00	0.448	0.00	0.00
CR constrained	0.903	0.33	1.378	0.00	0.246	0.00	0.00
CR unconstrained	0.971	0.79	1.259	0.07	0.162	0.00	0.00
HN1 vs. HN2		0.48		0.34		0.01	0.05
HN1 vs. CR1		0.93		0.05		0.28	0.19
CR1 vs. CR2		0.04		0.04		0.04	0.20

Note: The p-values in the top panel are from two-sided tests of whether the estimated markdown coefficient is equal to 1. The p-values in the bottom panel test the equality of estimated markdown coefficients from different models.

Looking at within-firm variation gives somewhat different results (Table 4)[14]. With respect to gender, there is some evidence of the "underpayment" of women within firms (as their productivity relative to men is roughly the same, but their wages tend to be lower), but this result is not statistically significant. Second, all of the models based on individual-level wage data show that better-educated workers are over-paid, as a increase in the proportion of workers who obtained at least a baccalaureate leads to an increase in wages but not in productivity. Third, the within-effects model reveals that older workers are less productive, but equally paid, leading to younger workers being significantly underpaid.

[14] The productivity and wage profiles of different worker types can be found in Appendix Table 5.

Finally, it has to be noted that the model HN model (using individual wages) and the unconstrained CR model lead to differences in the "markdown" parameter associated with education that is statistically significantly different from the constrained CR model. More specifically, the ratio of productivity and wages of more educated workers (relative to that of less educated workers) increases when one controls for the unobserved heterogeneity of workers in some way. This seems to imply (under the standard assumption that unobserved ability is positively correlated with education), that the productive advantage that comes form employing workers with higher unobserved ability is not fully compensated in terms of wages, or in other words, unobserved skills are not priced competitively.

Table 4: Panel data estimates, fixed effects model, 1997-2000 (n=3367)

	Female		*Educated*		*Experienced*		*All*
	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value	p-value
HN ind. wages	1.277	0.03	0.709	0.00	0.652	0.00	0.00
HN total wages	1.019	0.83	0.907	0.20	0.746	0.00	0.00
CR constrained	1.094	0.60	0.327	0.00	0.306	0.00	0.00
CR unconstrained	1.034	0.80	0.607	0.00	0.474	0.00	0.00
HN1 vs. HN2		0.00		0.00		0.01	0.00
HN1 vs. CR1		0.93		0.05		0.28	0.19
CR1 vs. CR2		0.24		0.00		0.00	0.01

Note: The p-values in the top panel are from two-sided tests of whether the estimated markdown coefficient is equal to 1. The p-values in the bottom panel test the equality of estimated markdown coefficients from different models.

Let me now summarize the results from the estimation of production functions with heterogeneous labor. These indeed show that relative wages did not reflect relative productivity during late socialism, with women and more educated being underpaied while older workers were overpaied. As for changes in wage differentials these were caused by a combination of factors.[15] Most of the doubling of the returns to education was caused by a large growth in the difference in productivity (and wages) between firms where better-educated and where less-educated work. By contrast, the most important factor behind the

[15] See Appendix Table 1 for the results of testing whether markdown coefficients differed across the before and the after transition period. There I show that the only the estimates for female labor differed across the two period, and this is true for both the between and the within models.

roughly 30 percent reduction in returns to experience was the (within-firm) reduction in older workers' relative productivity. Finally, the one-fourth reduction in the gender wage gap during transition came about as a result of two opposing processes. One the one hand, a fall in women's relative productivity (within firms) coupled with women increasingly working at low-productivity (and low-wage) firms lead to a decrease in women's pay. On the other hand there was an important reduction in discrimination that raised women's relative wages.

5 Robustness checks

Up to this point, I have been concerned with the estimation of the relative productivity and wages of different types of labor in the basic specification. Now I will consider relaxing various assumptions and the robustness of the estimates with respect to the most important econometric problems.

The specification of the labor input

As discussed in Section 2, the way labor input is specified in the basic model is restrictive in several ways. First, there was the assumption about the relative marginal productivities and wages (markdowns) and proportions of different types of workers; so for example, I assumed that the relative number of workers, the relative productivity, and the relative wages by gender are constant in each experience-schooling category. Second, I assumed that different types of labor have different marginal products, but are perfectly substitutable. Here I consider relaxing these assumptions.

To assess the robustness of the estimates of the basic (restricted) model, one would have to consider estimating production functions where one allows for potentially eight different types of labor. This exercise can only be done for the largest firms in my sample, as one needs to have a sufficiently large sample of workers in order to be able to find at least one worker of each type. Therefore, I looked at only firms with at least 500 employees, and I proceeded in three steps: first, I estimated the basic model for these firms, and I compared this to the esimates for medium-size (those between 100 and 500 employees). Second, I dropped all those firms that had no workers in either

category and reestimated the basic model; third I estimated the full (unrestricted) model on the above sample of firms. I carried out the estimation for the 1986-1989 and 1997-2000 panels, and in the discussion, I will concentrate on the estimates of the fixed-effects model.

The first step of this robustness check (disaggregating the sample into medium and large firms) already showed one important problem: the sample sizes for large firms were too small to permit statistical inference.[16] For the before-transition period, the markdown estimates for medium-size firms were in line with the results for the whole sample, and the estimates for the large firms were undistuingishable from these due to large standard errors. In fact, the only phenomenon that was statistically significant at the 10% level was the overpayment of older workers. The pattern of results for the after-transition period were much the same: the markdown coefficients for medium-size firms corresponded to the ones obtained for the whole sample, and the results for the large firms were plagued by large standard errors.

