

Rising Earnings Inequality in Sweden: The Role of Composition and Prices

David Domeij*

Stockholm School of Economics

SSE/EFI Working Papers in Economics and Finance

No. 639

October 4, 2006

Abstract

This paper decomposes the rise in cross-sectional earnings inequality in Sweden between 1990 and 2002 into changes in market prices of observable characteristics, changes in the composition of the labor force across demographic groups and industries, and changes in unobservables, and compares the Swedish experience with that in the U.S. The rise in earnings inequality is in both countries a consequence of rising upper tail dispersion. Contrary to the U.S. experience, where the rise is largely driven by changing market prices of observables and increased residual dispersion, shifts in the Swedish labor force composition have contributed positively to the rise in the P90-P50 gap. The rise in the Swedish P99-P90 gap is however entirely accounted for by changes in prices and residual dispersion.

Key words: Earnings, inequality, Sweden, United States.

JEL-classification: E24, J31

*Correspondence: Department of Economics, Stockholm School of Economics, Box 6501, SE-113 83 Stockholm, Sweden. E-mail: David.Domeij@hhs.se. Financial support from The Jan Wallander and Tom Hedelius Foundation is gratefully acknowledged.

1 Introduction

An extensive literature has documented the increase in U.S. earnings inequality in the last three decades. Important factors behind this increase are price effects associated with schooling and experience; in particular the college-high school premium increased substantially since the late 1970's. However, residual inequality - dispersion within well defined narrow demographic groups, is generally believed to account for most of the growth in overall inequality.¹ In a recent paper, Lemieux (2005) shows that compositional effects, i.e. changes in the distribution across demographic groups, have become more important over time and that a large part of the change in residual inequality in the 1990's is a consequence of a more educated and a more experienced population.

The reason, going back to Mincer (1974), is that earnings trajectories become more dispersed as workers gain more experience. Holding prices constant, compositional changes affects inequality by changing the employment shares of worker groups that have more or less dispersed earnings. Consequently, changes in the distribution of experience may mechanically cause changes in inequality. As wage dispersion differs across gender, educational levels and industries, similar effect occur when the share of women in the workforce changes, the population becomes better educated and when some industries contract and others expand.

In this paper I investigate changes in earnings inequality in Sweden between 1990 and 2002, and compare these changes with those in the U.S. Sweden is an interesting case as the earnings distribution in the two countries was rather different in 1990. After two decades of extreme wage compression both the Gini coefficient for family income and the college-high school premium had fallen by more than 30 percent in Sweden. During the same period, the U.S. equivalents increased by 8-9 percent.²

In the early 1990's, however, Sweden experienced a dramatic change in

¹See Juhn, Murphy and Pierce (1993) for the 1970's and 1980's, and Acemoglu (2002) and Katz and Autor (1999) for the 1990's

²See Johansson (2006) for data on Swedish Gini coefficients, Gottschalk and Danziger (2003) for data on U.S. Gini coefficients, and Domeij and Ljungqvist (2006) for data on the skill premium.

the macroeconomic climate; between 1990 and 1993 GDP fell by 10 percent and relative to 1970, GDP remained below trend from 1991 through 1998.³ From the mid-90's, however, Sweden experienced strong GDP growth and by 2000 GDP was approximately 3 percent above trend.⁴

In the 1990's, Sweden also experienced a trend reversal in cross-sectional inequality.⁵ Figure 1 graphs the evolution of the median, the first percentile, the tenth percentile, the ninetieth percentile and the ninety-ninth percentile of the real earnings distribution for 1990-2002. All series are normalized to zero in 1990. The figure shows that all workers but those in the ninety-ninth percentile experienced a fall in real earnings during the early 1990's. However, by 1997 earnings were larger than in 1990 across the distribution, and by 2002 earnings in the first percentile was 18 percent larger than in 1990. Figure 1 also suggests that while earnings growth was rather similar below the median, earnings growth was increasing in the percentile above the median. In the top percentile, earnings grew by 35 percent.⁶ Thus earnings became more dispersed in the upper half of the distribution, but remained essentially unchanged in the lower half.

Figure 2 graphs the change across the whole distribution over the entire period, and for the down-turn and the recovery of the Swedish economy separately. The Figure shows that earnings growth was rather similar for all percentiles below the 90th percentile during the recovery, and that the increase in inequality below the 90th percentile occurred during the down-turn. The increase in earnings dispersion at the top (the difference in earnings growth between the 90th and 99th percentile) was somewhat larger in the later period.

In the 1990's Sweden also experienced large changes across demographic groups in the working population. In particular, the share with at least some college education increased from 24 percent in 1990 to 34 percent in 2002. About half of these were college graduates (at least three years of university studies), and among college graduates the share of men fell from

³This is assuming a trend growth rate of 1.91 percent on an annual basis, which is what you find by imposing a linear trend on the data for 1970-2004. In 1990 Sweden was about 3 percent above trend, so that GDP in 1993 was 7 percent below trend

⁴Even though GDP growth slowed down during 2001-2003, GDP was still 2-3 percent above trend in 2005.

