PRELIMINARY DRAFT: PLEASE DO NOT CITE

Dispersion of Earned Income in Transition: The Robustness of Earnings Inequality Growth in the Hungarian Business Sector, 1986-2005

Gábor Antal*
Central European University
antalg@ceu.hu

(Sept 2008)

Abstract
Exploiting a large linked employer-employee dataset of 1.7 million observations on workers employed by 39,000 enterprises, I study earnings inequality of full-time employees in the Hungarian business sector. I find that the dispersion of real monthly earnings increased rapidly from 1989 to 2000, then declined significantly until 2002, and started to rise again afterwards. This general pattern is reflected by several inequality measures, however, the magnitude - and in some years even the direction - of changes are indicated to be very different. The 90-10 interdecile ratio rose by 75% for men between 1989 and 2000, whereas most social welfare based and entropy measures record a growth rate of more than 150%. Inequality measures that are more sensitive to the upper tail exhibit the largest rate of increase. The data include hours worked for the years 2002-2005, and inequality of computed real hourly earnings is lower than of monthly earnings in this period. Results do not change significantly when controlling for the changing size criteria of sample inclusion for companies across years, but firm composition does matter in general, suggesting that between-firm inequality is a main factor in explaining inequality trends. The contribution of skill composition effects to inequality changes fluctuates strongly over time.

Keywords: inequality, earnings, measurement, wage differentials, transition, Hungary
JEL: J31, P20

* The author is grateful to János Köllő, Mónika Bálint and their colleagues at the IE – HAS for providing, cleaning and harmonizing the Hungarian Wage and Earnings Survey. I also thank my supervisor, John Earle, and Álmos Telegdy for friendly and professional support. Katalin Antal, Anna Lovász, Judit Máthé and Mariann Rigó at the CEU Labor Project contributed with valuable comments and great efforts to assemble the data.

1. Introduction

Wage inequality got into the center of attention of labor economists in the late eighties, and the enthusiasm of researchers to understand the driving forces behind the evolution of wage and earnings dispersion has not dwindled in the last two decades. However, the focus of this enthusiasm was aimed mostly at the labor market of the United States and of developed countries. Transition economies remained out of the spotlight, partly due to the lack of large-scale databases that include individual observations on wages.[1] The low research output is especially regrettable since transition provides an ideal setting to investigate changes in the level and the structure of wages. In Central and Eastern European countries, wage determination became decentralized within a couple of years, the labor market and other factor markets turned more flexible and more open to international competition. Employers faced harder budget constraints, a tougher competitive environment, but at the same time easier access to liquidity and other resources. Employees had to deal with massive job destruction during early transition and changing skill requirements, but also found new opportunities of education and training.

Hungary stands out regarding the speed of transition and thus provides a particularly valuable terrain for the analysis of the wage structure. The country was ahead in the market liberalization and privatization processes, and it became in many dimensions – like openness of markets, degree of corruption, development of regulation, ownership structure and business environment – more similar to developed Western European economies than other countries in the region. Regarding wage structure, Rutkowski (1996) shows that Hungary displayed the largest level of earnings inequality before the start of transition. In the light of this fact, it is even more remarkable that according to the OECD (2007), Hungary exhibited the largest growth in

[1] Noteworthy exceptions are Rutkowski (1996), Keane and Prasad (2006), Ganguli and Terrell (2006) and Kertesi and Köllő (2000).

earnings inequality between 1994 and 2005, as measured by the change in the 90-10 interdecile ratio. By 2005, Hungary had a level of earnings inequality exceeding that of the United States. Rising wage inequality is interesting in itself, but it also has farther-reaching implications. Milanovic (1999) found that in transition economies, the most important driving force behind the rise in overall income inequality is the rapid increase in the inequality of earnings. Moving from a long period of the common notion of pervasive income and earnings equality to such a high level of inequality within less than two decades makes people feel that these differences in the wage and income structure are "unjustified" and feeds a general disappointment with the transition process.

In this paper I analyze the extent of inequality growth of real monthly earnings in the Hungarian business sector over two decades between 1986 and 2005, using a large linked employer-employee dataset of 1.7 million observations on workers employed by 39,000 enterprises. Empirical evidence on overall Hungarian wage inequality is scarce and relates mostly to the early nineties. Éltető (1996), Pudney (1994) and Rutkowski (1996) all document a huge increase in earnings inequality during the early phase of transition, using different datasets. The most thorough analysis of inequality is provided by Kertesi and Köllő (1997), but their last year of observation is 1996. In later years, there has been little attention paid to inequality indices, although some aspects of the wage structure were investigated. For example, Campos and Joliffe (2004) estimate the effects of market liberalization on gender wage gap; Kertesi and Köllő study skill differentials (2002) and industrial wage differences (2003a, 2003b); Neuman (1997, 2002) explores the effect of collective wage bargaining; and Köllő and Nagy (1996) study the effects of unemployment on earnings. I show that earnings inequality continued to rise in the late nineties until 2000 – though at a smaller pace than in the first half of the decades –, declined significantly between 2000 and 2002, and got close again to its 2000 peak level in the last three years of the sample. Besides the time pattern of inequality changes, I also explore how different parts of the distribution were affected in a given period, and I find that inequality above the median was on the rise constantly, while the dispersion among below-median employees declined due mainly to a 70 percent increase in the real value of the minimum wage. Low-skilled men lost the most in terms of real earnings and high-skilled women enjoyed the highest wage gains.

Most economists concerned with wage inequality tend to use a rather narrow set of inequality measures, typically interdecile differentials, the standard deviation (or variance) of log wages and the Gini coefficient. Beyond the fact that these measures do not fulfill some desirable properties, other measures – applied extensively in sociology and the income inequality literature – may yield different results empirically with respect to the direction and magnitude of earnings inequality changes. Studying the distribution of individual earnings in the United States between 1967 and 1986, Karoly (1992) demonstrated how, despite common general tendencies, various inequality indices imply different levels of inequality growth. In some cases, alternative measures even differ in how they rank yearly earnings distributions ordinally. To check the robustness of inequality growth, I use a broad set of inequality measures that differ *inter alia* in their sensitivity to changes in certain parts of the earnings distribution. The results in this paper are in line with Karoly's findings: alternative measures produce up to three-fold differences regarding the magnitude of yearly inequality growth rates, moreover, in some years, some indices display changes opposite in sign to what the majority of measures show.

Finally, I conduct a series of further robustness exercises. I check whether the varying size threshold of sample inclusion for employers affects results, and I also assess to some extent the effect of firm composition. For the last four years of the sample period (2002-2005), the data provide information on weekly regular and monthly paid hours, thus I can construct two measures of real hourly earnings, and confront results based on hourly earnings dispersion with values from monthly earnings distributions. I find that for the available years, inequality measures of hourly earnings are very close to what I get when working with monthly earnings. As a last step, I control for changes in the composition of the work force by applying the reweighting method of DiNardo et al. (1996) to construct counterfactual inequality measures where skill distribution is held constant at its start-of-period or end-of-period level. Composition effects do matter in all years, but their importance and nature varies constantly.

In the following section, I introduce the dataset used for the analysis, and explain how observations were weighted to account for different levels of representation of different groups. Section 3 gives a general overview on the evolution of wage inequality by using visual tools and commonly applied inequality measures. Both the time and the cross-section dimensions of earnings dispersion are discussed. In Section 4, I add various inequality and two earnings measures to the analysis, to check how results vary with the choice among these alternatives. I advert to the issue of composition effects in Section 5 where I reweight the data to isolate the impact of the changing skill distribution of the work force. Finally, Section 6 summarizes the results and sets some directions for future research.

2. Data Sources and Sample Construction

The dataset used in this paper is the Hungarian Wage and Earnings Survey (HWES), the most appropriate data source available in the country for wage dispersion analysis. The HWES is conducted at the level of firms, but its output is individual-level data on the employees of sampled companies. I also rely on firm-level data assembled by the Hungarian Tax Authority (HTA) when constructing sample weights. A short description of the weighting scheme and that of the HTA database follows later in this section.

The data host of the HWES is the National Employment Office, but data collection is carried out by the Central Statistical Office of Hungary (CSO). It is a matched employer-employee dataset, existing since 1986, containing yearly information on workers and establishments at the plant level. Although the survey was not executed in 1987, 1988, 1990 and 1991, the data still include two years prior to transition, which enables the researcher to address the question of how earnings inequality changed while the country was moving from a socialist to a capitalist regime. The last available year in the survey is 2005.

The HWES is based on a questionnaire filled out by a sample of Hungarian firms in May each year. Employers are requested to provide information on the size of their work force, on the number of blue- and white-collar workers, on their 4-digit NACE code, and on several characteristics of a sample of their employees selected according to different sampling guidelines for blue-collar and white-collar workers. Additional information on the geographical area of the plants' location is supplemented by the CSO.[2] Until 2001, only data on full-time employees were collected, part-time workers should have joined the target population from the following year according to the description of the database. However, the first year in the data when workers actually classify as part-timers based on their reported regular working hours is 2003. I define part-time employees as workers who reported less than 35 regular hours a week, and I excluded them from the sample of 2003, 2004 and 2005, thus all results in the paper consistently refer to full-time employees only.

[2] Regarding employer and geographical location information, only the number of employees was used in this paper to construct weights, and to cut the sample at different size thresholds.

The sampling frameworks for both employers and employees have changed several times during the 1986-2005 period, affecting the size, coverage and consistency of the dataset. In 1986 and 1989, all firms of the Hungarian business sector were surveyed. In all years starting with 1992, all companies with more than 20 employees were included in the sample. From 1994, in addition, a random sample of a sub-population of smaller-sized firms was selected, where this sub-population covered employers with 11-20 employees between 1994 and 1999, and those with 5-20 employees for the period 2000-2005

With respect to sampling of full-time workers, we can distinguish between three main regimes.[3] The first regime refers to sampling practice before transition (i.e. to the years 1986 and 1989), the second covers the years 1992-1993, and the third captures the sampling procedure between 1994 and 2005. During the first regime, all senior managers were included. A random sample of the rest of the professionals was selected consisting of the first and then every fifth person of groups formed by workers of similar qualification and working conditions in 1986, and the first and every tenth employee in 1989.[4] In case of manual workers, the survey covered the first and then every seventh worker of each group in 1986, while the first and every tenth person in 1989. In the three years of the second regime, every blue-collar worker born on the 5th or 15th of any month and every white-collar worker born on the 5th, 15th or 25th should have been included in the sample. This scheme was maintained during the third regime for firms above a certain size, however, all employees' information were required from sampled companies not exceeding that limit. The size threshold was 20 employees from 1994 to 2001, and it has been 50 employees since 2002. Foreign employees without a residence permit, pensioners working

[3] Since most analyses in this paper only deal with full-time employees, the sampling of part-timers is not described here in detail.
[4] Under the socialist regime, all employees were classified into so-called "tariff categories", based on qualification of the worker and working conditions of the job. Membership in the categories determined one's wage, hence the term "tariff".

full-time, employees working abroad (except for a delegation), employees on loan and on exchange (at another firm or from another firm), and employees who were not receiving wage for more than three days in May were not sampled.

In order to account for the different degree of representation of the two occupational groups of employees within firm, I constructed weights for each group separately – called individual weights henceforth – as given by the ratio of the number of employees of either type on payroll in May reported by the firm to the actual number of workers in the sample.[5] Individual weights thus show how many individuals are represented by one observation within any given firm. The HWES contains a lot of mistakes regarding reported May employment, which were cleaned where possible using time-series information on the same variable and on average yearly number of employees as well. Whenever the construction of individual weights was impossible due to either missing data or to data errors, I used average individual weights of firms of similar size, or in some cases theoretical weights, that is, the inverse of the probability that a given employee was selected into the sample.

In addition to within-firm weights, to render the sample representative of the non-public sector of the national economy, I am also applying company-level weights computed by dividing the total number of employees in arbitrarily defined size categories by the sum of individual weights within the corresponding size category. The figures on total employment by size are gained by adding up firm-level employment numbers found in the HTA database, which virtually contains every firm registered in Hungary that conducts double-entry bookkeeping for the years

[5] The weighting procedure relies heavily on the ideas of and preliminary work done by researchers at the Institute of Economics of the Hungarian Academy of Sciences.

1992-2005.[6] For 1986 and 1989, I used employment information from annual labor market statistics.[7] The final weight of each worker-year observation is then simply the product of the individual and the company weight, and it approximates the number of employees of the Hungarian business sector represented by a single worker in the sample.

The complete HWES dataset comprises 1,901,159 full-time worker-year observations, and a total of 39,373 unique firms that employ at least one full-time employee. However, there are several reasons to further restrict the sample. First, data files include firms also below the sampling size limits pointed out earlier in this section (21, 10 and 5 employees, depending on the period), and their size distribution unfortunately rejects the possibility that they may have gotten into the sample randomly. Second, only a few firms with less than 21 employees are observed in 1994 and 1995, although a random sample of enterprises with 10 or more workers should be included according to sampling guidelines, so I dropped all firms below or at the twenty-employee threshold. Third, I wanted to keep workers with a relatively strong labor market attachment, and also to account for the changes in retirement age and early retirement schemes during the 1986-2005 period. At the end, I end up with a sub-sample of 1,723,213 individual observations for employees aged 19-54, who have less than 40 years of experience, work full-time and earn at least half the 1997 real minimum wage. These individuals are employed by a total of 39,088 employers.