Second, I experimented with a specification where workers who have at least a baccalaureate and those without are imperfect substitutes (which seems sensible given that these educational categories correspond well to working in blue collar versus white collar jobs). This was done by introducing two labor aggregates, one for each education category, in the Cobb-Douglas production function. The assumption that the fraction, the relative productivity and wage of workers with different gender and experience is constant across education categories is maintained. I carried out the estimation for the 1986-1989 and 1997-2000 panels. The main qualitative results remained unchanged. During late socialism women were both sorted into firms that paid workers less than their productivity and were underpaid within firms, while after transition neither of these was true: women and men worked for firms where the relation between productivity and wage was similar, and their relative productivity and wages were on par within firms. As for experience, older workers were both sorted into "high paying" firms, and their relative

[16] For the 1986-1989 period, I had 776 observations for 483 firms in the large-firm sample and 935 observations for 621 firms in the large-firm sample. In the 1997-2000 panel there were 2688 observations of 1163 firms in the medium-size firm sample, and the large-firm sample was comprised of 677 observations for 267 firms.

productivity did not match their relative wages during late socialism; and these tendencies became even more pronounced after transition.

Endogeneity and sample selection problems

It is well known there are endogeneity problems inherent in the estimation of production function if firms have additional information about demand conditions etc. over and above what is observed by the econometrician and are able to act upon this when choosing inputs. The estimation of a Crépon et al. type augmented production function comes with the additional problem that if markets are not perfectly competitive, wages might be endogenous. The second problem that is more specific to panel data analysis stem from the fact that in general there is entry and exit of firms into the sample. Now if these decisions are based on unobservables, and input choices are also correlated with these factors, then we face the problem of non-random selection that will also bias the coefficients on inputs. In this section I first look at the robustness of the results with respect to endogeneity due to unobserved productivity shocks, then go on to test for selection bias. I will consider both of the above econometric problems based on the unbalanced panel from 1997-2000, as the data from the late socialist period is not rich enough (more specifically it only contains two years, and the number of outside instruments is limited).

In order to look at the sensitivity of the results with respect to endogenetiy problems I used a Blundell-Bond system GMM estimator (see Blundell-Bond (2000)), that uses suitably lagged levels of the variables as instruments for the differenced equations, and suitably differenced first differences of the variables as instruments for the equations in levels. I carried out the estimation for the unrestricted version of the Crépon et al. procedure, considering the version where all of the original restrictions regarding proportions, relative productivity and wages (as well as substitutability) were assumed. In the specification that was not rejected by the data (using Hansen overidentification tests), average wages were treated as endogenous, and all other variables were treated as predetermined. I also experimented with using the unemployment rate in the region where the firm operates (more specifically, its natural logarithm), and the market share of the firm in its 3 digit industry as outside instruments,

where I considered the first one as predetermined and the second one as endogenous. The results of this estimation procedure gave the same qualitative results as the fixed effects (or first-differenced) estimation. This means that within firms, the relative productivity and relative wages of women are on par, that more educated workers are overpaid (although the significance of this result is sensitive to the choice of instruments), and that older workers are significantly overpaid (relative to their younger counterparts).

To address the problem of selection bias, or more specifically, to test for nonrandom selection, I follow Verbeek-Nijman (1992) and Semykina-Wooldridge (2005). These authors show that the fixed effects and the fixed-effects 2SLS estimators are consistent if selection is strictly exogenous, or in other words, if selection can be correlated with time-invariant unobservables (but not with time-varying unobservables). Verbeek-Nijman propose a Hausman-type test that compares the fixed effects estimators from the balanced and unbalanced panels, as well as some variable addition tests, where they proposed to test whether some time-varying function of the selection indicators are significant when added to the equation of interest (Semykina-Wooldridge also use similar simple variable-addition tests).

I have carried out both types of tests for both the fixed effects estimator as well as the Blundell-Bond estimator, where besides comparing the balanced and unbalanced panels, I experimented with adding the number of times a given firm was observed after the given year to the main equation (in the unbalanced panel). The Hausman-type tests did not indicate significant differences between parameter estimates from the balanced and the unbalanced models for neither the fixed effects estimator not the Blundell-Bond estimator.[17] As for the number of times the given firm was in the sample

[17] It is also worth looking at the direction of the change in coefficients between the balanced and the unbalanced panels in order to get an idea of the possible effect of selection. When using the unbalanced panel, the coefficient on capital and average wages increases, while the coefficient on labor decreases. This indicates that selection leads to a negative correlation between the productivity shocks and capital, which means that firms with larger capital stocks are able to withstand larger negative shocks. The same is true for average wages, which seems sensible if one presumes that firms with higher wages have workers with high unobserved ability (or in these firms workers exert more effort). It is less clear why selection might induce a positive correlation between productivity shocks and labor.

after the given year, this variable had a positive coefficient for both the fixed effects and the Blundell-Bond estimators, and while it was not significant in the first case, it was significant at the 5% level in the second case. This indicates that a positive shock to production will lead to a given firm surviving for longer.

6 The effect of competition

The most marked and robust result of the estimation of production functions with heterogenous labor has been that while in late socialism women's relative wages did not match their relative productivity, we find no sign of such discrimination for post-transition. In this section, we take a first attempt at testing the hypothesis (based on Becker (1957)) that increased competition in the product market would reduce discrimination against women. According to Becker's model, as discriminating employers forego profits to indulge in their distaste for employing women, this will only be possible if they earn rents, and an increase in product market competition reduces rents and forces discriminating employers out of the market.