⁵See for example Edin and Fredriksson (2000), Gustavsson (forthcoming) and Johansson (2006).

⁶The top 0.1 quantile of the earnings distribution grew by more than 50 percent.

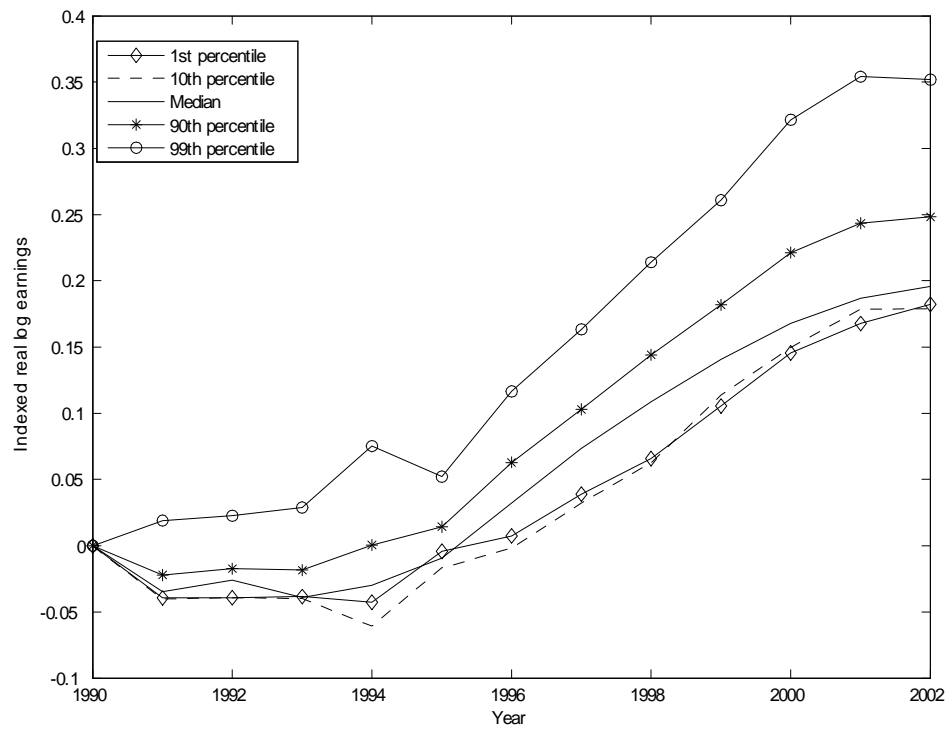


Figure 1: Indexed log real Swedish earnings by percentile, 1990-2002.

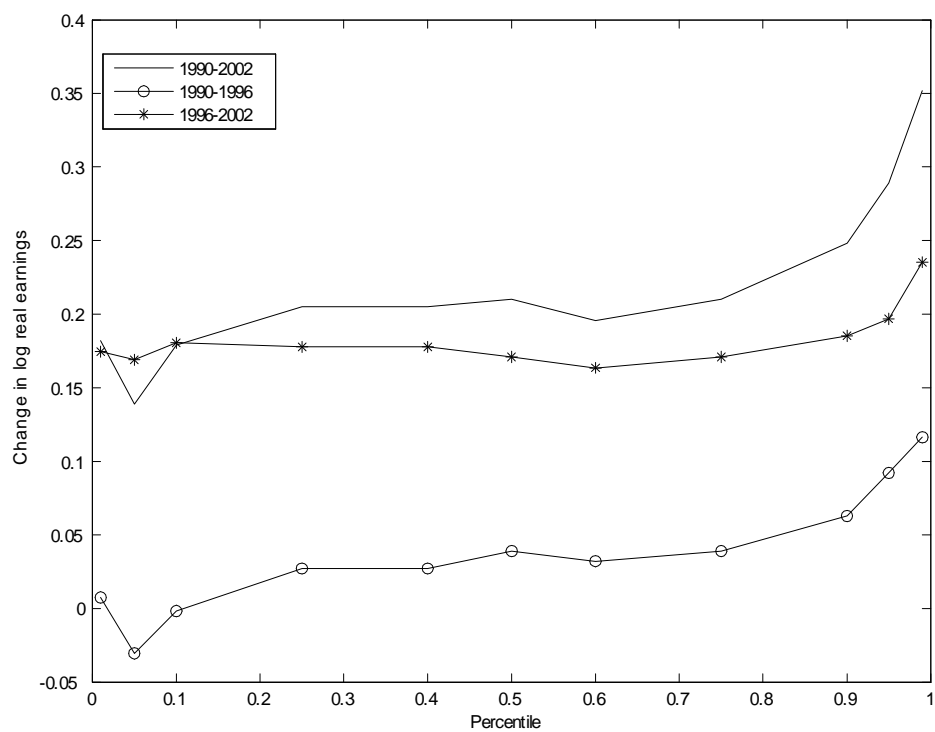


Figure 2: Log real Swedish earnings changes by percentile.