Both individual- and firm-level data were cleaned thoroughly by researchers and research assistants at the Institute of Economics of the Hungarian Academy of Sciences (IE-HAS) and at

[6] There is a discrepancy regarding number of employees between the two sources in that the HWES provides information on employment in May, while firms report their end-of-financial-year average statistical employment to the Tax Authority. Thus, a firm may fall into different categories when calculating the numerator and denominator of company weights.
[7] Information on employment by size categories does not exist prior to 1989, thus I assumed that the size distribution of employers remained unchanged through the pre-transition period and imputed total employment numbers for 1986 accordingly.

the CEU Labor Project. We managed to ensure to a great extent the continuity and consistency of the database. In addition to the aforementioned process of cleaning variables necessary for the construction of sample weights, extensive cleaning efforts were made to get rid of spurious company exits and entries by detecting longitudinal linkages among exiters and entrants[8]; the content of several variables was harmonized across years; roundtripper values (implausible one-period changes) were fixed and missing values were filled where the imputation was obvious.

Table 1 summarizes year-by-year information on the unweighted size of the sample, and on the number of workers and firms represented by the sample conditional on individual and company weights. Note that the large increase in the number of represented firms is due to changes in the sampling design already described earlier in this section. In order to account for these changes, I will present results that rely on a consistent sample that includes only firms which have at least 20 employees in a given year, and also on one that includes continuing firms only.

The Wage Survey contains information on worker characteristics such as age, experience, highest education completed, gender, current occupation, and detailed data on various earnings categories, which allows checking the robustness of my results regarding direction and magnitude of the change in earnings dispersion by using several earnings definitions. For baseline measurements, however, I am using the broadest gross monthly earnings category that is consistently available for all years in the HWES, and which is defined as the sum of all payments to the employee in May (September in 1986) at the expense of the employer's wage cost account,

[8] Spurious exit and entry are common to establishment-level datasets, and also present to a great extent in the Hungarian data. The cleaning procedure benefited first of all from a registry compiled by the CSO, which revealed valuable data until 2002 on boundary changes of companies, such as mergers, acquisitions, split-ups and spin-offs, and also provided information on spurious changes in continuing firms' identification numbers due to re-registration or bankruptcy, for example. In addition, we found longitudinal links in the data by matching exiting and entering firms comparing their employment, settlement and industry codes, ownership and net sales revenue.

including base salary, allowances (for overtime, night shift, language proficiency, work abroad, etc), regular monthly premia, bonuses and commissions, and one-twelfth of total premia, bonuses, commissions and thirteenth-month salary passed in the previous year. Since the personal income tax system was only introduced January 1, 1988 in Hungary, I grossed reported 1986 monthly earnings using the 1988 income tax brackets for 1986 values inflated to 1988 Hungarian forints. By measuring inequality in terms of gross earnings, the redistributive effects of personal income tax influencing the level of inequality are not taken into account, which is in line with the main goal of this paper being a documentation of earnings, and not income, dispersion.

3. Evolution of Earnings Dispersion: Baseline Results

To get a broad, though instructive picture on the evolution of earnings inequality in the last two decades in Hungary, Figure 1 presents estimates of kernel density functions for the distribution of log real gross monthly earnings by gender. The change in the shape of the distribution is remarkable even when comparing years only visually. Pre-transition years exhibit a tightly compressed density function for both genders, which supports our image of a non-competitive labor market, and a centrally controlled wage-setting regime biased towards social equality leaving no or very little scope for individual companies and employees to negotiate wages. Starting in 1992, we see the earnings distributions spreading out gradually over the years, reaching the highest level of dispersion somewhere at the end of the nineties, and then a seemingly decreasing pattern of the variance after 2000. As it might have been expected, the mean of the earnings distribution is higher for men than for women, but the gap seems to have been narrowing constantly since the first year of observation. This trend is analogous to the findings of Campos and Joliffe (2004) who show that the gender wage difference in log wages declined from 0.31 to 0.19 between 1986 and 1998 in Hungary. It is also in line with other countries' experience in Central and Eastern Europe – Brainerd (2000) documented a reduction in the gender wage gap during transition in Slovenia, the Czech Republic, Poland, Estonia, East Germany and Slovakia. Nonetheless, within-group inequality was on the rise in Hungary for both genders as demonstrated later in the paper. Note also the large difference between the earnings of men and women in 1986 and 1989, which seems query the widely stressed socialist claim of gender equality on the labor market, but of course we cannot draw substantive conclusions from this unconditional wage gap.

For each year, I marked the logarithm of the prevailing real minimum wage by a solid vertical line, except for 1986 when the institution not yet existed. The minimum wage and monthly earnings are not directly comparable, since the latter includes non-wage elements, like overtime pay, bonuses, premia, and so forth, but even so the minimum wage seems to play an important role in determining the shape of the earnings distribution. From 1994 on, the distributions are two-modal, and following a significant rise in the real value of the minimum wage between 2000 and 2001, it becomes the highest frequency point of the kernel. Based on this visual evidence, the impact on inequality of this institution is as pronounced as in the United States over a different period, 1979-1988, just working in the opposite direction. DiNardo et al. (1996) point out that during the eighties the decline in the real minimum wage explains about twenty-five percent of the change in the standard deviation of the male log wage distribution, and up to thirty percent for women. It is also evident that such a high count of workers earning the minimum wage does not seem plausible. In Hungary, reporting minimum wages is one of several common tax avoidance practices, and it sheds some doubt on the direct interpretation of inequality measures.

Figure 2 plots the evolution of five selected points of the real earnings distribution: the 90th, the 75th, the 25th, the 10th percentiles and the median. Except for the top-earners and for the year 1994, we can observe a declining trend in real earnings for the first ten years, and even after for workers below the median. The 90th percentile employee enjoyed real wage gains in all but 6 years, with a particularly pronounced average growth of 4.6 percent per year following 1996. Real earnings at the 75th, the 50th and the 25th percentiles started to recover in 1997, 1998 and 1999, respectively, at a pace getting smaller as we move from the top to the bottom of the distribution. Despite the pervasive increasing tendency from the second half of the nineties, even by the end of the period nobody above the median earned more in real terms than the equivalent employee in the 1986 distribution, with losses ranging from 15.9 to 2.4 percent, while workers at the two upper quantiles gained 16 and 42.7 percent. However, even for 75th-percentile-employees, the first year above the 1986 level was 2002. The gap between the most and the less skilled was high in 2005, but 2002 shows the largest difference, when the 90th percentile was up by 12.1 percent compared to its 1986 level, while the 10th percentile was down by almost 40 percent. The narrowing of the gap is due to the surprising behavior of the value of the 10th percentile, which is explained by the fact that the minimum wage became the 10th percentile of the distribution. The nominal value of the minimum wage was officially raised by almost 57 percent from 2000 to 2001 and by an additional 25 percent in the following year, which resulted in yearly increments of 44 percent and 19 percent in real terms, respectively. As a consequence, 16.6 percent of the employees in the weighted sample earned not more than the minimum wage by 2002, whereas the figure was 5.2 percent in 2000 (and 11.8 percent in 2001).[9]

In Figure 3a, I show how the changes in real earnings at selected percentiles translate to changes in earnings inequality by following the evolution of the difference between the 90th, 50th

[9] Without weighting, the figures are somewhat higher: 17, 6.1 and 14 percent, respectively.

and 10th percentiles of the log earnings distribution separately for women and men. Within just fifteen years, between 1986 and 2000, both women and men experience an almost steady increase in the 90-10 log earnings differential of 74.4 percent and 69.9 percent, respectively, with a slight decrease in 1993 for men, and in 1995 for both genders. About 40 percent of the total growth is concentrated between 1989 and 1992, the first years of transition. Translating logs to levels, the ratio of the earnings of the most skilled (ninetieth percentile) to that of the least skilled (tenth percentile) reached, at its peak in 2000, 4.9 in case of women and 5.6 in case of men. The general pattern of the rate of wage dispersion growth in this period is closely reflected in both the top and the bottom ends of the distribution, with slightly decreasing differences in the levels. On average, inequality within the more skilled group is higher by 20 percent for women and by roughly 13 percent for men.

After this steep climb, inequality drops sharply for two consecutive years after 2000 for both groups, hitting even its 1993 level for female workers in 2002. The observed parallel paths in the 1986-2000 period for employees below and above the median now diverge clearly. Since top end inequality keeps on rising at the same pace as before, the fall in overall inequality is completely due to the declining 50-10 differential. As I mentioned in the previous paragraph, the minimum wage became the tenth percentile regarding the whole distribution, and since this is also true when splitting the sample by gender, the rise in the real minimum wage corresponds to an almost one-to-one decline in the 50-10 differential, because the median increased by less than 0.1 log points in this period.

After 2002, overall inequality is back to its pre-2000 growth path with continuously increasing top end dispersion and a turn of the decreasing trend in the bottom part of the earnings distribution. Although the minimum wage remains the tenth percentile, its real value declines in all the last years but 2005, while that of the median does not, except for men in 2004. By the end of the series, the level of wage dispersion measured by the 90-10 interdecile differential does not reach the maximum observed in 2000, but it is still very high even by international standards, exceeding 1.4 in case of women and 1.6 for men.

As a consequence of the sampling design outlined in Section 2, inequality is not directly comparable across years in the sense that the employment threshold of sample inclusion was changing and firms with less than 21 employees are not represented in each year. Figure 3b charts results on a sample that only includes enterprises which employ more than 20 people in a given year. Note that we may expect differences only beginning with 1996, since the HWES contains no plausible information on companies with less than 21 employees in earlier years. Generally speaking, for both genders, the pattern of inequality changes is very similar to that computed on the full sample, however, the magnitude of changes in periods of both increasing and decreasing dispersion is smaller. For women, the peak in the 90-10 gap is lower by around six percent and by 2002, inequality drops to a level four percent higher than in the complete sample. The end-of-the-period numbers are almost identical. Practically the complete difference in the two series comes from the lower end of the distribution, since the 90-50 differentials are very close, which is not surprising considering the fact that the major driving force in the evolution of the 50-10 (and the 90-10) gap is the change in the real minimum wage in this period, and smaller firms are more affected by it. For men, discrepancies are even bigger. 2005, not 2000, is the year with the highest level of inequality in the consistent sample, mainly because the 90-10 measure is lower by 8.5 percent in 2000, but also it is higher by one percent in 2005. The drop between 2000 and 2002 is less dramatic in both the 90-10 and the 50-10 earnings differentials. The rate of growth of earnings dispersion in the upper half of the distribution is slightly attenuated.

To consider another aspect of sample consistency, Figure 3c shows the history of the same three measures, this time only for companies which answered the survey in at least 14 years. This means that all old firms (i.e. those which existed prior to transition) followed for at least 12 years after transition, and all new enterprises (i.e. those which are observed only after 1990) with a complete spell are included. After changes in firm composition are ruled out in this very crude way, the picture is very different, especially in case of males. For females, the previous pattern is to some extent recognizable, although the highest level of the 90-10 differential is only 1.32 (recorded in 2005); inequality starts to fall already in 1998, but from a much lower level; and since the lowest point in 2002 is the same as in the preceding analysis, the rate of decrease is only moderate. Except for the jump from 1989 to 1992, and a very modest increase between 1992 and 1997, the paths of all measures are smoothed out for men. This observation, combined with the results on the full sample, suggests that almost all of the growth in earnings inequality is a within-firm phenomenon in early transition, but basically all changes occur between firms afterwards.[10]

So far, except for the visual evidence of kernel density estimates, I have been concentrating only on some points of the earnings distribution. Figure 4 departs from this exercise and displays how real earnings changed over the complete distribution (with the exception of the top and bottom three percents) during the complete sample period and in five sub-periods. The first panel in the upper row gives an overview by plotting earnings differences between 1986 and 2005. Real earnings growth is a convex U-shaped function of the position in the distribution

[10] Certainly, the proper way to address this issue would be a decomposition of inequality chenges to within-firm and between-firm components, but it is beyond the scope of this paper.

with employees located between roughly the tenth and thirtieth percentiles being at the bottom of the U. In case of women, this is the only group which saw its earned income decline compared to the base year, while for men, everybody under the 67 percentile mark earned less in 2005 than the respective employee in 1986. The slope of the curve becomes very steep after the ninetieth percentile, earnings of the most skilled women more than doubled and we can also observe growth rates in the 60-80 percent range in the topmost male earnings brackets. Notice also the closing gender wage gap which was already visible on Figure 1. On average, women gained by thirty percentage points more over the whole 1986-2005 period at all points of the real earnings distribution.

We discover very different trends when we split the sample into sub-periods based on the nature of inequality changes suggested by the previous graphs (Figure 2 and 3a). As demonstrated by the second panel in the first row of Figure 4, in the pre-transition period there are no significant differences in the real earnings changes across percentiles, with the exception of women above the ninetieth percentile. Male workers at all percentiles of the distribution experienced a roughly equal decline of around ten percent, while most female workers only 2-3 percent, and of the whole population only the top twenty percent of the female distribution had higher earnings than three years before. As mentioned earlier, 40 percent of the growth in the 90-10 differential is concentrated in the 1989-1992 period. The third panel indicates clearly that inequality growth affected the complete distribution evenly, since changes in real earnings are an almost linear – positively sloped – function of the position in the earnings distribution. Roughly speaking, only women above the median gain from transition in terms of higher real earnings, those in the bottom half see their earned income decline, the least skilled by nearly twenty percent. Men are in a worse situation with sharply decreasing real earnings for all but the highest ten centiles. Employees in the bottom twenty percent loose more than twenty percent on average. This linear pattern is very similar to what Juhn, Murphy and Pierce (1993) find regarding changes in real weekly wages in the late 1970s and in the 1980s in the United States, and to the results of Kertesi and Köllő (1997) for real monthly earnings in the Hungarian economy in the same period.