Looking at Hungary (and other transition countries) to provide evidence on Becker's claim seems fruitful in several respects. First, resulting of the rapid introduction of liberalization measures, the level of competition increased significantly within a few years as firms faced competition from domestic entrants, foreign companies, and in the form of imports.[18] In the meantime, there was also an important drop in the gender wage gap (it reduced by one-fourth). Thus we might hope to be able to pick up the effects of competition on discrimination, if there is any. Second, I argue that it is possible to circumvent the usual econometric problems resulting from the simulateneity between market shares and output, given that transition can be interpreted as a

[18] Halpern-Körösi (2001) show that during transtion in Hungary import competition had a negative effect on firms' market share while the industry-level concentration had a positive effect on this variable. Moreover, their result that technically more efficient firms gained market share implies that in face of increasing competition firms could only keep their market share if they upgraded the organizion of their production, which might in turn imply a reduction in taste discrimination.

"natural experiment".[19] In this sense, one might provide "cleaner" results than studies using data from full-fledged market economies.

Table 5: Estimated markdown coefficients for women by industry concentration, 1986-1989 and 1997-2000 panels (CR unconstrained methodology)

	Competition in current period				**Competition in 1986**			
	1986-1989		*1997-2000*		*1986-1989*		*1997-2000*	
	Coeff.	p-val.	Coeff.	p-val.	Coeff.	p-val.	Coeff.	p-val.
Four-firm share=0	2.033	0.254	0.851	0.446	1.872	0.345	1.160	0.551
Four-firm share=100	5.402	0.015	1.376	0.249	5.519	0.010	0.850	0.661
Test of equality of markdown		0.161		0.271		0.123		0.571

Note: The p-values in the top panel are from two-sided tests of whether the estimated markdown coefficient is equal to 1. The p-values in the bottom panel test the equality of estimated markdown coefficients in high-concentration versus low-concentration industries.

Studying the effect of competition on discrimination against women, I examined two kinds of questions. First, a static one: can we find a smaller disparity between the relative productivity and wages of women in firms that operate in more competitive industries? To answer this question, I estimated with fixed effects the unrestricted version of the Crépon et al. equation seperately for the 1986-1989 and the 1997-2000 panels. I used the share of the four largest firms in the 2 digit industry output as a measure of product market competition, and interacted this variable with the shares of women in the wage bill, and then tested whether the gap between women's productivity and wages was lower in more-competitive than in less-competitive industries. The results, presented in Table 5, reveal that internal product market competition does lead to a decrease in the discrimination against women, and that the association between how concentrated an industry is and the gap between the wages and productivity of women in firms operating in this industry was particularly pronounced in late socialism. On the other hand,

[19] The argument for this interpretation is that the conditions which detemined the level of competition prior to 1990 were completely different from those after the introduction of liberalization measures; and that the transition shock was to a large part unexpected. See Brown-Earle (2000) who use this argument.

these results are not significant, which is probably due to the fact that a rather crude measure of competition.[20]

Second, I look at the dynamic implications of Becker's insights: did discrimination against women decrease more in industries that experienced more increase in competition than in industries where competition did not become more intense? In this case, I esimated the change in the markdown coefficients of women (using the 1986-1989 and 1997-2000 panel, and adding firm fixed effects), and contrasted industries that were relatively competitive already in 1986 with concentrated ones. Thus, in essence, I used a difference-in differences strategy to identify the effect of changes in product market competition on gender wage discrimination. I find only very weak support for Becker's hypothesis using the dynamic specifications: the discrepancy between women's relative productivity and wages decreased more in concentrated industries than ones that were already competitive towards the end of socialism, but this result is again not statistically significant.

7 Conclusion

The basic objective of this paper was to assess how competitive was the labor market in late socialism and during transition in Hungary. This was done by using matched employer-employee data from the manufacturing sector from 1986 to 2000 to simultaneously estimate production functions where we control for the composition of labor in firms and different types of workers have different marginal products along with firm-level earnings functions, and to test whether workers' relative productivity corresponded to their relative wages. This exercise is also important as we can look at what was behind the marked changes in wage differentials during transition: did the the relative productivity of different types of labor change, or rather did the introduction of markets bring about an adjustment of relative wages to relative productivity?

---

[20] When using the four-firm concentration ratio measured at the 3 digit industry level, the difference between concentrated and non-concentrated industries is actually significant int the before-transition panel. Two-didit industries needed to be used because the industry coding of before 1990 could only be harmized at this level with NACE coding used from 1997.

We found that in late socialism, relative wages were not on par with the relative productivity of workers: women and more educated workers were underpayed, while older workers were overpaied. After transition we can still reject that firms determine wages as if the labor market was competitive, this is manifested in the fact that older workers are paied higher wages than their younger counterparts, even though they are equally productive. This means that the decrease in the returns to experience during transition was because the organization of work has changed in a way that led to a appreciation of younger workers productivity. Surprisingly, we do not find the same result for the productivity of more educated workers, as the large increase in the returns to schooling seems to stem from more educated workers being sorted to more productive (and high pay) firms, but increasing the proportion of better educated workers in a given firm does not lead to higher productivity. By contrast, we find that the introduction of market forces lead to an adjustment of women's wages to their relative productivity. More precisely, women are still paied less than men, but this is because the firms they are working for less are productive and women are less productive than men within firms, but discriminatory wage setting towards women has disappeared. We examined whether this latter phenomenon was due to higher competition in the product markets disciplining employers to behave in non-discriminatory way, but only found weak evidence of this effect.