53 percent in 1990 to 45 percent in 2002. The working population also became more experienced. The share of workers below age 30 fell from 24 percent to 18 percent, and the share of workers above age 50 increased from 23 to 31 percent. There were also reallocations between industries from manufacturing towards services. This shift is by the popular press and by some authors (Bluestone and Harrison, 1988), characterized as a shift towards high-paid jobs in professional services and low-paid jobs in trade and other services, resulting in increased earnings dispersion.

Using a register database covering the whole Swedish working population, I show that these compositional changes account for almost 50 percent of the overall change in earnings dispersion. Changes in the education level and the industry composition are particularly important. I also find that the composition effect is the most important factor in understanding changes in the mid-range of the distribution, and in particular in the P90-P50 range. Price effects and residual dispersion on the other hand explain changes in the tails of the distribution. In the top tail these effects tend to re-enforce each other and the P99-P90 gap increased by more than 0.10 log-points. At the lower end, changes in unobservables instead more than cancel the price effects and the lower tail became more compressed.

Classifying individuals according to gender and education levels reveals that men experienced a much larger increase in earnings dispersion and that this difference is accounted for by a much larger increase in residual inequality. The difference is particularly large for college graduates and at the top end of the distribution; for men the P99-P90 ratio increased by 0.25 log-points but only by 0.08 log-points for women.

To compare the Swedish experience with the U.S. experience, I perform the same decomposition on U.S. earnings. The U.S. data is drawn from the *Current Population Survey* (CPS).⁷ Three observations stand out. First, the compositional effect was the most important factor in accounting for changes in the P90-50 gap in Swedish earnings, but in U.S. earnings it mainly affected dispersion below the median. Second, the increase in earnings inequality is in both countries a consequence of increased upper tail dispersion. In the U.S. this increase is mainly accounted for by changing market prices and increased residual dispersion. In Sweden on the other hand, compositional effects account for most of the change in the P90-P50 gap, while changing

⁷The results presented in this paper are very similar to those in recent studies like Lemieux (2005) and Autor et al. (2005b).

market prices and increased residual dispersion account for all of the change in the P99-P90 gap. Third, relative labor demand of both low skilled workers and the most skilled workers increased in the U.S. In Sweden, private sector labor demand mimics the U.S. changes in labor demand. But the increase in private sector demand for low skilled workers in Sweden was more than wiped out by a large fall in public sector demand.

The remainder of the paper is organized as follows. Section 2 describes the data and the method. Section 3 presents results for Sweden. Section 4 compares the Swedish development with that in the U.S. Section 5 concludes.

2 Data and Method

The Swedish data is drawn from the LOUISE database, which is a register database covering all individuals between 16 and 64 years of age for the period 1990-95 and all individuals above age 16 thereafter. This study restricts attention to workers between ages 20-64. In the sample there are between 3.3 and 3.9 million individuals each year. All results are based on total earnings. See appendix A.1 for further details.

Changes in the distribution of earnings are decomposed into changes in three components using the method in Lemieux (2005). First, prices of observable characteristics of workers may have changed. Second, the composition of the workforce may have changed and third, prices and the composition of unobserved characteristics such as intelligence and social skills, may have changed.

The starting point for the analyses is a version of Mincer (1974)'s human capital earnings function

$$\log w_{it} = x_{it}\beta_t + e_{it},$$

where w_{it} is earnings of individual i in time period t , x is an exhaustive set of dummy variables dividing the sample in a set of J cells according to observable characteristics (age, sex, education, race and industry), β is the price of these characteristics and e is an error term assumed to have a zero conditional mean; $E(e_{it}|x_{it}) = 0$. Consider some distributional statistic $V_t = V(\log w_{it})$ of the distribution of $\log w_{it}$. The change in V_t between two time periods t and s is decomposed as follows. First, consider the OLS estimate b_t of β_t . The estimated regression is then

$$\log w_{it} = x_{it}b_t + u_{it},$$

where u is the regression residual. Second, construct the counterfactual wage $\log \hat{w}_{it}$ which would prevail in period s if the price of observable characteristics were the same as in s