Linearity disappears in the 1990s, and the relationship between the change in earnings and position in the distribution becomes convex above the median and below the tenth percentile, and concave between the tenth percentile and the median. Workers of both genders in the lower three quartiles of the distribution experience real earnings losses in this period, and again it is male employees below the twentieth percentile who suffer the largest decrease. It is also clear from the fourth panel of Figure 4 that the pervasive advantage of women in real earnings growth now disappears in the upper half of the distribution, but remains close to ten percentage points in the very bottom. (Of course "advantage" here means only a smaller decline.) As pointed out earlier, the years between 2000 and 2002 are special in the sense that inequality changes seem to be driven by the minimum wage. For percentiles which already earned the minimum wage in 2000, the 71 percent increase in the real value of the minimum wage obviously translates directly into the measured real earnings change as apparent from the second panel in the lower row of Figure 4. However, also the above average growth rates of the bottom twenty percent of both distributions can be explained by the fact that these employees all earned the minimum wage in 2002 that was way higher in real terms than what they had gotten in 2000. It is also noteworthy that although the earnings growth rates of the two gender groups were equalized in this period at most parts of the distribution, women between the tenth and fortieth percentiles still enjoyed a higher climb in salaries than their male counterparts. The very skilled earned ten percent more at the end of the period than in the base year. Finally, in the last three years of the sample time series, we may observe linear profiles similar to the 1989-1992 period, but now the rates of increase are practically the same for both genders.

For a striking comparison with tendencies in the United States, I took the sample of men and pooled the sub-periods in Figure 4 to end up with only two. The resulting graph, Figure 5, is in many aspects very similar to Figure 4 in Lemieux (2007), which illustrates changes in the real hourly wages of men by percentile on May/ORG CPS data. Lemieux divides the thirty years between 1974 and 2004 into two periods of equal length. In the second half of the 1970s and in the 1980s, the growth in inequality was pervasive, meaning that the higher we climb along the percentiles of the distribution the larger (the less negative) is the growth rate of wages, moreover, the relationship was linear (except for the upmost eight percent, where it was convex). In stark contrast, during the nineties and after the turn of the century, the curve of wage growth plotted against location in the distribution became U-shaped. Workers between the twentieth and fiftieth percentiles experienced the smallest gains, and were outpaced by both upper-tail and lower-tail workers (with the highest increase in the top fifteen percent). Figure 5 tells a similar story for Hungary. The only differences are in the magnitudes and the time horizon. The linear pattern is present in the second half of the eighties and in the first years of transition, but earnings changes range between -31 and 8 percent (for the U.S., these numbers are -20 and 12) within a time span of only six years. The interception with the zero line is at the 96th percentile, not at 79th, as on the graph of Lemieux. We can also see a pronounced U-shaped curve in the 1992-2005 period, which is even more asymmetric than the CPS counterpart with workers at the bottom reaching earnings growth rates of more than thirty percent, while those at the top enjoying twice that pace. The interval of the "worst" percentiles is slightly shifted to the left compared to the U.S. graph, since employees between the tenth and thirty-fifth percentiles have to settle with earnings losses, while at all other parts of the distribution we see positive differentials. It is an interesting question whether the same mechanisms work behind the changing nature of inequality growth that were listed by Lemieux as possible explanations for the American trends, or something inherently different is happening in Hungary that leads to observationally equivalent outcomes.

4. Robustness of Inequality Changes: Alternative Inequality and Earnings Measures

When we ask the question how earnings inequality has changed over time, we may think of an ordinal and a cardinal dimension of the answer. In the former sense, we would simply like to rank distributions according to the measured level of their dispersion, while in the latter case we are also interested in the magnitude of the inequality change from one year to another. No matter which dimension we are concerned with, the answer will depend on the choice of the method of measurement. This method may be a visual tool, like a density or a cumulative distribution function, or a Lorenz-curve; or we may want to grasp the level of inequality by just one single measure as in the case of percentile differentials, the standard deviation or the Gini coefficient. It can be shown that it is only possible to rank two distributions unambiguously according to dispersion, if the corresponding Lorenz-curves do not intersect, that is, one of them lies closer to the 45-degree-line representing perfect equality at all points on the axis of cumulative frequencies.[11] If, however, the Lorenz-curves do intersect – and this is what we see most often in real data –, we can always find two inequality measures that will give different answers to our

[11] See Atkinson [1970], Cowell [1977]

ordinal question. And even if we are dealing with two ordinally equivalent measures, they might have very different cardinal properties.

Given the wide abundance of inequality measures applied in the income inequality literature, and the contingency of results on the choice of measure, it is to some extent surprising that most economists analyzing wage and earnings inequality tend to use a limited set of inequality measures, usually the interdecile differentials, the standard deviation of log wages and the Gini coefficient. Using CPS data, Karoly (1992) has shown empirically that different measures yield different implications regarding the direction and the magnitude of changes in wage inequality in the United States from 1967 to1986. Of course, it is always to a great extent a subjective decision which inequality measures the researcher might want to apply, since different measures possess dissimilar properties with respect to how they respond to transfers between different points of the distribution. The 90-10 gap, for example, is not affected by any earnings changes in the middle part of the distribution, as long as these do not influence the value of the tenth and the ninetieth percentiles. Also, the 90-10 differential may record a huge growth in wage inequality if its two components get further away over time, even if accompanied by equalizing redistributive processes around the mean which might cause other measures to document a more moderate growth rate, or even a decline. Thus, the choice of inequality measure is not only subjective, but also indicative of the nature of the mechanisms driving inequality changes. Once we know the sensitivity properties of these measures, we are able to infer what parts of the distribution contributed most to the general inequality trends. Figures 4 and 5 serve the same purpose in this respect, but the advantage of single inequality measures is that since they compress information to a single number they help answering policy questions related to the magnitude of changes – of course, only conditional on the researcher's direction of attention to a certain section of the earnings distribution.

We can do more than just rely on a subjective choice if we *ex ante* postulate some desirable properties of inequality measures. Cowell (1977) lists three of such principles that may help to reduce the set of possible choices. First, he argues that an inequality measure should be independent of scale, that is, if everybody in the population earns a scalar multiple of what he or she earned in the previous period, then inequality should not change. Second, an inequality measure should satisfy the principle of population. If the population is merged with an identical one, measured inequality should remain the same. (Of course, we can think of any finite replication.) Finally, a desirable property may be the weak principle of transfers which states that as a consequence of a positive transfer from a higher to a lower part of the distribution (in a way that the recipient does not become "richer" than the "donor") inequality should decrease. The last principle is static in the sense that it assumes constant total income between two periods that is simply redistributed. A dynamization would involve a skilled and a less-skilled worker, both of them enjoying (additive) wage gains from one period to the next, but the increase for the less skilled is higher. Other things being equal, this should result in a drop in inequality. Karoly (1992) mentions an even stronger version of this idea, the principle of equal additions, which, unlike independence of scale, concerns disproportionate simultaneous changes in everybody's earnings: If every member of the population earns by a constant amount more than before, then measured inequality should fall. However, the principle of equal additions does not imply the principle of transfers.

Concerning the most frequently used measures, the standard deviation (of levels) fails the first two principles, while the standard deviation of logs satisfies both, but fails the weak principle of transfers. The 90-10 measure is independent of scale and of replicating the population, but just fails the principle of transfers. "Just fails" means that it *may not* decrease following a transfer specified in the previous paragraph. (It will never record any change unless one of the participants of the transfer is at the tenth or ninetieth percentile of the distribution.) The Gini passes all the above tests, but is very sensitive to transfers around the mode of the distribution. In this section, I compare the evolution of eleven inequality measures over time, including Theil's index and entropy measure (THEIL, ENT – Theil [1967]), Atkinson's measure with inequality aversion parameters of 0.5, 1 and 2 (ATK 0.5, ATK 1, ATK 2 – Atkinson [1970]), the mean logarithmic deviation (MLD), the coefficient of variation (CV), the relative mean deviation (RMD), the Gini coefficient (GINI), the standard deviation of logarithms (SDL), and finally, the interdecile differential as a benchmark (90-10). All Theil and Atkinson measures and the MLD, CV and RMD measures are independent of scale and satisfy the principle of population and that of equal additions. Concerning reaction to transfers from the rich to the poor, and setting aside already discussed measures for a moment, all measures but the RMD satisfy the weak principle of transfers. The RMD just fails, similarly to the 90-10 differential.

Atkinson's measures contain a parameter ε that represents the society's degree of inequality aversion. Increasing ε has two consequences: the higher will be measured inequality in any given year, and the more sensitive will Atkinson's measure become to changes over time in the bottom part of the distribution. Similarly, the Theil index, Theil's entropy measure and the MLD are members of the General Entropy Class (GE) of inequality measures with parameter α, where α is 1, -1and 0, respectively. The higher the value of α, the more sensitive are GE measures to transfers at the top of the distribution relative to changes at the lower end. RMD and the Gini are both sensitive to redistribution around the centre of the distribution, centre meaning the arithmetic mean in case of the former, and the mode for the latter. In fact, the RMD has the special feature of being insensitive to redistribution on the same side of the mean. The coefficient of variation attaches no extra weight to changes in any part of the distribution.

I plotted standardized values of these measures on Figures 6a and 6b by gender, with 1986 being the base year. Both figures consist of three panels, since I grouped measures that exhibit similar time paths. Each panel includes the 90-10 differential as a benchmark, and as reference to the previous section. In general, the message of alternative inequality measures is very similar to that displayed by the 90-10 differential: Inequality was growing from 1989 to 2000 – with a short break in 1995 –, then it declined sharply until 2002, but was on an increasing path again afterwards. However, after closer inspection, we find differences not only in the magnitude, but also in the direction of changes. First, consider the ordinal ranking of yearly earnings distributions by the various inequality measures. In case of women, there are three measures that in some years point into different directions than the others, the coefficient of variation, Theil's index and the Atkinson measure with an inequality aversion parameter of one half. Out of this group, the CV is out of line most frequently, it gives different ordinal ranking of subsequent distributions in four years. The biggest controversy occurs in 1994, all of the three measures document decreasing inequality – THEIL and ATK 0.5 only a modest fall, but the CV one of -9.5 percent – while according to all other measures inequality was rising by three percent on average. This suggests that redistribution in the upper end of the distribution resulted in greater equality, but it was not captured by measures that are not particularly sensitive to such processes. Interestingly, the 90-50 gap also increased, albeit only slightly. Another odd year is 2002, when all measures but the CV clearly indicate a reduction in earnings inequality (mostly of at least three percent, close to five percent on average). Also next year, inequality seems to be on the rise again, but not as measured by the Theil index and the CV. For male distributions, differences in the signs of inequality changes are even richer. Already between 1986 and 1989, where the 90-10 differential is falling by 1.2 percent and the standard deviation of log earnings does not change (and it was hard to take away anything meaningful from the flat profile of earnings changes by percentile on Figure 4), all other measures do indicate an ascent in inequality. Those sensitive to the bottom and the middle show only a very modest one (0.5-1 percent), but the others a rate close to four percent. Similarly, in 1993, the 90-10 gap is the only measure to exhibit declining inequality, while measures sensitive to the top end change by 7-10 percent. Finally, we saw that 2002 was the second year of sharply decreasing inequality, which is supported by a reduction of 7.6 percent in the interdecile differential, 2.2 percent in SDL and 3-4 percent in measures very sensitive to the bottom of the distribution. However, the ATK 0.5, THEIL, CV and RMD measures tell us that inequality was actually growing, while the Gini coefficient does not change at all. Thus, the reduction in inequality in this year was probably concentrated in the very low end of the earnings distribution.

Even if all inequality measures point into the same direction, though, the question of by how much inequality has changed between two periods may be difficult to answer. At some points in the discussion of the previous paragraph, I have already touched briefly upon the issue of cardinal differences, and it is already apparent from Figures 6a and 6b that the measured magnitudes of changes are widely dispersed. The standard deviation of log earnings of women, for example, grew by "only" 63.5 percent between 1986 and 2000, while the Theil index more than tripled. After 2000, the 90-10 differential (and some other measures) documents a sharp decline of 16 percent within two years, but the ATK 0.5 measure decreases only by half that much and the RMD by only a little more than two percent. In Table 2, I exploited the sensitivity properties of these measures, and computed arithmetic averages of their yearly changes by grouping them according to their different reactions to redistribution in different parts of the earnings distribution. ENT and ATK 2 are used to get to the values in the "Bottom" columns; MLD, RMD and ATK 1 are the measures sensitive to changes around the mean; and the "Top" columns include averages of changes in THEIL and ATK 0.5. I omitted the Gini coefficient, because as we saw on Figure 1, the mode (or one of the two modes) is not necessarily in the middle of the distribution in later years. The CV was also dropped from this analysis due to its odd behavior compared to other measures, and its lack of any particular direction of sensitivity. It is important to note that if earnings only grow, say, in the top end of the distribution, then measures sensitive to the bottom end will also display an increase in inequality, but measures particularly responsive to the top will show an even larger rate of growth. Thus, it is the relative magnitude of numbers across columns in a given row of Table 2 that is giving us hints about the nature of inequality changes.