A more refined analysis of the link between product market competition is left for future work. There are two paths that seem fruitful to explore. First, to use variation in competitive pressure in different industries that arise from the imports of foreign goods. Second, to allow for a differential impact of product market, by looking at whether firms in industries that were more seriously financially constrained at the beginning of transition and therefore could allow for less inefficient behavior in the form of discrimination were more sensitive in their wage setting to changes in competition.

References

Abowd, J. and Kramarz, F. (1998): Econometric Analyses of Linked Employer-Employee Data. *Labour Economics*, Vol.6, 53-74.

Ackerberg, D., Benkard L., Berry, S. and Pakes, A. (2007): Econometric Tools for Analyzing Market Outcomes. In.: Heckman, J. (ed.) *Handbook of Econometrics*, Vol. 6, Elsevier.

Basu S., Estrin S. and Svejnar J. (2004): Wage Determination under Communism and in Transition: Evidence from Central Europe, *IZA Discussion Paper* No. 1276

Bayard, K., J. Hellerstein, D. Neumark, and K. Troske (2003) : New evidence on sex segregation and sex differences in wages from matched employee–employer data. *Journal of Labor Economics*, Vol. 21, 887-922.

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

Black, S. and Brainerd, E. (2006): Importing Equality? The Effects of Increased Competition on the Gender Wage Gap, *Industrial and Labor Relations Review*, Vol. 57, 540-589.

Blundell, R. and Bond, S. (2000): GMM estimation with persistent panel data: an application to production functions., *Econometric Reviews*, Vol. 19, 321-340.

Boeri, T. and Terrell, K. (2002): Institutional Determinants of Labor Reallocation in Transition. *Journal of Economic Perspectives*, Vol. 16, 51-76.

Brown D. and Earle J. (2000): Competition and Firm Performance: Lessons from Russia. *CEPR Discussion Paper*, No. 2444.

Brown D., Earle J. and Telegdy Á. (2006): The Productivity Effects of Privatization: Longitudinal Estimates from Hungary, Romania, Russia, and Ukraine. *Journal of Political Economy*, Vol. 114, 61-99.

Campos N. and Jolliffe D. (2005): Does Market Liberalisation Reduce Gender Discrimination? Econometric Evidence from Hungary, 1986-1998, *Labour Economics*, Vol. 12, 1-22.

Crépon B., Deniau N. and Pérez-Duarte S. (2003): Wages, Productivity, and Worker Characteristics: A French Perspective. *CREST Working Paper*, No. 2003-04.

Frazer Garth (2001): Linking Firms and Workers: Heterogeneous Labor and Returns to Education. mimeo, Yale University.

Griliches, Z. and J. Mairesse (1998): Production Functions: The Search for Identifcation.. In: *Econometrics and Economic Theory in the Twentieth Century: The Ragnar Frisch Centennial Symposium*, ed. S. Strøm. Cambridge, UK: Cambridge University Press.

Halpern, L. and G. Körösi (2001): Efficiency and Market Share in Hungarian Corporate Sector; *Economics of Transition*, Vol. 9, No. 3.

Hellerstein, J., and D. Neumark (2006): Using Matched Employer-Employee Data to Study Labor Market Discrimination, in *Handbook on the Economics of Discrimination*, William Rodgers, (Ed)., 29-60.

Hellerstein J., Neumark D., and K.. Troske (1999): Wages, Productivity, and Worker Characteristics: Evidence from Plant-Level Production Functions and Wage Equations. *Journal of Labor Economics*, Vol. 17, 409-46.

Kertesi G. and Köllő J. (2003): Economic transformation and revaluation of human capital: Hungary 1986-1999. In: de Grip A., van Loo and Mayhew K. (eds.) *The Economics of Skills Obsolence*, Elsevier Science, Oxford, 235-273.

Kézdi, Gábor (2002): Two Phases of Labor Market Transition in Hungary: Inter-Sectoral Reallocation and Skill-Biased Technological Change. *Budapest Working Papers on the Labour Market*, No. 2002/ 3.

Köllö J. (ed.) (2002): Wages: a decade of transformation. In: *The Hungarian Labour Market 2002*. MTA KTK, Budapest.

Münich, D., Svejnar J. and K. Terrell (2005): Returns to Human Capital Under the Communist Wage Grid and During the Transition to a Market Economy. *Review of Economics and Statistics*, Vol. 87, 100-123.

Sabirianova Peter K. (2003): Skill-Biased Transition: The Role of Markets, Institutions, and Technological Change. *IZA Discussion Paper*, No. 893.

Semykina A. and Wooldridge J. (2005): Estimating Panel Data Models in the Presence of Endogeneity and Selection: Theory and Application. Mimeo, Michigan State University.

Svejnar, Jan (1999): Labor Markets in the Transitional Central and Eastern European Economies, in *Handbook for Labor Economics*, Volumes 3. O. Ashenfelter and D. Card, eds., Elsevier Science, North-Holland, pp. 2809– 857.

Verbeek, M. and T. Nijman, (1992): Testing for selectivity bias in panel data models, *International Economic Review*, Vol. 33, 681-703.