$$\log \hat{w}_{it} = x_{it}b_s + u_{it},$$

where b_s is the OLS estimate of β_t . Calculate the counterfactual statistic $\hat{V}_t = V(\log \hat{w}_{it})$. Third, calculate the counterfactual statistic $\bar{V}_t = V(\log \hat{w}_{it})$ weighting each observation by $\hat{\omega}_{it} = \omega_{it}\theta_{js}/\theta_{jt}$, $i \in j$, where ω_{it} is the sample weight and θ_{jt} is the sample share of demographic group j in period t . As shown by Lemieux (2005) this statistic accounts for changes in observables. The overall change in the statistic V_t can then be written as

$$V_t - V_s = (V_t - \hat{V}_t) + (\hat{V}_t - \bar{V}_t) + (\bar{V}_t - V_s)$$

The first term captures the change that is due to changes in observed prices. The second term captures the effect of changes in observed quantities, i.e., a change in the composition of demographic groups. The last term captures the contribution of changes in unobserved prices and quantities, i.e., the unexplained part of earnings.

To implement this I ran separate regressions for each year, by gender and education (less than high school, high school graduate, some college and college graduate), to remove age (9 age dummies, five-year interval between ages 20-64) and industry effects (17 industry dummies).

3 Changes in Swedish Earnings Inequality

Table 1 shows how several different distributional statistics have changed between 1990 and 2002, and also the extent to which these changes are accounted for by changes in prices, quantities or in unobservables. Table 1 calls for several remarks.

First, between 1990 and 2002 Sweden experienced a large increase in earnings inequality as measured by the cross-sectional variance. The decomposition shows that only 12 percent was due to changes in observed prices. The change in the distribution across demographic groups accounted for 48 percent of the change and 40 percent was due to changes in the variance of unobservables. Second, the rise in earnings inequality is a consequence of rising upper tail dispersion. Whereas the lower tail, captured by the P10-P1 gap or the P50-P10 gap remained essentially unchanged, the upper tail

became much thinner. In particular, at the very top, the P99-P90 gap increased by more than 0.10 log-points. Third, the decomposition shows that different effects account for the changes in the upper and the lower tail of the distribution. Price effects only matter at the very top and the very bottom of the distribution and tend to increase inequality. But whereas changes in unobservables cancel the price effects at the lower tail, they re-enforce the same effects at the very top. Changes across demographic groups are most important factor for the mid-range of the distribution, and in particular in the P90-P50 range.

Accounting for compositional changes is thus important. But which of these compositional changes are most important? Table 2 shows what happens to the residual variance when different changes in the distribution across demographic groups are accounted for. It shows that the variance of the regression residual is reduced by 41 percent when all distributional changes across demographic groups are accounted for. It also shows that accounting for education and industry composition is important. The change in educational composition accounts for one third of the reduction in residual inequality and the change in industry composition for the remaining two thirds. Changes in the age and gender composition had no effect on the residual inequality.

3.1 *Gender and Education*

Even though changes in the gender composition is not important, Tables 3 and 4 show that there are some interesting differences between sexes. Table 3 shows that men experienced a larger increase in the variance of total earnings, and this difference was largely due to a much larger increase in residual inequality. Table 4 shows that this increase did not take place across the entire male earnings distribution. Rather, while residual dispersion was the main factor behind the increase in the P50-P10 and P99-P90 gaps, it worked to compress earnings at the low end of the distribution, and it was the least important factor in the P90-P50 range.

For women price and composition effects were relatively more important. Larger price effects (and less negative changes in unobservables) for women led to less compression in the lower tail. Similarly, much larger composition effects (and somewhat smaller changes in unobservables) led to more dispersion at the top end of the distribution (see Table 4).

Classifying individuals according to educational levels, reveals that the

largest differences between men and women occur for high school drop-outs and for college graduates. Among high-school drop-outs, changes in composition and unobservables increased male inequality, in particular in the P50-P10 range, but compressed female earnings, in particular in the P90-P50 range. Among college graduates on the other hand, the difference is due to changes in unobservables. Male college graduates experienced a large increase in the residual variance and the P99-90 gap increased by 0.25 log points, out of which 80 percent was due to increased residual inequality. For female college graduates the residual variance was essentially unchanged. For these women, the largest change occurs in the P90-P50 gap and was primarily driven by price changes.

The largest increases in the cross-sectional variance were for men and women with some college education. For women, almost two-thirds of this increase is explained by a large price effect and the remaining third is due to increased residual inequality. For men with some college education on the other hand, there was a large negative composition effect, counteracted by an even larger increase in residual dispersion. These effects were particularly large in the P50-P10 range.

Finally, price effects of observables were rather similar when comparing men and women across educational levels.