Until 1994, we can observe a pervasive inequality growth, affecting all parts of the distribution of both genders. The only exception for women is 1994, when measures responsive to top-end changes show an attenuated level of inequality. Besides similarities across genders in that inequality growth seems to be driven by changes in both ends of the distribution – to a somewhat greater extent in the top end –, there are differences in the timing of growth. The biggest rates for women are recorded in 1992 and 1993; while for men in 1992 and 1994, 1993 is a year of modest inequality growth for them. 1995 is a year of decreasing inequality for both groups: In case of women, equalizing processes seem to be rather concentrated in the bottom end; while it is the top end that is primarily affected in case of men. In the second half of the nineties, the growth paths are very similar (except for some minor differences in timing), and changes cover the whole distribution with some bias towards the most skilled. As already discussed, the sudden fall in inequality in 2001 affects mostly the least skilled. For men, only measures sensitive to the bottom signal a further drop in inequality in the next year, while the results for women are unambiguous. The last three rows of Table 2 are particularly interesting: As inequality returns to its growth path, first, changes occur in the lower half, but are then gradually shifted to the upper part of the distribution. A possible explanation for this phenomenon may be the return of the labor market to the wage differentials that prevailed before the minimum wage increase. As the real value of the minimum wage rises, wage differentials between minimum wage and non-minimum wage jobs narrow, hence, all things being equal, wages have to grow in order to get back to the old premia paid due to higher skills, compensating differentials, discrimination, efficiency incentives, union pressure, etc. Since wage difference between jobs closer to each other in the distribution are more visible to both employers and employees, the increase in minimum wage may first push wages in the bottom end higher, and then the effect triggers a chain reaction up the distribution.

Up to this point, I have analyzed the distribution of real monthly earnings. Starting in 2002, employers were asked to provide information on working hours of their employees. The HWES contains two variables measuring working hours: the first captures reported weekly regular working hours; while the second counts hours actually paid in the month of the survey. Although the HWES provides no information on hourly wages, I am able to construct two proxies using the aforementioned two measures. Since it is possible that workers who earn more over the month also work more hours, real hourly wages would be the appropriate level of aggregation to measure inequality. Figures 7a, 7b and 7c compare inequality measures computed for the distribution of real monthly earnings and of the proxy hourly wage rates based on weekly regular and on monthly paid hours, for the last four years of the sample when working hour information is available.

Figure 7a displays the evolution of interdecile differentials (measured in logs). Inequality in the top part of the distribution is almost exactly the same, no matter which measure of earnings we consider. The absolute difference between the 90-50 differential measured by hourly earnings and by the respective hourly earnings measure is less than 0.5 percent in every year, except for two years in the series of males where it is around 0.7 percent. We can only observe larger differences in the bottom part of the distribution. Inequality of hourly earnings according to weekly regular hours is virtually the same as inequality of monthly earnings, but dividing earnings by monthly paid hours generates a real difference. Figures based on this latter variable are lower on average by 3.7 percent in case of women and by 3 percent in case of men. These are significant discrepancies concerning that in this period the average growth rate of the 50-10 measure of real monthly earnings was 8.9 percent for the female distribution and 6.3 percent for men. About one third of the above difference carries over to the 90-10 differential, that is, the monthly-paid-hour-based proxy for the dispersion of real hourly earnings in the entire population is by one percent lower on average than the dispersion of monthly earnings. Since the interdecile differentials are independent of scale, if the majority of full-time workers report similar weekly regular hours, than dividing monthly earnings by this measure should not introduce large changes into measured dispersion. According to the evidence in Figure 7a, this seems to be the case; and we may also conclude that inequality of monthly earnings is to some extent an artifact of top-earners working many hours and low-earners working less, as supported by the series based on actual monthly paid hours, and this observation is especially pronounced for below-median employees.

Figures 7b and 7c repeat the exercise, but this time, for Atkinson measures with inequality aversion parameters 0.5, 1 and 2. Computed indices of inequality now differ even less than in the case of interdecile differentials, the absolute discrepancy between measures based on monthly and on proxy hourly earnings is mostly below half percent. The exception in case of women is 2005 when the hourly Atkinson values are lower by close to one percent than the ones relying on monthly earnings. The gap is widening as we move towards higher inequality aversion, and from regular weekly hours towards monthly paid hours. For men, differences are again negligible in most of the cases. In 2004 and 2005 however, the lower dashed line of the Atkinson index based on monthly paid hours with unit inequality aversion, and especially with an aversion parameter of two, deviates clearly from the solid line representing inequality of monthly earnings (and also from the upper dashed line plotting the level of inequality when dividing monthly earnings by weekly regular working hours). The difference in case of medium inequality aversion (ATK 1) is close to one percent, while for ATK 2 – that is especially sensitive to wage changes in the bottom part of the distribution – it is around 1.4 percent. So, the conclusion drawn from these (rather small) deviations is the same as in the previous paragraph: Controlling for actual working hours tends to reduce measured inequality of earnings primarily in the bottom part of the distribution, albeit only slightly.

5. Robustness of Inequality Changes: Composition Effects

So far, I computed inequality measures with the actual weights of the yearly samples and for yearly earnings distributions resulting from the actual skill composition of workers. However, during transition, the composition of the work force changed substantially. During the early years of the transition process, many workers dropped out of the labor market – many of them for good – due to inability to adjust their skills to changing market conditions and due to bad incentives (e.g. flawed unemployment insurance and social policies). Kertesi and Köllő (1997) estimated individual job loss probabilities for the years 1986-1996 based on the 1993 Labor Force Survey to approximate the selection bias in earnings inequality. Their results show that over this period, the distribution of job loss propensities shifted towards workers with low transition probabilities to unemployment. Köllő (2005) also showed that most workers who lost their jobs during this time became permanently inactive afterwards. Other workers managed to update their skills during a temporary state of inactivity by taking part in education and training, which also affected composition from two sides: by the transitory inactivity and by changing the educational composition after returning to the labor market. The ratio of people with high school completed and with higher education, and the ratio of females increased during the sample period, and population aging also had an effect on the age and experience distribution of the work force.

Lemieux (2006) demonstrated how the increasing share of worker groups with large wage dispersion can bias inference with respect to the causes of increasing residual wage inequality. In this paper, I only analyze inequality in general without paying attention to residual dispersion in real earnings; however, I applied the reweighting method Lemieux uses to control for composition. The method was first proposed by DiNardo et al. (1996) and it serves to create counterfactual wage distributions by holding the skill composition of the work force constant.[12] First, between any two given year, t and t-k, I ran a logit regression on the pooled sample of the two years, by including education and experience variables on the right hand side and with an indicator variable of being present in the sample of, say, year t. Then, I predicted the probability

[12] In the Eastern European context, Ganguli and Terrell (2006) applied the method of DiNardo et al. (1996) to construct counterfactual kernel densities of wages for Ukraine and to do a Lemieux (2002) and Juhn et al. (1993) type of decomposition analysis..

of being sampled in t conditional on the above individual characteristics. After having the predicted values, $\hat{p}$, the new sample weights for year t-k are calculated as

$$W_{t-k} = \frac{\hat{p}}{1-\hat{p}} w_{t-k},$$

where W_{t-k} represents the new set of weights and w_{t-k} the old one; and $W_t = w_t$. Of course, we can hold the skill composition of year t-k constant by switching t and t-k in the above description. After having constructed the counterfactual skill distributions, one can compute inequality indices based on the new and old set of weights, any difference between the two series measures to which extent composition effects influence the dispersion of earnings.

Figure 8 demonstrates the results of this exercise for the standard deviation of log monthly earnings (SDL). The evolution of the actual standard deviation is familiar since it is very similar to that of the 90-10 differential as displayed in Figure 3a. When we freeze the skill composition in its 1986 or 2005 structure, a rather different picture emerges, especially in case of women. While the actual SDL increased by 60 percent between 1986 and 2005, the counterfactual dispersion grew by only 27 percent, holding the 1986 skill distribution constant (the figure is 40 percent when keeping the distribution of skills at the 2005 level). It means that one-third to half of the growth in overall inequality is due to composition effects. For men, changing composition plays a less significant though not negligible role: against a 60 percent growth in actual standard deviation we are left with a 38 percent change when fixing the 1986 skill structure, and with 54 percent when sticking to the end-of-period composition.

Figure 8 is also instructive concerning the dynamics of composition effects, albeit only visually. Judging from the slopes of the three curves, we may suppose that the importance of changes in composition is different across sub-periods. Tables 3a and 3b quantify this relationship by calculating "rolling-base" counterfactual standard deviations. For each sub-period, the skill distribution is fixed at its beginning-of-period level in order to have a common base for the actual and counterfactual SDL. The first two columns list the resulting changes in the standard deviation over the given sub-period. We see that periods when the numbers void of composition effects are lower than the actual change alternate with periods when the realignment in skill composition actually alleviated the rise in inequality (i.e. the change in the counterfactual SDL is higher than in the actual). To what extent composition effects matter exactly is shown in the third column of the tables, which contains the deviation of the counterfactual change from the actual in percentages. The closer the numbers are to zero, the less important are composition effects in the given period. Positive figures mean that composition effects were proeffective, that is, they intensified the effects of other factors explaining inequality changes. Negative values on the other hand represent countereffective composition effects that either mitigate a rise (printed in italic) or a fall (printed in bold) in inequality.

We see that composition effects do always matter: even in periods when they are the least important, their contribution to the explanation of the variation in earnings dispersion is between ten and twenty-five percent. Except for 1993, composition effects seem to add to the growth in inequality during the early years of transition before 1994. This is an important period, since 88 and 65 percent of the total change in SDL is concentrated to these eight years, for women and men respectively. When being proeffective, the ratio explained by composition effects ranges from one-fifth (1986-1992) to four-fifths (1993-1994) for females, and from one-fifth to one-half for males. Skill composition had a huge countereffective impact on male inequality in 1993. According to Köllő (2005), it has been the least-skilled male working population that has been most severely affected by job losses and permanent inactivity, which indeed may have had a tempering effect on earnings inequality growth. The yearly rate of inequality growth that would have prevailed in the absence of composition effects is rather large: it is equivalent to an 0.09 six-year change which is comparable in magnitude (and in absolute terms) to the 0.1 rise between 1986 and 1992. This countereffective pattern repeats for the years 1995 and 1996 to about the same extent for men and on a smaller scale for women. The question why this phenomenon is concentrated to those three years only needs further investigation beyond the scope of this paper.

In the second half of the nineties, we observe again a significant rise in earnings dispersion, and composition effects account for around twenty percent of it; so in these terms, the nature of inequality change is similar to that of the 1986-1992 period. Following the rise in the minimum wage in 2000 and 2001, both the actual and the counterfactual standard deviations decline as expected, but surprisingly, composition effects have a countereffective impact on inequality, and the magnitude of the effect is about twice as large for women as for men. In their paper about the employment effects of the minimum wage regulation, Kertesi and Köllő (2004) show that it was workers of firms with 5-20 employees and workers paid around the minimum wage who suffered the most from an increased propensity to get unemployed. Intuitively, this should work in the opposite direction to what we observe in Tables 3a and 3b. I will not try to resolve the puzzle in this paper.

More interestingly, once the trend in earnings inequality turns from a descent to a climb in 2002, the nature of composition effects changes as well. Compared to the first sub-period of rising earnings dispersion, the explanatory power of composition doubles for males and quadruples for females. (The growth rate of inequality is at the same time smaller than during the beginning of the sample period: it is one-third of the 1986-1992 pace in case of women and three-quarters for men.) To sum up, composition effects do matter, and accounting for them alters the conclusions drawn concerning the evolution of real earnings inequality in Hungary. Figure 8 and Tables 3a and 3b are suggestive, but only a careful decomposition analysis could shed light on the exact effects of a changing skill distribution.

6. Conclusions

TO BE ADDED

References

Autor, David H., Lawrence F. Katz, and Melissa S. Kearney, "Measuring and Interpreting Trends in Economic Inequality." *AEA Papers and Proceedings*, Vol. 96(2), 189-194, May 2006.

Campos, Nauro and Jolliffe, Dean. "Does Market Liberalization Reduce Gender Discrimination? Econometric Evidence from Hungary 1986-1998." William Davidson Working Paper Institute Working Paper No. 678, 2004.

Cowell, F. A., *Measuring Inequality*. Oxford: Philip Allan Ltd., 1977.

DiNardo, John, Nicole M. Fortin, and Thomas Lemieux, "Labor Market Institutions and the Distribution of Wages, 1973-1992: a Semiparametric Approach." *Econometrica*, Vol. 64(5), 1001-1044, September 1996.

Éltető, Ödön "A keresetek színvonala és szóródása az 1990-es években Magyarországon" *Statisztikai Szemle*,11, 1996.

Ganguli, Ina, and Katherine Terrell, "Institutions, Markets and Men's and Women's Wage Inequality: Evidence from Ukraine." *Journal of Comparative Economics*,34, 200-227, February 2006.

Juhn, Chinhui, Kevin M. Murphy and Brooks Pierce, "Wage Inequality and the Rise in returns to Skill." Journal of Political Economy, Vol. 101(3), 410-442, 1993.

Karoly, Lynn A., "Changes in the Distribution of Individual Earnings in the United States: 1967-1986." *Review of Economics and Statistics*, Vol. 74(1), 107-115, February 1992.

Katz, Lawrence F., and David H. Autor, *Changes in the wage Structure and Earnings Inequality.* Chapter 26., Handbook of Labor Economics, Amsterdam: Elsevier Science B. V., 1999.