Appendix

Appendix Table 1: Tests of the equality of markdown coefficients between the before and after transition period (p-values)

Between effects	*Female*	*Educated*	*Experienced*	*All*
HN ind. wages	0.007	0.999	0.573	0.051
HN total wages	0.071	0.998	0.478	0.272
CR constrained	0.009	0.666	0.137	0.021
CR unconstrained	0.002	0.873	0.265	0.008
Fixed effects	*Female*	*Educated*	*Experienced*	*All*
HN ind. wages	0.018	0.132	0.492	0.061
HN total wages	0.014	0.100	0.502	0.044
CR constrained	0.000	0.079	0.994	0.001
CR unconstrained	0.000	0.929	0.260	0.003

Appendix Table 2: Panel data estimates, wage and productivity profiles, 1986-1989

	BE				**FE**			
	Productivity		*Wages*		*Productivity*		*Wages*	
	Coeff.	St. Err.	Coeff.	St. Err.	Coeff.	St. Err.	Coeff.	St. Err.
Individual wages								
Female	0.104	0.104	-0.243	0.030	0.219	0.073	-0.251	0.025
Educated	0.510	0.128	0.364	0.037	0.056	0.084	0.277	0.029
Experienced	-0.230	0.113	0.100	0.033	-0.097	0.063	0.190	0.021
Ln(capital)	0.311	0.022			0.185	0.033		
Ln(labor)	0.543	0.034			0.164	0.030		
Aggregate wages								
Female	0.053	0.104	-0.009	0.065	0.243	0.073	0.141	0.057
Educated	0.558	0.128	0.353	0.080	0.173	0.084	0.378	0.065
Experienced	-0.228	0.113	-0.075	0.071	-0.050	0.063	0.189	0.049
Ln(capital)	0.204	0.020			0.135	0.032		
Ln(labor)	0.693	0.032			0.436	0.030		

Appendix Table 3: Panel data estimates of Crépon et. al model, 1986-1989

	BE				**FE**			
	Constrained		*Unconstrained*		*Constrained*		*Unconstrained*	
	Coeff.	St. Err.	Coeff.	St. Err.	Coeff.	St. Err.	Coeff.	St. Err.
Female	0.263	0.106	0.379	0.109	0.311	0.070	0.351	0.072
Educated	0.265	0.114	0.115	0.119	-0.001	0.071	-0.071	0.076
Experienced	-0.245	0.114	-0.305	0.114	-0.134	0.062	-0.168	0.063
Ln(capital)	0.307	0.023	0.310	0.023	0.181	0.033	0.178	0.033
Ln(labor)	0.578	0.035	0.546	0.035	0.192	0.028	0.163	0.030
Ln(avg. wage)			1.011	0.109			0.363	0.073

Appendix Table 4: Panel data estimates, wage and productivity profiles, 1997-2000

	BE				**FE**			
	Productivity		*Wages*		*Productivity*		*Wages*	
	Coeff.	St. Err.	Coeff.	St. Err.	Coeff.	St. Err.	Coeff.	St. Err.
Individual wages								
Female	-0.373	0.082	-0.378	0.044	0.024	0.076	-0.193	0.034
Educated	1.233	0.097	0.859	0.052	-0.040	0.064	0.330	0.029
Experienced	-0.447	0.086	0.096	0.047	-0.238	0.057	0.025	0.026
Ln(capital)	0.180	0.014			0.110	0.016		
Ln(labor)	0.749	0.038			0.711	0.030		
Aggregate wages								
Female	-0.416	0.082	-0.357	0.044	0.031	0.076	0.020	0.030
Educated	1.282	0.096	0.828	0.052	-0.029	0.064	0.062	0.025
Experienced	-0.452	0.086	0.021	0.047	-0.230	0.057	-0.050	0.023
Ln(capital)	0.140	0.013			0.079	0.015		
Ln(labor)	0.833	0.034			0.789	0.029		

Appendix Table 5: Panel data estimates of Crépon et. al model, 1997-2000

	BE				**FE**			
	Constrained		*Unconstrained*		*Constrained*		*Unconstrained*	
	Coeff.	St. Err.	Coeff.	St. Err.	Coeff.	St. Err.	Coeff.	St. Err.
Female	-0.079	0.077	-0.021	0.079	0.054	0.071	0.024	0.070
Educated	0.308	0.081	0.192	0.089	-0.385	0.058	-0.276	0.060
Experienced	-0.614	0.076	-0.622	0.076	-0.398	0.058	-0.370	0.057
Ln(capital)	0.173	0.015	0.175	0.015	0.124	0.016	0.106	0.016
Ln(labor)	0.814	0.031	0.742	0.038	0.573	0.023	0.702	0.030
Ln(avg. wage)			0.940	0.050			0.364	0.038

Appendix Table 6: Coefficient estimates, unconstrained CR model, labor with different levels of education as imperfect subsitutes

	1986-1989				**1997-2000**			
	BE		*FE*		*BE*		*FE*	
	Coeff.	St. Err.	Coeff.	St. Err.	Coeff.	St. Err.	Coeff.	St. Err.
Female	0.433	0.113	0.348	0.116	-0.022	0.083	-0.148	0.117
Experienced	-0.368	0.116	-0.162	0.103	-0.575	0.080	-0.258	0.103
Ln(capital)	0.327	0.024	0.178	0.052	0.188	0.016	0.244	0.025
Ln(educated)	0.201	0.025	0.063	0.025	0.324	0.023	0.190	0.029
Ln(noneducated)	0.263	0.068	0.047	0.054	0.263	0.029	0.391	0.034
Ln(a. wage ed.)	0.287	0.033	0.095	0.037	0.171	0.042	0.055	0.043
Ln(a. wage none.)	0.707	0.106	0.313	0.118	0.854	0.062	0.227	0.070