3.2 *Reallocation across industries*

As noted above, accounting for the composition across industries is important when analyzing changes in the earnings distribution. Table 5 shows the eight most important industries in terms of contribution to the overall change in residual dispersion. These industries account for approximately 87 percent of the change in residual inequality but only 54 percent of employment. To what extent are these changes in residual dispersion a result of changes in labor demand towards more (or less) skill intensive industries? Table 6 shows the fractions of workers in the top, middle and bottom 10 percent of the earnings distribution that is employed in each industry. The table shows that there are significant differences in skill composition across industries, and that there was a shift towards less skill intensive industries such as services and more skilled intensive industries such as professional services. In fact, the five least skill intensive industries saw their share of employment increase by 4.7 percent, and the five most skill intensive industries saw their share of the employment increase by 19 percent. Table 5 also

indicates that industries characterized by large employment changes experienced large changes in residual dispersion. The correlation between the change in employment shares and the contribution to the change in residual dispersion is 0.29.

To further investigate the importance of these employment shifts across industries we follow Juhn et al (1993) and construct a labor demand index for workers of different skill levels. Let output in industry j , be given by

$$Y_j = F_j(L_{1j}, \dots, L_{100j}),$$

where L_{pj} is the number of workers in percentile p . Assuming that the production function F_j displays constant returns to scale, the 100×1 vector of labor demand in industry j is given by

$$\mathbf{L}_j = Y_j \times \mathbf{D}_j(W_1, \dots, W_{100}),$$

where \mathbf{D}_j is a vector of unit demand functions and W_p is the wage rate at percentile p . Changes in the industrial structure then implies that labor demand changes as follows

$$d\mathbf{L}_j = \sum_j \frac{dY_j}{Y_j} \times \mathbf{L}_j.$$

To construct the demand index, I measure the change in output by the relative change in value added by industry,⁸ and \mathbf{L}_j is the distribution of workers across industries by percentile over the entire sample.

Figure 3 displays the changes in relative demand of workers across the distribution that are driven by reallocation of workers across industries. The figure shows that the 1990's has been characterized by rising demand for the most skilled workers and a modest fall for all other skill-levels (except for those around the 20th percentile who experienced a larger fall in demand). At first, this suggests that recent demand changes can not be characterized as both a shift towards high-paid jobs in professional services and a shift towards low-paid jobs in trade and other services. However, Figure 3 also shows that changes in relative demand by the private sector differs substantially relative to those of the whole economy. In particular, demand for low skilled individuals has increased in the private sector. Among high skilled

⁸Data is from Statistics Sweden: www.scb.se

individuals the increase in demand in the private sector is confined to the very top of the skill distribution.

The difference is due to the development in the public sector.⁹ Several authors (e.g. Lindbeck (1997), Freeman et al. (1997)) have emphasized that the dramatic expansion of the public sector in Sweden between 1970 and 1990 disproportionately involved the hiring of low skilled workers. Lindbeck for example (1997, pp. 1280, 1311) argues that the expansion was "a consequence of the government serving as 'an employer of last resort'," keeping up the demand the demand for low and medium-skilled workers.

The figure indicates that a change occurred in the early 1990's. The low-skilled intensive parts of the public sector (Health and social services) contracted and the skill-intensive parts (Education and Public administration) expanded.¹⁰ As the demand for low-skilled workers in the public sector fell more than it increased in the private service sector, the net outcome for the whole economy was a slight fall. In sum, this suggests that employment changes in the private sector indeed have been characterized by increased demand for both low and high skilled workers, but that the development in the public sector have counteracted the former change and re-enforced the latter.

4 A comparison between Sweden and the United States

In this section I compare results for Sweden with changes in the U.S. earnings distribution. As changes in the U.S. wage distribution are well documented elsewhere (recent studies include Autor et al. (2005b) and Lemieux (2005)), I will highlight similarities and differences between the Swedish and the U.S. experience. Table 7-10 shows the equivalent statistics for the U.S. earnings distribution as those that were displayed in Tables 1-3 and 5 for the Swedish earnings distribution. The results for the U.S. are derived using data from the Current Population Survey (CPS). I use the same sample selection criteria as Autor et al. (2005a) and have slightly more than 60 000 observations

⁹The public sector is defined as Health and social services, Education and Public administration.

¹⁰There are two ways this change could have arisen. Either, the public sector was no longer willing to act as an employer of last resort, or the political process determined to change the composition of goods produced by the public sector.