Keane, Michael P., and Eswar S. Prasad, "Changes in the Structure of Earnings During the Polish Transition." *Journal of Development Economics* 80, 389-427, 2006.

Kertesi Gábor, and Köllő János, "Wage Inequality in East-Central Europe" Budapest Working Papers on the Labour Market 2000/7, October 2000.

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

Kertesi Gábor, and Köllő János, "A 2001. évi minimálbér-emelés foglalkoztatási következményei" *Közgazdasági Szemle*, Vol. 60, 293-324., April 2004.

Kertesi Gábor, and Köllő János, "Reálbérek és kereseti egyenlőtlenségek, 1986-1996." *Közgazdasági Szemle*, Vol. 44(4), 612-634, July-August 1997.

Kertesi Gábor, and Köllő János [2003a], Industrial Wage Differences in Hungary, Pt. 1 (In Hungarian: Ágazati bérkülönbségek Magyarországon 1. rész: Az ágazati járadékképződés alternatív modelljei) . *Közgazdasági Szemle*, Vol. 50(11), 923-938, 2003.

Kertesi Gábor, and Köllő János [2003b], Industrial Wage Differences in Hungary, Pt. 2 (In Hungarian: Ágazati bérkülönbségek Magyarországon 2. rész: Járadékokon való osztozkodás koncentrált ágazatokban, szakszervezeti aktívitás jelenlétében). *Közgazdasági Szemle*, Vol 50(12), 1049-1074, 2003.

Köllő János, "A foglalkoztatottak összetétele az ezredfordulón" Budapest Working Papers on the Labour Market 2005/2, 2005.

Köllő János, and Nagy Gyula, "Earnings Gains and Losses from Insured Unemployment in Hungary" *Labour Economics*, Vol. 3(3), 279-298, 1996.

Lemieux, Thomas, "Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising demand for Skill?." *The American economic Review*, Vol. 96(3), 461-498, June 2006.

Lemieux, Thomas, "The Changing Nature of Wage Inequality." NBER Working Paper No. 13523, October 2007.

Milanovic, Branko, "Explaining the Increase in Inequality during Transition." *Economics of Transition*, Vol. 7(2), 199-341, 1999.

OECD Policy Brief, Paris: OECD Publications, June 2007.

Pudney, Stephen, "Earnings Inequality in Hungary: A Comparative Analysis of Household and Enterprise Survey Data." *Economic Change and Restructuring*, Vol. 27(3), 251-276, 1994.

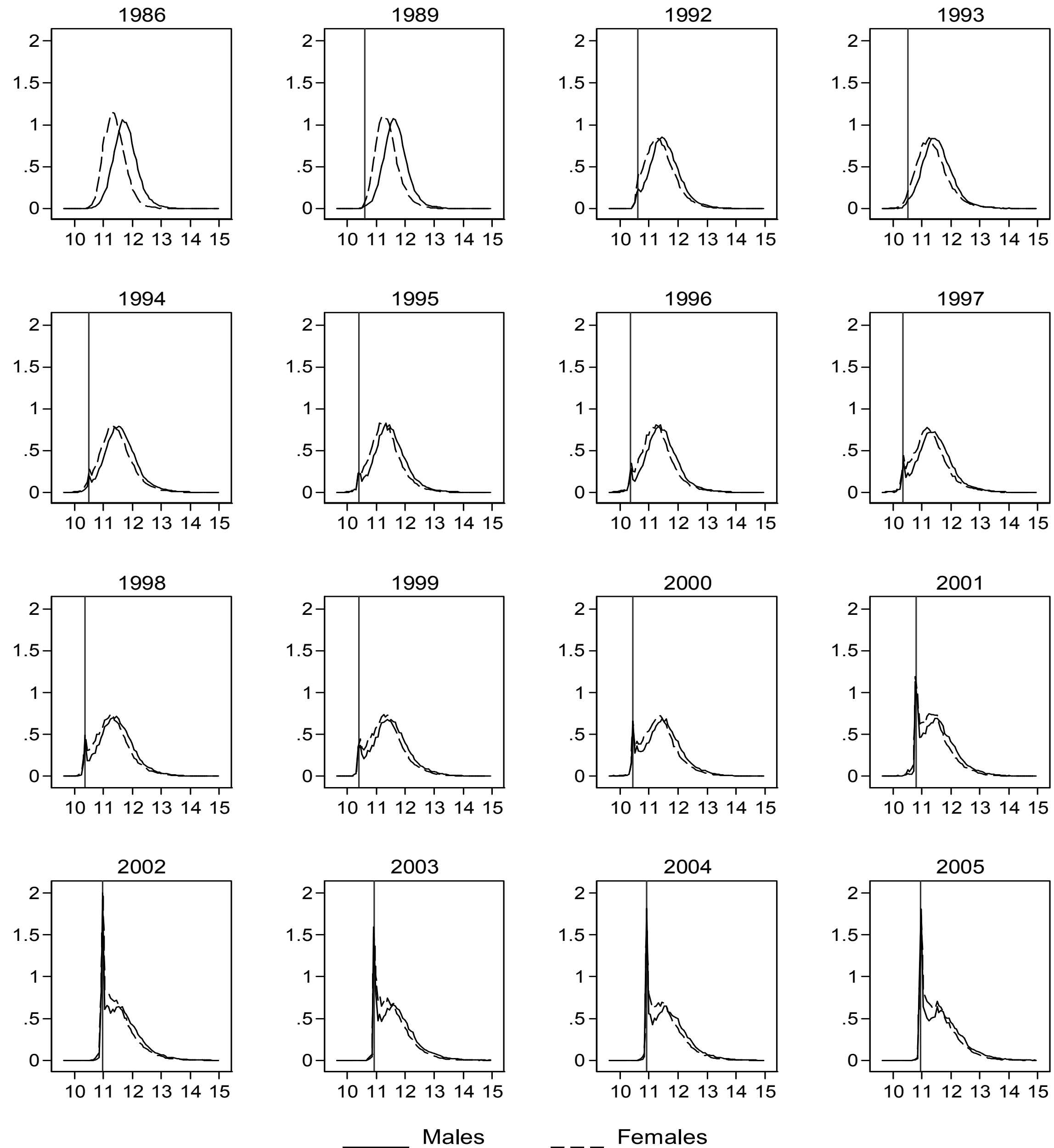
Rutkowski, Jan J., "Changes in the Wage Structure During Economic Transition in Central and Eastern Europe." World Bank Technical Paper, Washington D.C.: The World Bank, 1996.

Table 1: Sample Size by Year

Year	Unweighted: Workers Observed (thousands)	Unweighted: Firms Observed	Weighted: Workers Represented (thousands)	Weighted: Firms Represented
1986	88.9	3,242	2,435.2	9,823
1989	108.0	4,584	3,254.0	14,045
1992	91.1	6,534	2,129.4	10,409
1993	91.5	6,831	1,702.6	11,010
1994	103.5	8,491	1,701.6	11,973
1995	102.7	8,411	1,626.5	11,628
1996	98.2	7,793	1,593.1	18,130
1997	88.6	7,706	1,600.1	19,596
1998	96.2	7,452	1,641.9	21,121
1999	97.8	8,220	1,618.6	22,267
2000	116.1	10,551	1,755.4	42,810
2001	117.4	11,111	1,757.3	45,474
2002	122.7	8,822	1,754.4	78,005
2003	119.3	8,662	1,757.2	81,738
2004	134.6	9,732	1,808.4	104,820
2005	146.5	10,151	1,828.3	85,229

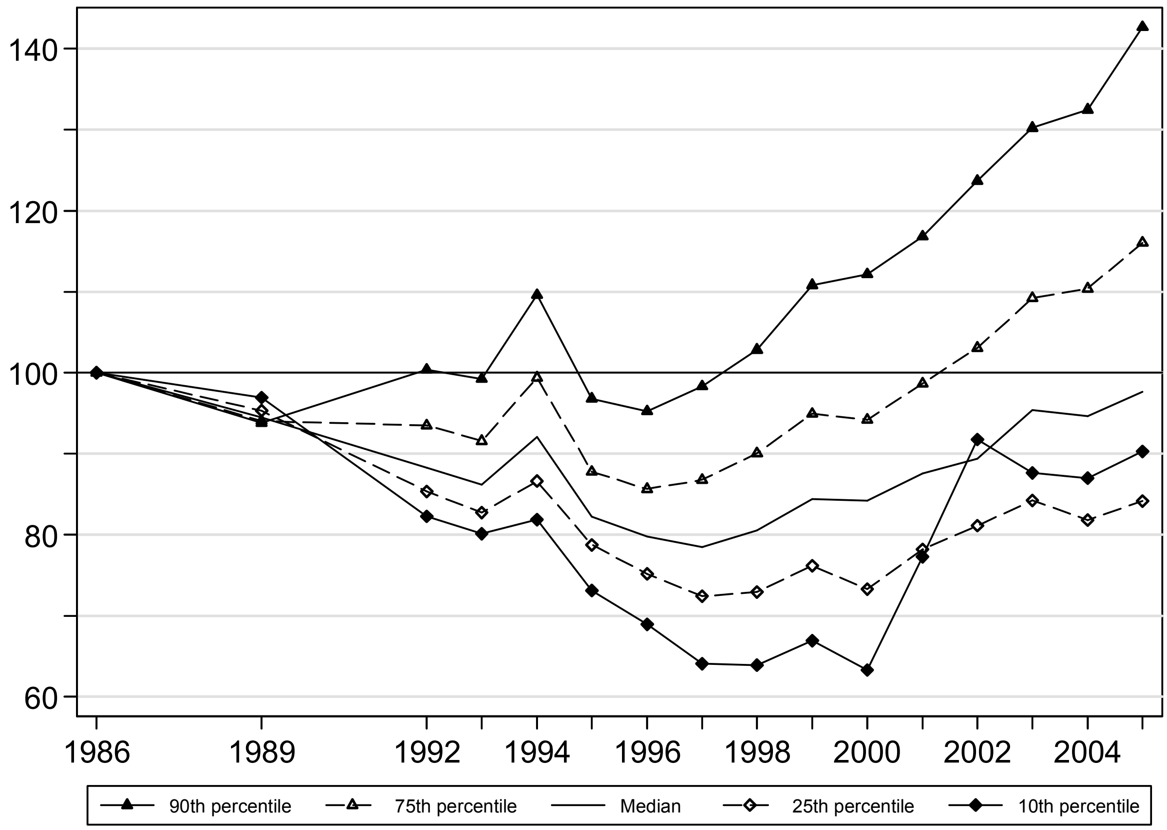
Notes:

Figure 1: Kernel Density Estimates of Log Real Earnings by Gender, 1986-2005



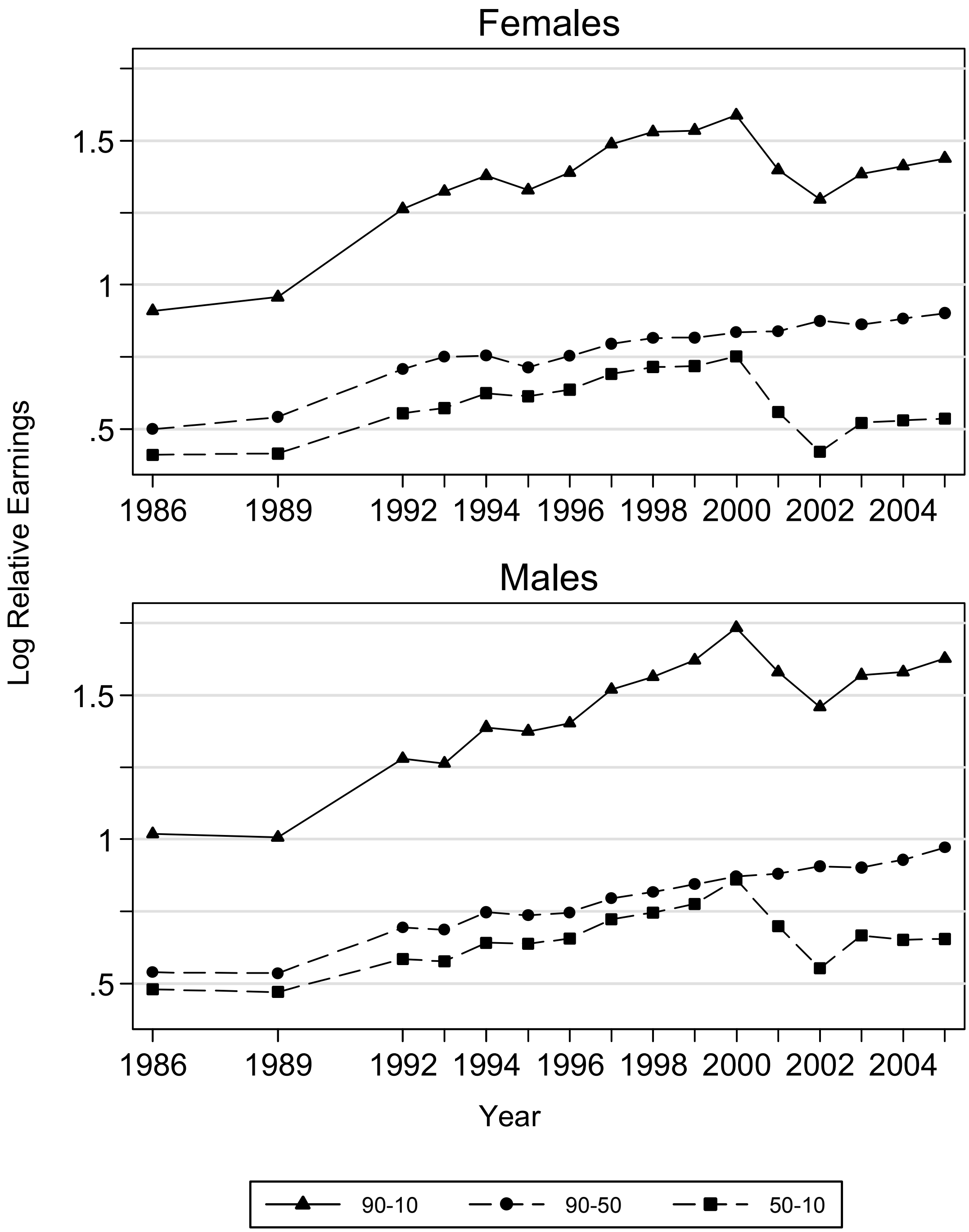
Notes: Log real earnings are expressed as the logarithm of gross monthly earnings in 2005 Hungarian forints. Earnings refer to all payments defined as earnings by the Hungarian Statistical Office, received by the employee in the month of the survey at the expense of the employer's wage cost account, plus one-twelfth of all non-regular premia, bonuses, commissions and thirteenth-month salary earned in the previous year. Results are weighted.