Appendix Table 7: Coefficient estimates, unconstrained CR model, medium firms

	1986-1989				**1997-2000**			
	BE		*FE*		*BE*		*FE*	
	Coeff.	St. Err.	Coeff.	St. Err.	Coeff.	St. Err.	Coeff.	St. Err.
Female	0.420	0.117	0.380	0.063	-0.030	0.082	0.050	0.072
Educated	-0.051	0.123	-0.030	0.070	0.176	0.091	-0.249	0.062
Experienced	-0.373	0.116	-0.205	0.056	-0.615	0.078	-0.375	0.058
Ln(capital)	0.315	0.025	-0.015	0.035	0.183	0.016	0.108	0.017
Ln(labor)	0.437	0.060	0.105	0.034	0.674	0.046	0.685	0.039
Ln(avg. wage)	0.894	0.119	0.445	0.066	0.897	0.052	0.367	0.040

Appendix Table 8: Coefficient estimates, unconstrained CR model, large firms

	1986-1989				**1997-2000**			
	BE		*FE*		*BE*		*FE*	
	Coeff.	St. Err.	Coeff.	St. Err.	Coeff.	St. Err.	Coeff.	St. Err.
Female	0.309	0.233	0.161	0.185	0.057	0.274	0.175	0.272
Educated	0.361	0.267	-0.328	0.183	-0.160	0.350	-0.420	0.212
Experienced	-0.078	0.291	-0.401	0.166	-0.937	0.310	-0.281	0.220
Ln(capital)	0.327	0.050	0.321	0.053	0.208	0.045	0.092	0.045
Ln(labor)	0.625	0.074	0.100	0.069	0.759	0.099	0.734	0.075
Ln(avg. wage)	0.952	0.236	0.618	0.165	1.468	0.167	0.376	0.130

Appendix Table 9: Coefficient estimates, unconstrained CR model, quality labor term with education and age interaction

	1986-1989				**1997-2000**			
	BE		*FE*		*BE*		*FE*	
	Coeff.	St. Err.	Coeff.	St. Err.	Coeff.	St. Err.	Coeff.	St. Err.
Female	0.143	0.252	0.321	0.191	-0.013	0.299	0.454	0.282
Noned. Old	-0.529	0.423	-0.104	0.237	-0.940	0.716	-0.156	0.405
Educ. Young	-0.202	1.004	0.383	0.727	0.196	0.997	-0.520	0.621
Educ. Old	-0.116	1.416	-2.410	0.860	-1.605	0.989	-1.448	0.662
Ln(capital)	0.311	0.055	0.364	0.054	0.203	0.047	0.101	0.047
Ln(labor)	0.643	0.076	0.159	0.070	0.690	0.085	0.725	0.079
Ln(avg. wage)	0.851	0.255	0.667	0.161	1.503	0.178	0.210	0.142

Appendix Table 10: Coefficient estimates, unconstrained CR model, quality labor term with gender and education interaction

	1986-1989				**1997-2000**			
	BE		*FE*		*BE*		*FE*	
	Coeff.	St. Err.	Coeff.	St. Err.	Coeff.	St. Err.	Coeff.	St. Err.
Experienced	-0.252	0.315	-0.411	0.171	-0.794	0.330	-0.338	0.234
Male, Educated	-0.702	0.877	-0.522	0.520	0.047	0.685	-1.009	0.496
Female, Noned.	-0.209	0.383	0.538	0.292	0.237	0.469	0.427	0.435
Female, Educ.	1.401	1.417	-1.833	1.135	0.306	1.404	0.042	0.918
Ln(capital)	0.301	0.054	0.364	0.054	0.206	0.047	0.095	0.047
Ln(labor)	0.653	0.075	0.145	0.070	0.685	0.085	0.726	0.078
Ln(avg. wage)	0.882	0.253	0.697	0.163	1.513	0.174	0.215	0.141

Appendix Table 11: Coefficient estimates, unconstrained CR model, quality labor term with gender and age interaction

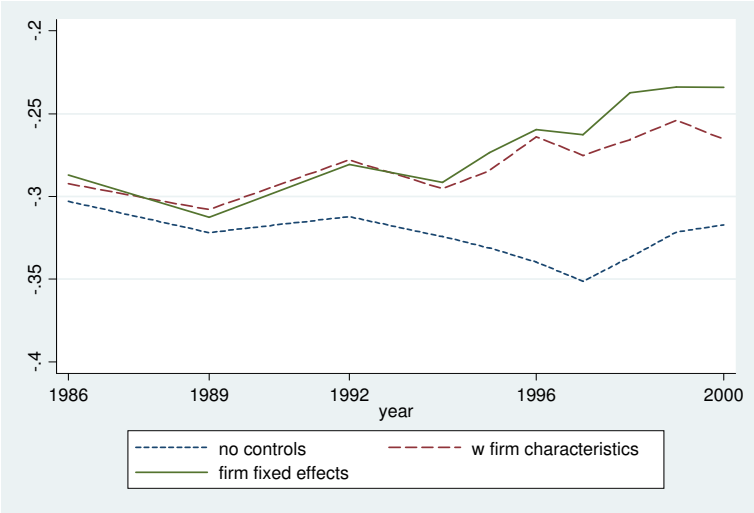
	1986-1989				**1997-2000**			
	BE		*FE*		*BE*		*FE*	
	Coeff.	St. Err.	Coeff.	St. Err.	Coeff.	St. Err.	Coeff.	St. Err.
Educated	0.212	0.283	-0.228	0.189	-0.049	0.380	-0.572	0.226
Male, old	-0.925	0.560	-0.623	0.281	-0.591	0.501	-0.334	0.401
Female, young	-0.425	0.786	0.986	0.562	0.786	0.852	0.960	0.733
Female, old	-0.460	0.674	-0.271	0.411	-1.216	0.785	-0.409	0.691
Ln(capital)	0.304	0.055	0.360	0.054	0.211	0.047	0.096	0.047
Ln(labor)	0.647	0.075	0.161	0.070	0.683	0.086	0.725	0.078
Ln(avg. wage)	0.824	0.252	0.679	0.164	1.514	0.180	0.221	0.140