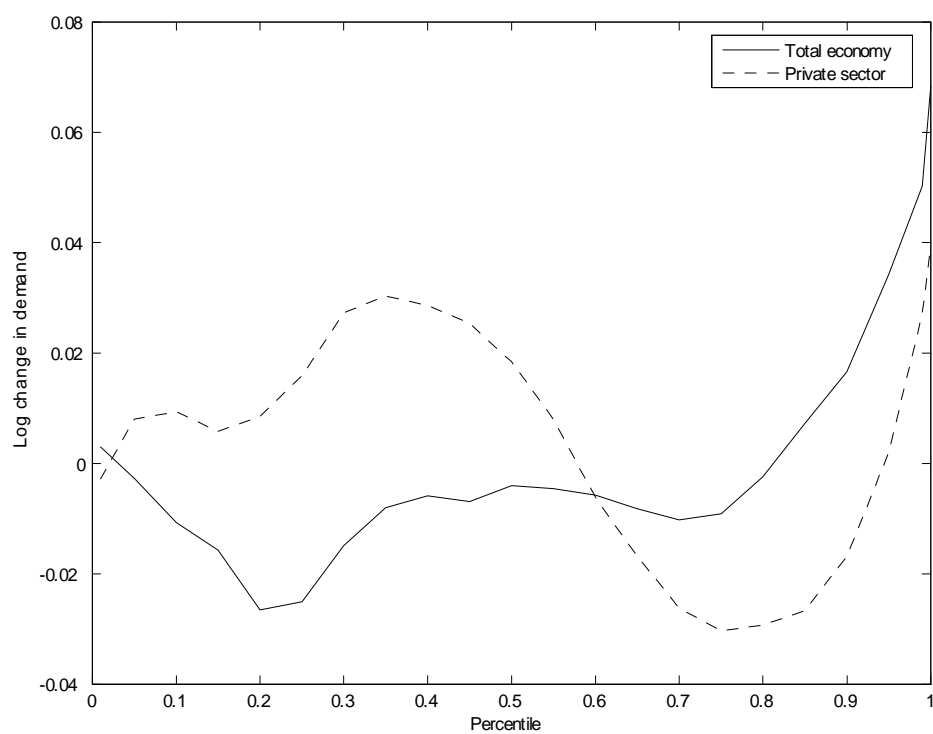


Figure 3: Changes in labor demand by earnings percentiles.

per year. See Appendix A.2 for further details. To increase the number observation in each demographic group in the U.S. data, I pooled the years 1989 to 1991 and 2000 to 2002. As for Sweden I focus on total earnings and full-time workers.¹¹ In the regressions on U.S. data, I add controls for race (white/non-white), and year dummies (to handle year effects associated with pooling of the data).

Several observations are worth noting. First, Table 7 shows the well-know fact that earnings inequality is much larger in the U.S. than in Sweden. The cross-sectional variance is twice as large and the P90-P50 and P50-P10 gaps are approximately 0.2 log-points larger. Second, the change in earnings inequality is qualitatively similar but larger in the U.S. Composition effects and changes in residual inequality explain most of the change, but their relative importance differ. Changes in residual inequality was more important in the U.S., and compositional change was more important in Sweden.

Third, as noted by Autor et al. (2005b) changes in prices and residual inequality led to increased dispersion above the median, and more compression below the median in the U.S. Compositional changes re-enforced these changes above the median, but counter-acted them below the median. As a result, changes in the earnings distribution mainly occurred above the median. In Sweden, the largest changes in the earnings distribution also occurred above the median, but changes in prices and residual dispersion only accounts for changes at the very top. The rise in the P90-P50 gap is largely due to shifts in the Swedish labor force composition.

Fourth, Table 8 shows that accounting for compositional changes across demographic groups has reduced the residual variance in the U.S. and Sweden by a similar absolute magnitude. The effect of accounting for education and industry is also similar. In contrast to Sweden, U.S. changes in the gender (and race) composition increased the residual variance whereas changes in the age composition decreased it.

Fifth, Table 9 shows that there are much smaller differences between men and women in the U.S. than in Sweden, both in terms of overall changes and in terms of which factors that account for these changes. Finally, Table 10 shows that the overall change in residual dispersion is more evenly spread across U.S. industries. The 8 industries contributing most to the overall change in residual dispersion account for approximately 69 percent of

¹¹I have also used all workers with similar results, albeit somewhat smaller in magnitude.

the change in residual inequality, and employs 72 percent of the workforce. Moreover, the change in residual inequality was larger in contracting industries in the U.S., whereas the opposite was true in Sweden. The correlation between the change in employment shares and the contribution to the change in residual dispersion is -0.45 in the U.S., and as noted above 0.29 in Sweden.

Finally, Figure 4 graphs the changes in relative demand of workers across the distribution, that are driven by reallocation of workers across industries.¹² The figure shows that the 1990's has been characterized by modest demand changes except at the very ends of the distribution. It is noteworthy, however, that demand has increased for workers below the median and for the most skilled, and fallen for upper 'middle-skilled' workers. Moreover, contrary to the Swedish case, there are no major differences between the private and the public sector in the U.S. If anything, the increase in demand for low skilled workers was larger in the public sector than in the private sector in the U.S., which stands in stark contrast to the Swedish experience.