Figure 2: Real Monthly Earnings by Percentile
(1986=100)



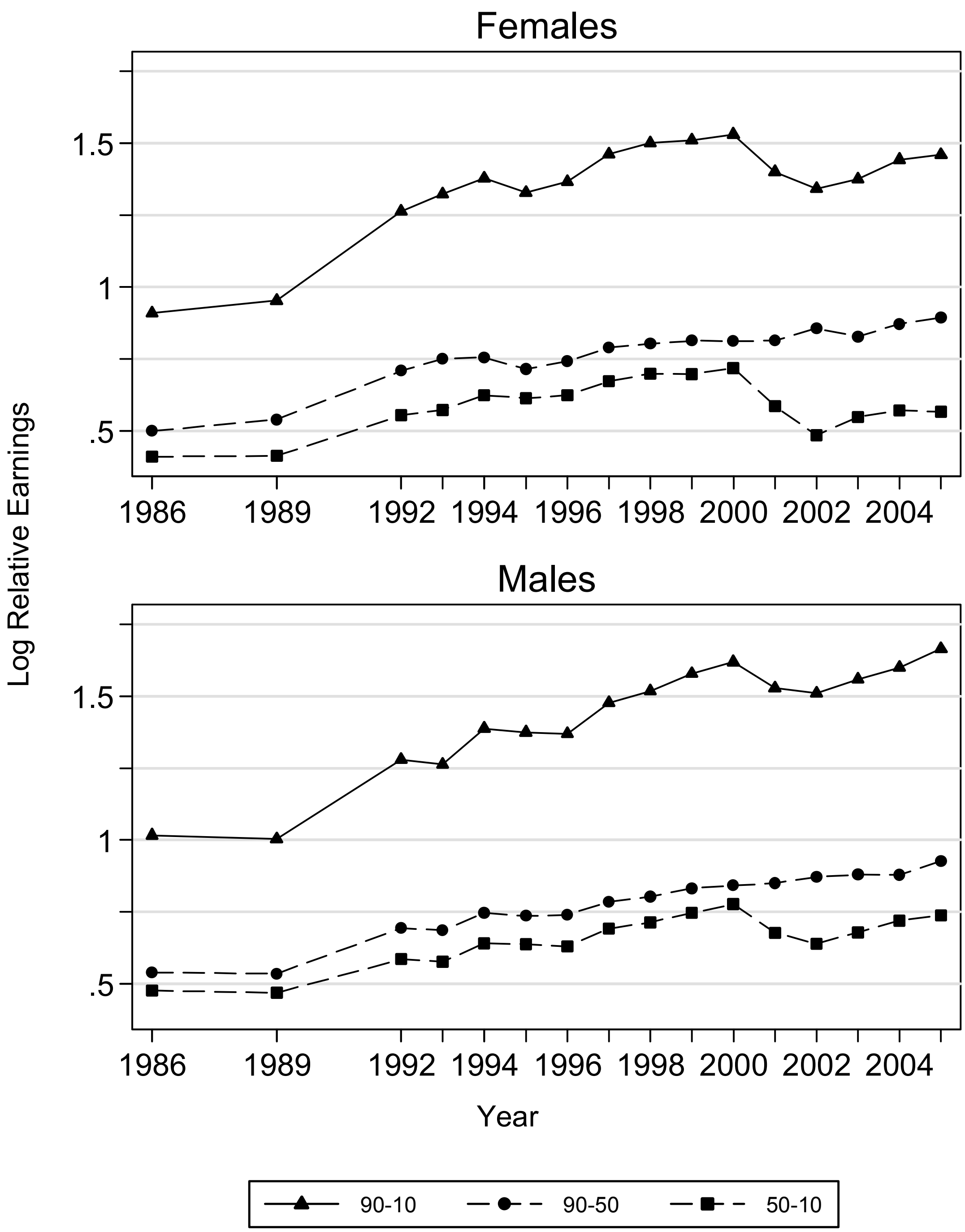
Notes: Log real earnings are expressed as the logarithm of gross monthly earnings in 2005 Hungarian forints. Earnings refer to all payments defined as earnings by the Hungarian Statistical Office, received by the employee in the month of the survey at the expense of the employer's wage cost account, plus one-twelfth of all non-regular premia, bonuses, commissions and thirteenth-month salary earned in the previous year.

Figure 3a: Interdecile Differentials by Gender, 1986-2005



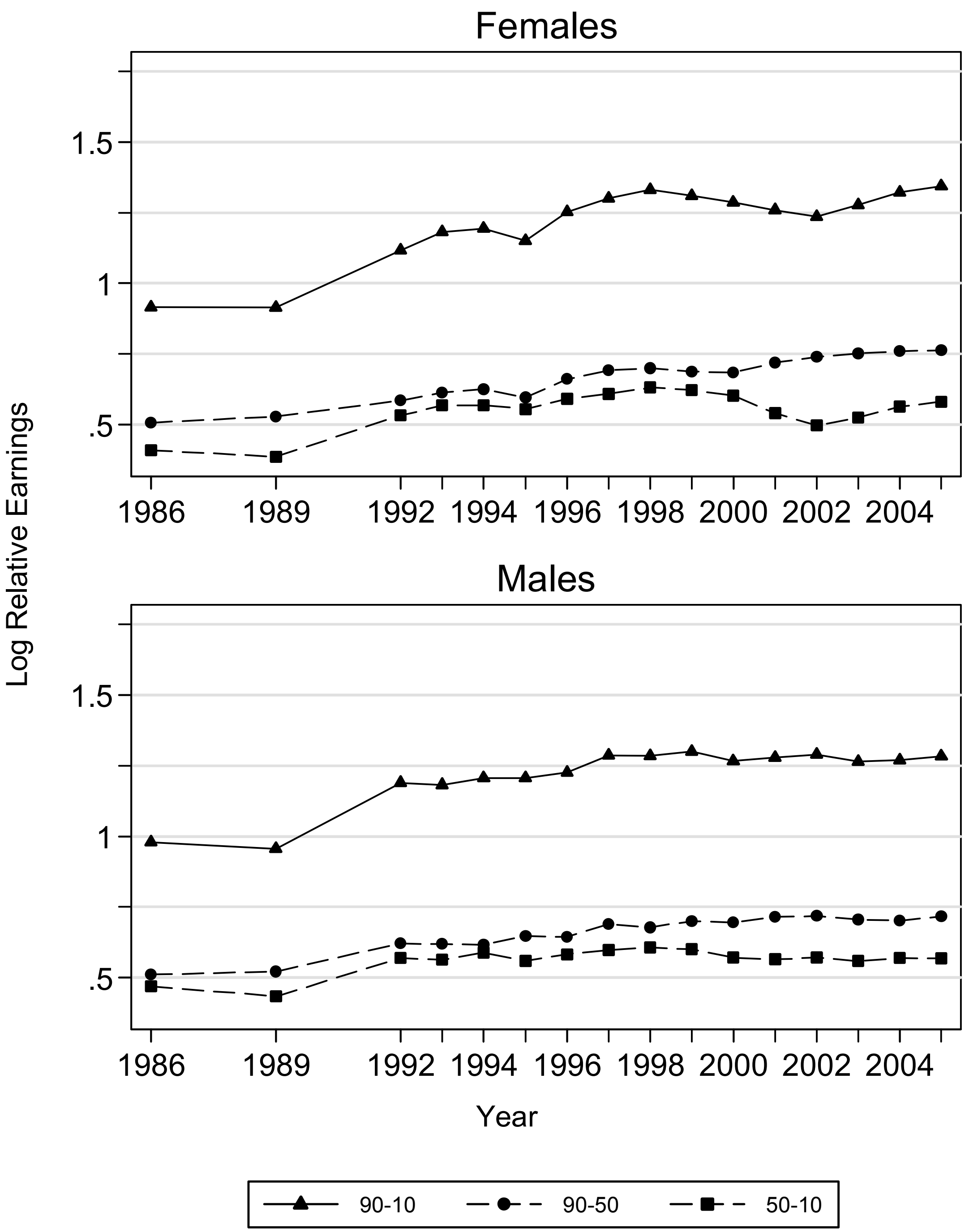
Notes: See Figure 2.

Figure 3b: Interdecile Differentials on a Sample of Enterprises with more than 20 Employees Only, 1986-2005



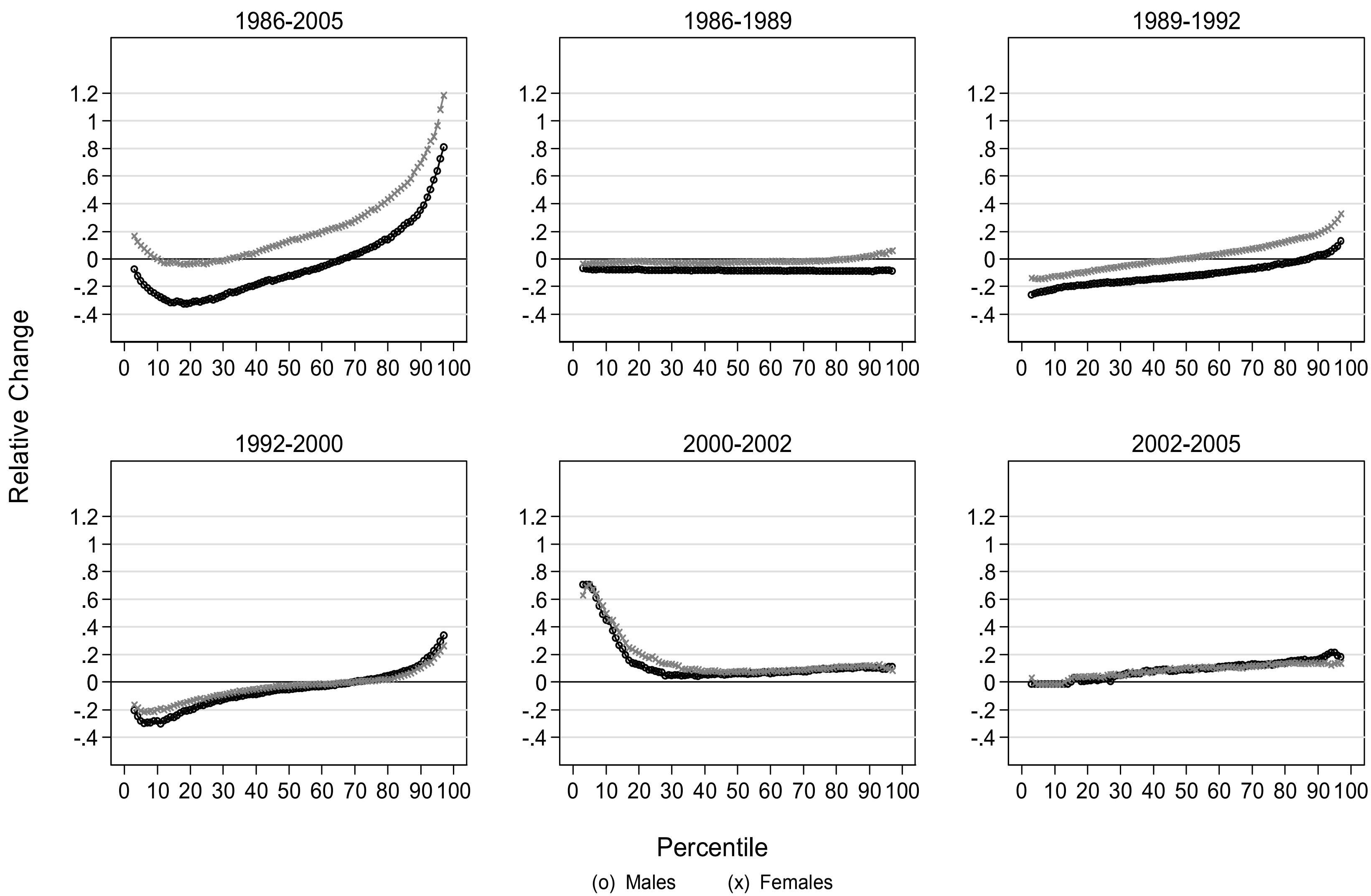
Notes: See Figure 3.