Appendix Table 12: Effect of firm selection, 1997-2000

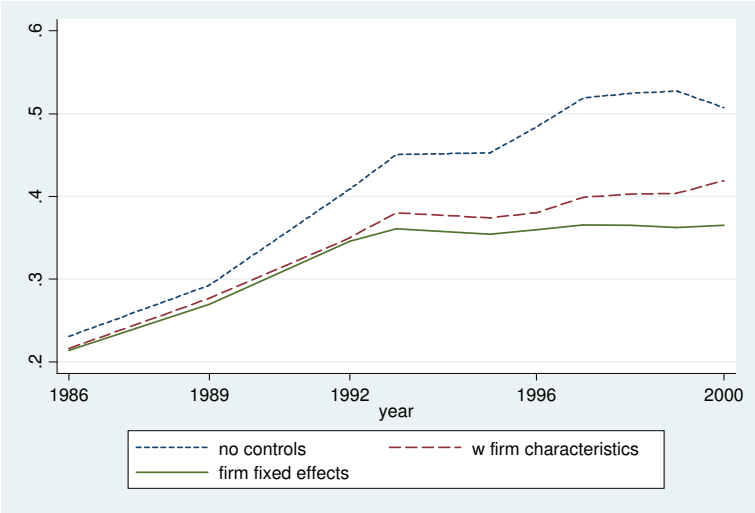
	FE				**With n. of times after added**			
	Balanced		*Unbalanced*		*FE*		*GMM*	
	Coeff.	St. Err.	Coeff.	St. Err.	Coeff.	St. Err.	Coeff.	St. Err.
Female	0.098	0.116	0.014	0.093	0.009	0.093	-0.058	0.235
Educated	-0.148	0.117	-0.236	0.087	-0.234	0.087	-0.424	0.222
Experienced	-0.137	0.111	-0.275	0.102	-0.278	0.103	-0.343	0.169
Ln(capital)	0.092	0.034	0.093	0.023	0.094	0.023	0.086	0.055
Ln(labor)	0.782	0.073	0.722	0.067	0.720	0.066	1.026	0.172
Ln(avg. wage)	0.232	0.082	0.317	0.080	0.316	0.080	1.091	0.318
N. of times firm appears after *t*					0.048	0.037	0.073	0.032

Appendix Table 13: First-difference and GMM estimation of Crépon et al. model

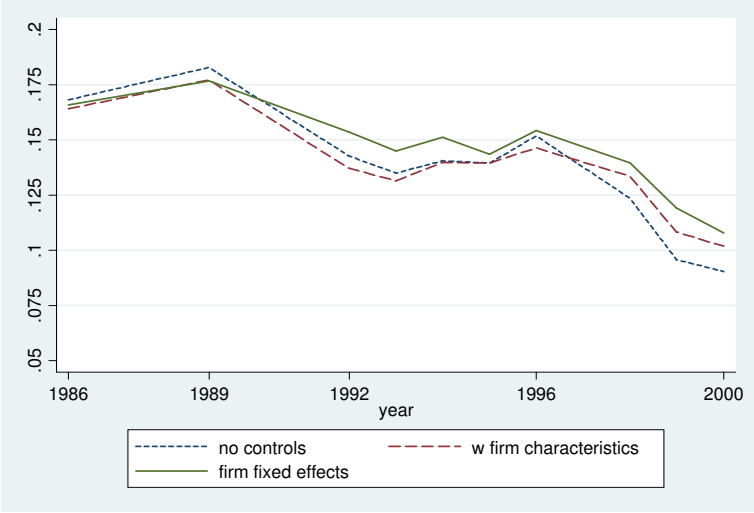
	Constrained				**Unconstrained**			
	FD		*System GMM*		*FD*		*System GMM*	
	Coeff.	St. Err.	Coeff.	St. Err.	Coeff.	St. Err.	Coeff.	St. Err.
Female	0.154	0.085	-0.115	0.256	0.098	0.080	-0.137	0.245
Educated	-0.281	0.078	-0.425	0.172	-0.158	0.077	-0.289	0.194
Experienced	-0.219	0.077	-0.342	0.168	-0.167	0.076	-0.377	0.155
Ln(capital)	0.109	0.028	0.103	0.058	0.074	0.026	0.128	0.056
Ln(labor)	0.457	0.042	1.077	0.115	0.715	0.058	1.043	0.148
Ln(avg. wage)					0.183	0.049	0.982	0.292