5 Concluding Remarks

This paper decomposes the rise in cross-sectional earnings inequality in Sweden between 1990 and 2002 into changes in market prices of observable characteristics, changes in the composition of the labor force across demographic groups and industries, and changes in unobservables, and compares the Swedish experience with that in the U.S. The rise in earnings inequality is in both countries a consequence of rising upper tail dispersion. Contrary to the U.S. experience, where the rise is largely driven by changing market prices of observables and increased residual dispersion, shifts in the Swedish labor force composition have contributed positively to the rise in the P90-P50 gap. The rise in the Swedish P99-90 gap is however entirely accounted for by changes in prices and residual dispersion.

The diverging changes in the upper and lower half of the earnings distribution are not well understood. Autor, Levy and Murnane (2003) hypothesis that routine tasks performed by workers in the mid-range of the skill distribution computers are easier to substitute for by computers, than non-routine jobs, either low-skilled jobs in services or high-skill jobs in professional services. The increased usage of computers may then have had an asymmetric

¹²Data on U.S. output changes by industries are from the the Bureau of Economic Analysis: www.bea.gov.

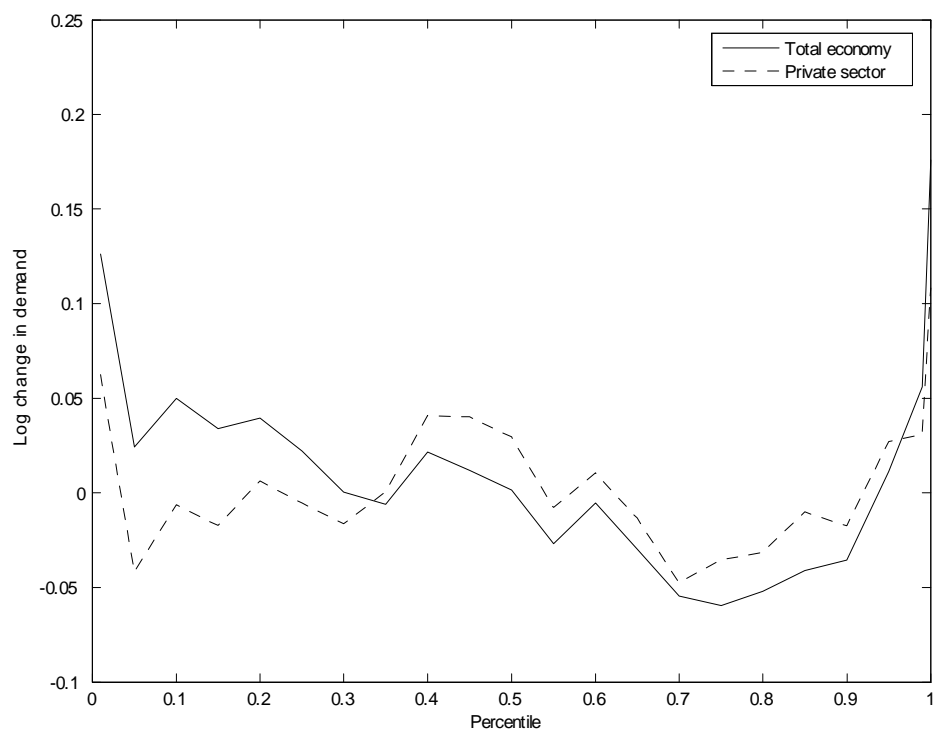


Figure 4: Changes in U.S. labor demand by earnings percentiles.

impact on the demand for skills.

This paper finds evidence that the demand for 'middle-skilled' workers have decreased, which supports the idea that computerization may explain the diverging trends. But the findings also suggest that employment decisions made by the public sector can have a significant impact on the earnings distribution. Employment in the public sector is used either to produce public goods that differ from what would be demanded and produced if the workers were instead employed in the private sector, or when used to produce close substitutes to private sector output, most likely in a less efficient way.¹³ In either case, public sector employment changes are determined by potentially other factors than those driving employment outcomes in the private sector. That public sector employment can be an important determinant for wage outcomes has previously been emphasized by Edin and Topel (1997) and Domeij and Ljungqvist (2006) and is a finding that merits further analysis.

References

- [1] Acemoglu, Daron. 2002. "Technical Change, Inequality, and the Labor Market." *Journal of Economic Literature*. Vol. 40, No. 1, pp. 7-72
- [2] Autor, David H., Lawrence. F. Katz and Melissa S. Kearney. 2005a. "Trends in U.S. Wage Inequality: Re-Assessing the Revisionists." *NBER Working Papers*, No. 11627, National Bureau of Economic Research.
- [3] Autor, David H., Lawrence. F. Katz and Melissa S. Kearney. 2005b. "Rising Wage Inequality: The Role of Composition and Prices." *NBER Working Papers*, No. 11628, National Bureau of Economic Research.
- [4] Autor, David H., Frank Levy and Richard J. Murnane. 2003. "The Skill Content of Recent Technological Change: An Empirical Investigation." *Quarterly Journal of Economics*, Vol. 118, pp. 1279-1333.
- [5] Bluestone, Barry, and Bennett Harrison. 1988. "The Growth of Low-Wage Employment: 1963-86" *American Economic Review Papers and Proceedings*, Vol. 78, pp. 124-128.