Figure 3c: Interdecile Differentials on a Sample of Enterprises with at least 14 Years of Existence Only, 1986-2005



Notes: See Figure 2.

Figure 4: Changes in Real Earnings by Percentile and Gender, 1986-2005



Notes: See Figure 2.

Figure 5: Changes in Real Earnings by Percentile, Men

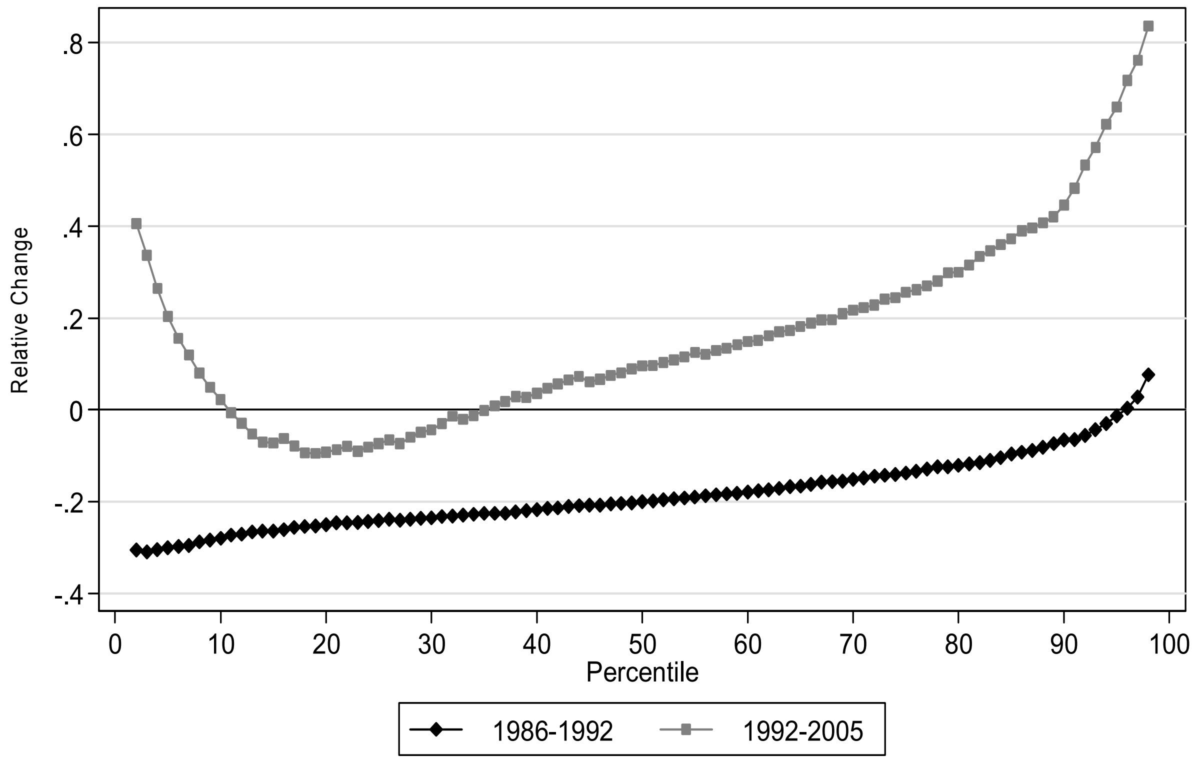


Figure 6a: Comparison of Inequality Measures, Females
(1986=100)

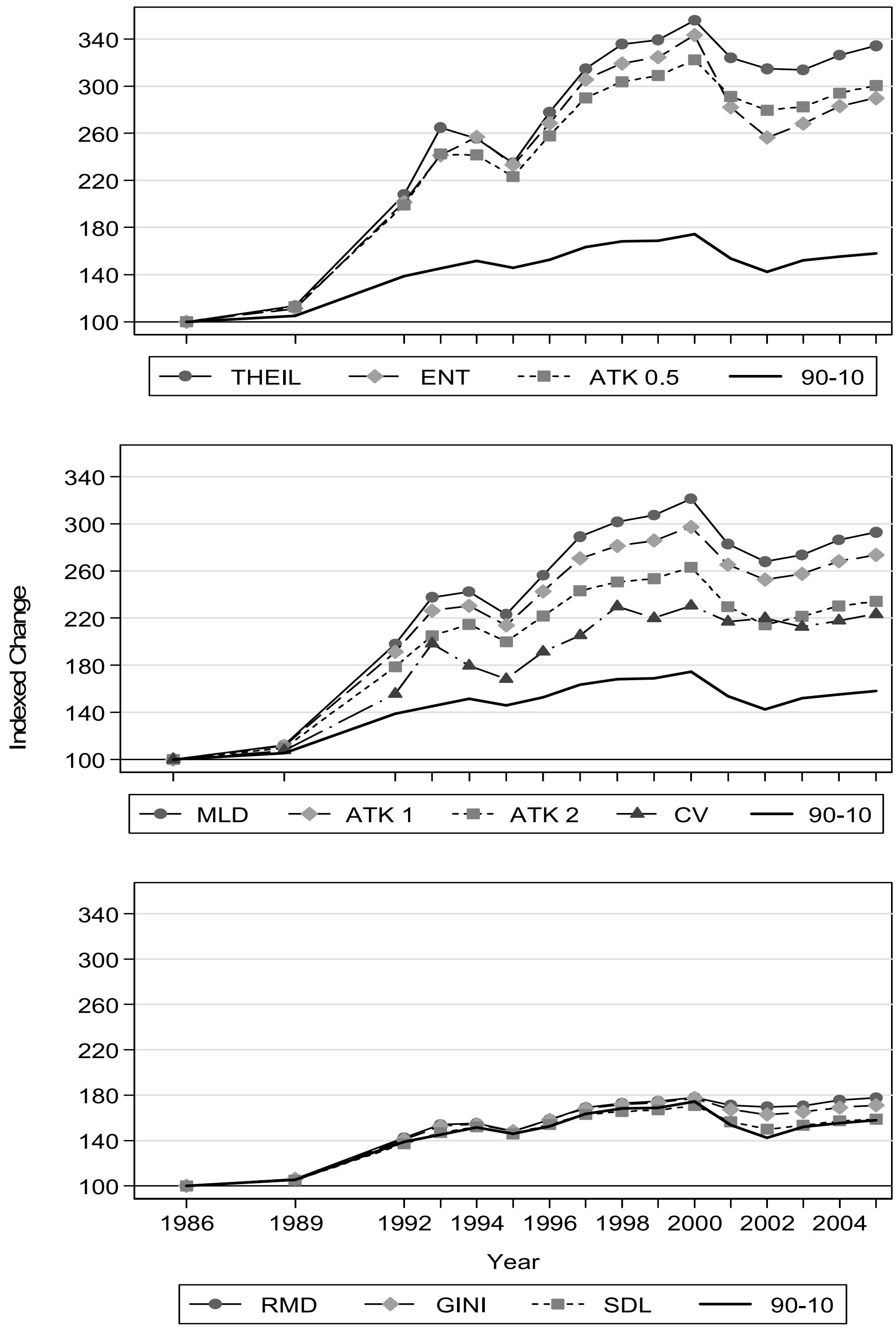


Figure 6b: Comparison of Inequality Measures, Males
(1986=100)

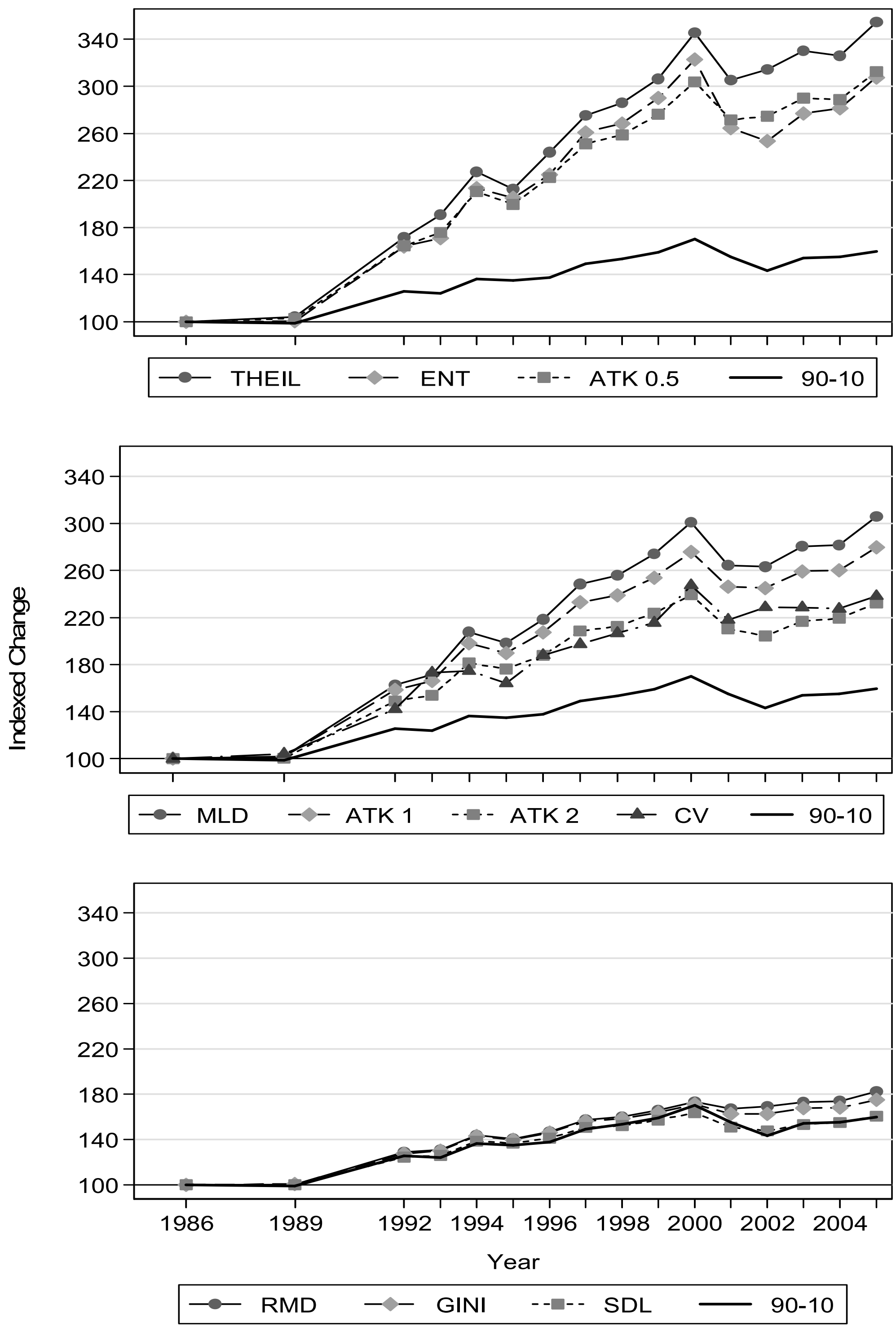


Table 2: Changes in Earnings Inequality by Sensitivity Groups of Inequality Measures (Percentage change between t-1 and t)[a]

	Women			Men		
Year	**Bottom**	**Middle**	**Top**	**Bottom**	**Middle**	**Top**
1989	10.5	10.0	13.1	0.4	1.3	3.4
1992	72.0	60.7	79.8	55.9	48.0	62.4
1993	17.3	15.5	24.6	3.8	4.0	9.0
1994	5.6	1.5	-1.9	21.3	16.8	19.5
1995	-8.1	-6.5	-7.9	-3.3	-3.6	-5.8
1996	13.1	11.7	16.9	8.1	7.9	13.1
1997	11.8	10.4	12.9	13.5	11.1	12.8
1998	3.7	3.4	5.6	2.4	2.4	3.5
1999	1.4	1.5	1.4	6.7	5.7	6.9
2000	4.9	3.5	4.6	9.1	7.6	11.4
2001	-15.3	-8.8	-9.3	-15.1	-8.8	-11.1
2002	-7.9	-3.6	-3.4	-3.5	0.2	2.0
2003	3.9	1.5	0.4	7.7	4.8	5.3
2004	4.8	3.9	4.0	1.4	0.5	-0.8
2005	2.0	1.8	2.3	7.6	6.9	8.5

[a] 1989 and 1992 figures refer to changes over three years.

Notes: Table values are unweighted averages of period-by-period changes of ENT and ATK 2 in the "Bottom" column; of MLD, RMD and ATK 1 in the "Middle" column; and of THEIL and ATK 0.5 in the "Top" column.