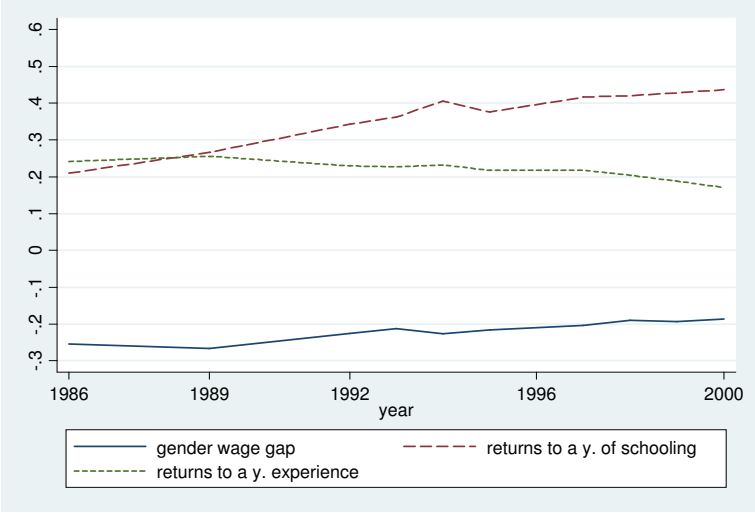
(a) gender wage gap



(b) education wage differential

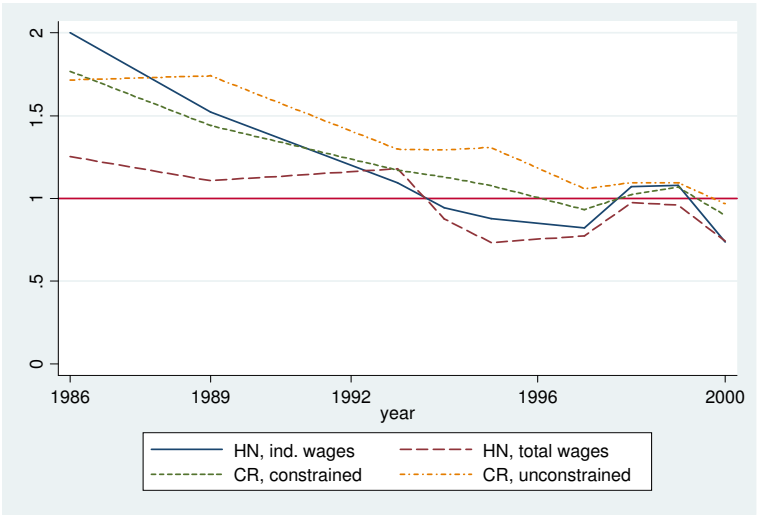


(c) experience wage differential

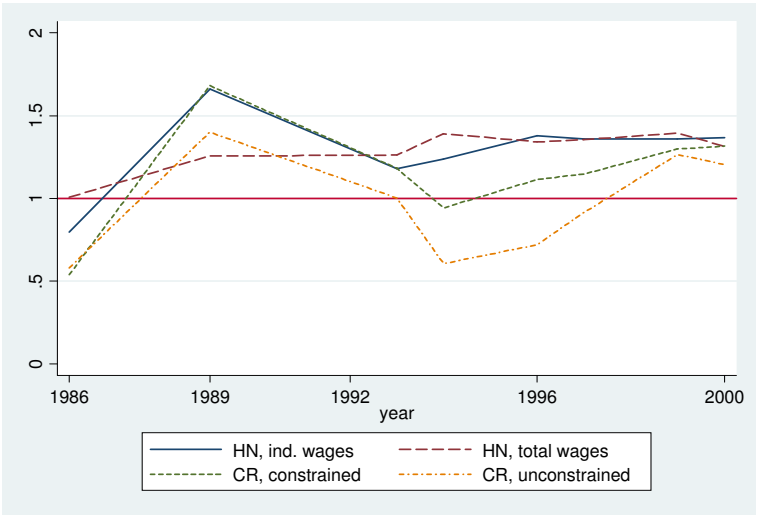


(d) detailed analysis

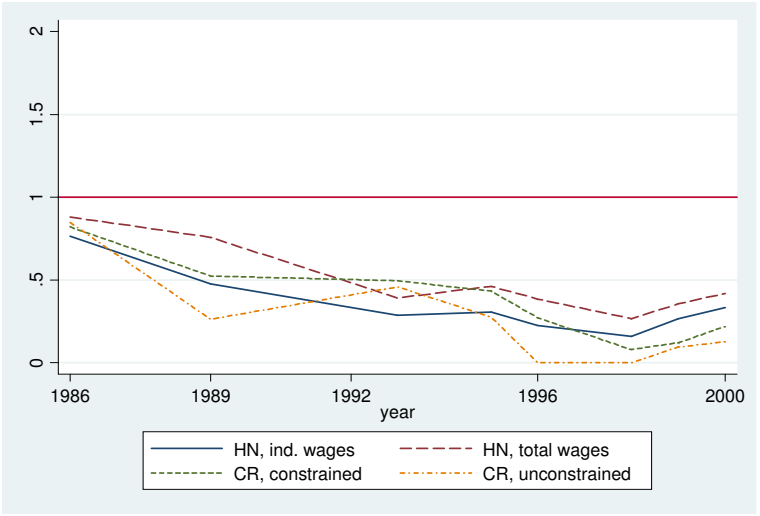
Figure 1: Evolution of relative wages, 1986-2000 (based on individual-level wage regressions)



(a) Female



(b) Educated



(c) Older

Figure 2: Evolution of estimated markdown coefficients of different types of labor 1986-2000 (four different specifications)