Figure 7a: Dispersion of Monthly and Hourly Earnings, 2002-2005
Interdecile Differentials
(From top to bottom: 90-10, 90-50 and 50-10 differential)

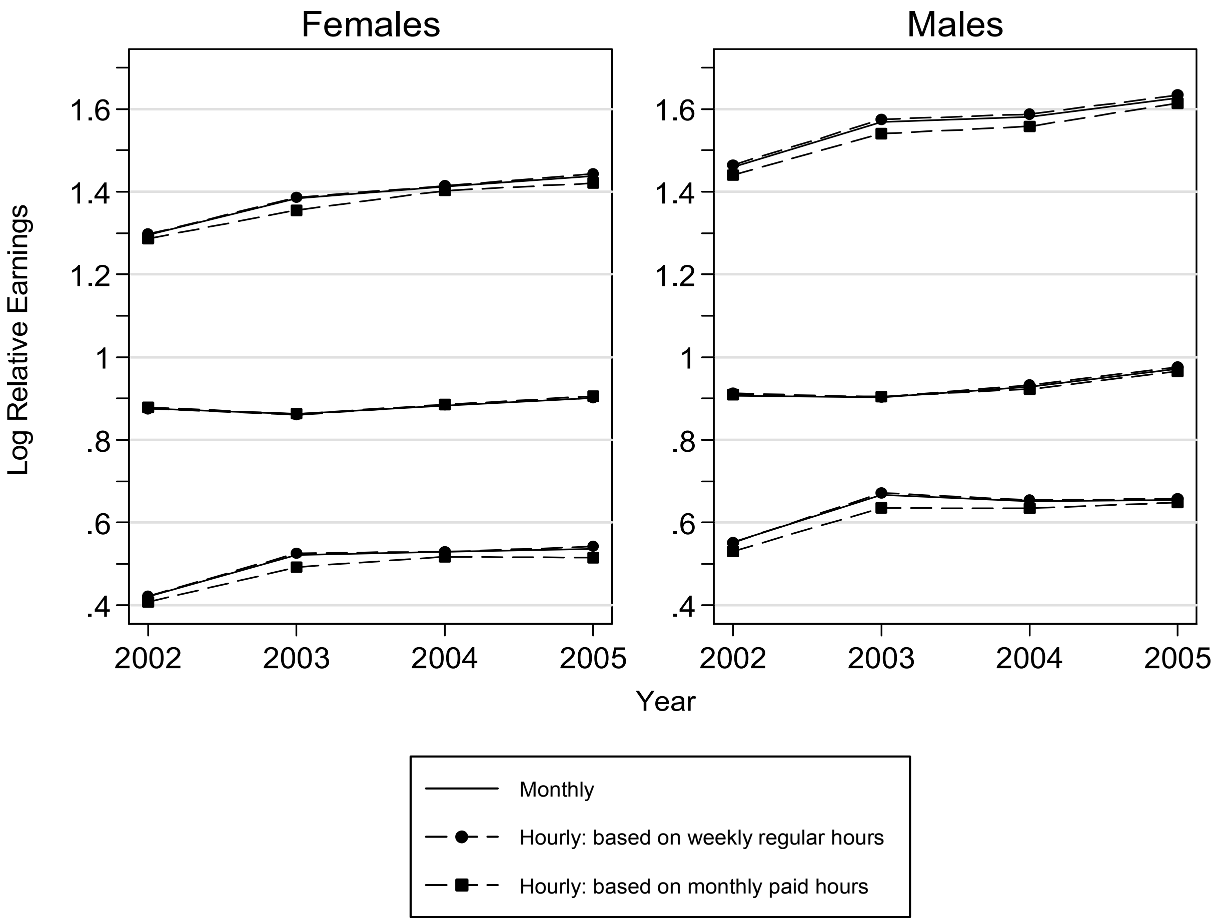


Figure 7b: Dispersion of Monthly and Hourly Earnings, 2002-2005
Atkinson Measures, Females

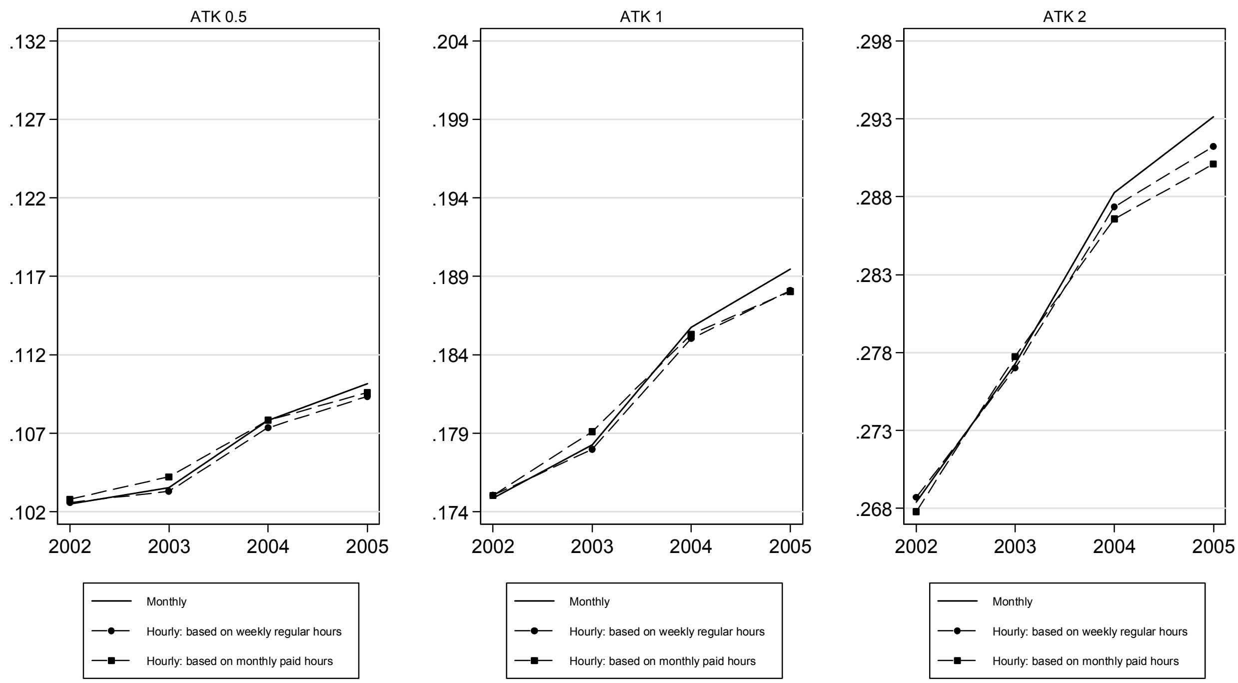


Figure 7c: Dispersion of Monthly and Hourly Earnings, 2002-2005
Atkinson Measures, Males

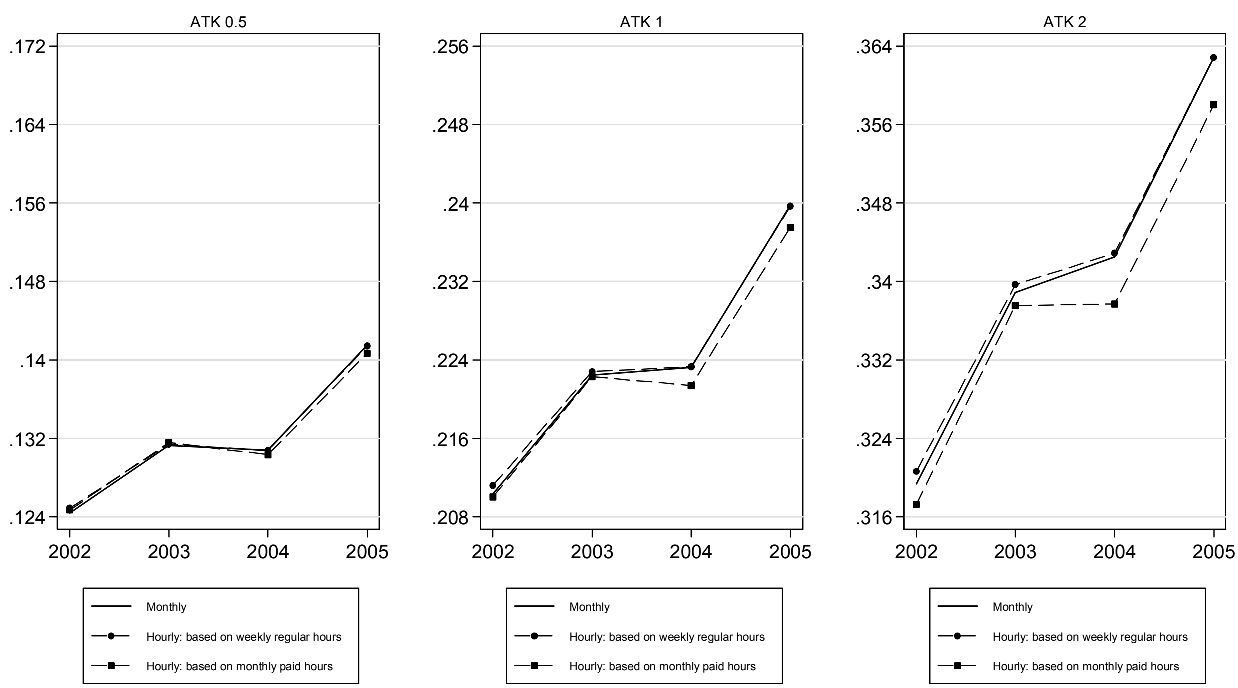


Figure 8: Evolution of the Actual and the Counterfactual Standard Deviation of Log Earnings, 1986-2005

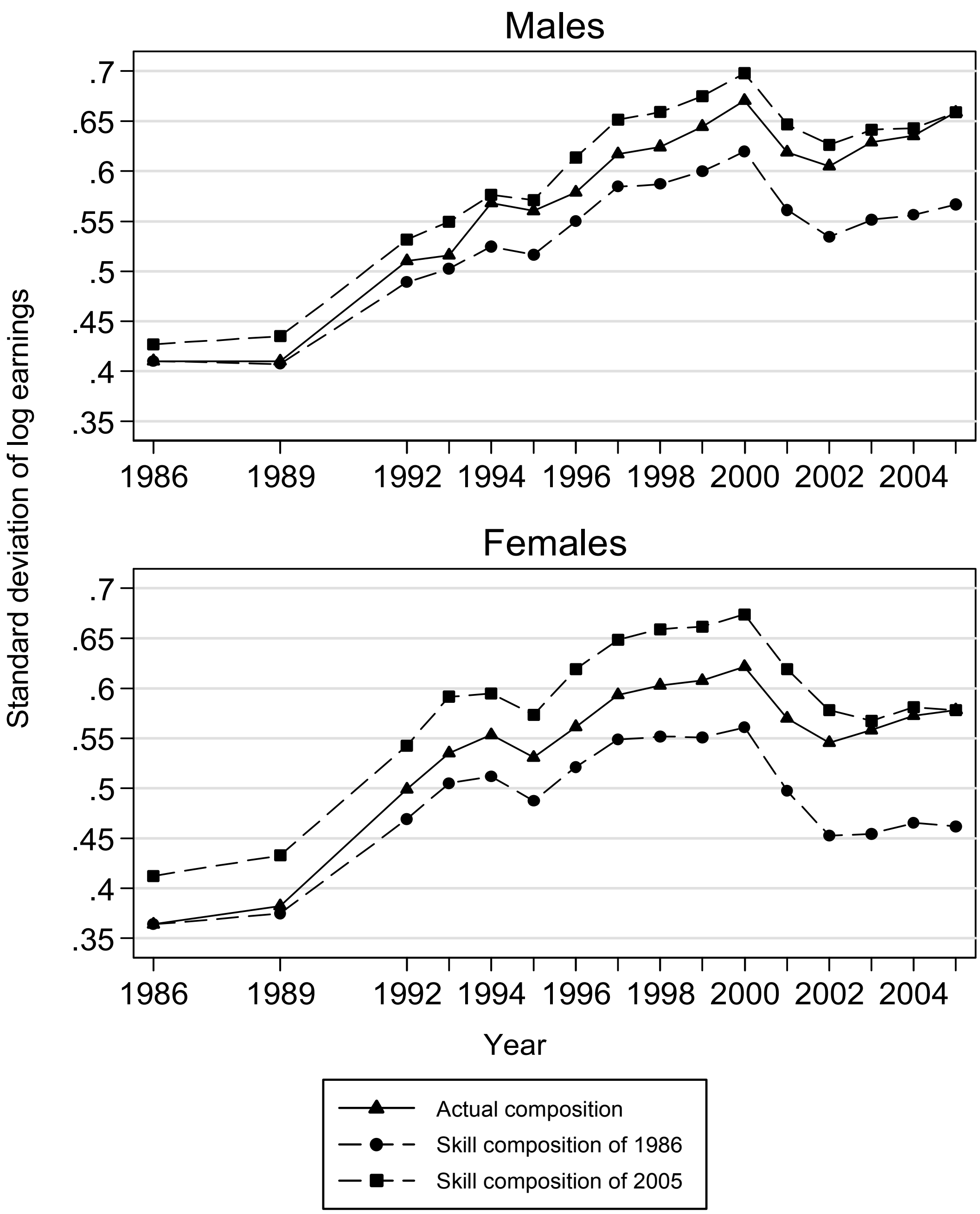


Table 3a: Change in the Standard Deviation of Log Earnings and Percent Explained by Composition Effects
Females

	Change in SDL		Percent Explained by Composition Effects
	Actual	Controlling for Composition	
1986-1992	0.135	0.105	22
1992-1993	0.036	0.040	*-12*
1993-1994	0.018	0.004	77
1994-1996	0.008	0.014	*-77*
1996-2000	0.060	0.046	24
2000-2002	-0.076	-0.099	**-31**
2002-2005	0.033	0.005	86

Table 3b: Change in the Standard Deviation of Log Earnings and Percent Explained by Composition Effects
Males

	Change in SDL		Percent Explained by Composition Effects
	Actual	Controlling for Composition	
1986-1992	0.100	0.079	21
1992-1993	0.006	0.015	*-166*
1993-1994	0.052	0.024	55
1994-1996	0.011	0.033	*-211*
1996-2000	0.092	0.076	17
2000-2002	-0.066	-0.075	**-14**
2002-2005	0.054	0.033	39