¹³Rosen (1997) for example concludes that there is an excessive supply of publicly provided child-care services in Sweden, than would have been if household services had been paid for privately and transacted through the market.

- [6] Domeij, David and Lars Ljungqvist. 2006. "Wage Structure and Public Sector Employment: Sweden versus the United States 1970-2002", Stockholm School of Economics, mimeo
- [7] Edin, Per-Anders and Peter Fredriksson. 2000. "LINDA - Longitudinal INdividual DAta for Sweden", Working Paper 2000:19, Department of Economics, Uppsala University
- [8] Edin, Anders and Robert Topel 1997, "Wage Policy and Restructuring: The Swedish Labor market since 1960" in Richard B. Freeman, Robert Topel and Birgitta Swedenborg (eds.) *The Welfare State in Transition: Reforming the Swedish Model*. Chicago: The University of Chicago Press, pp.155-202.
- [9] Freeman, Richard B., Robert Topel and Birgitta Swedenborg, 1997."Introduction" in *The Welfare State in Transition: Reforming the Swedish Model*. Chicago: The University of Chicago Press, pp. 1-32.
- [10] Gottschalk, Peter and Sheldon Danziger. 2003. "Wage Inequality, Earnings Inequality and Poverty in the U.S. Over the Last Quarter of the Twentieth Century." Boston College, mimeo.
- [11] Gustavsson, Magnus. "The 1990s Rise in Swedish Earnings Inequality - Persistent or Transitory?", *Applied Economics*, forthcoming.
- [12] Johansson, Mats. 2006. "Inkomst och ojämlikhet i Sverige 1951-2002." Working paper 2006:3 Institutet för Framtidsstudier.
- [13] Juhn, Chinhui, Kevin M. Murphy and Brooks Pierce. 1993. "Wage Inequality and the Rise in Returns to Skill." *The Journal of Political Economy*, Vol. 101, No. 3, pp. 410-442.
- [14] Katz, Lawrence F. and David H. Autor. 1999. "Changes in the Wage Structure and Earnings Inequality" in Orley Aschenfelter and David Card (eds.) *Handbook of Labor Economics* Vol. 3A, Elsevier: Amsterdam, pp. 1463-1555.
- [15] Lemieux, Thomas. 2002. "Decomposing Changes in Wage Distributions: A Unified Approach." *Canadian Journal of Economics*. Vol. 34, No. 4, pp. 646-688.

- [16] Lemieux, Thomas. 2005. "Increasing Residual Wage Inequality: Composition Effects, Noisy data, or Rising Demand for Skill?" University of British Columbia, mimeo.
- [17] Lindbeck, Assar. 1997. "The Swedish Experiment." *Journal of Economic Literature*, Vol. 35, No. 3, pp 1273-1319.
- [18] Mincer, Jacob. 1974. *Schooling, Experience and Earnings*. New York, National Bureau of Economic Research.
- [19] Rosen, Sherwin. 1997. "Public Employment, Taxes and the Welfare State in Sweden" in Richard B. Freeman, Robert Topel, and Birgitta Swedenborg (eds.) *The Welfare State in Transition: Reforming the Swedish Model*. Chicago: The University of Chicago Press, pp. 79-108.

A Data

A.1 Swedish data

The Swedish data is drawn from the LOUISE database. This study restricts attention to workers between ages 20-64. Self-employed workers are not included. All results are based on total earnings defined as wages and salaries paid by the employer (the variable Arbetsinkomst in LOUISE). It also includes the following taxable work-related income transfers; maternity and paternity leave payments and sickness payments. To focus on full-time workers I remove those with income less than 25 percent of mean income as a way eliminating those with short work spells.¹⁴ In the sample there are between 3.3 and 3.9 million individuals each year; for 1990 and 2002 the sample consists of 3,801,079 and 3,628,958 individuals respectively.

I use the following classification scheme for education. I divide individuals into 4 groups; (i) less than high school graduates, (ii) high-school graduate, (iii) some college, and (iv) college graduate. Note that a high-school diploma can be obtained in Sweden after 11 or 12 years of schooling and I classify both groups as high school graduates. Moreover, obtaining a traditional Swedish college degree requires 3-5 years of studies depending on field of specialization. I thus classify individuals who have completed a 3 year university or college degree as college graduates.

¹⁴This means that we remove approximately 2 percent of the sample each year.

Finally, I classify individuals into 17 industries¹⁵; Agriculture, Hunting, Forestry and Fishing (01111-05025), Mining (10100-14500), Utilities (40110-41002), Construction (4511045500), Manufacturing, nondurables (15111-19300, 2111-25240), Manufacturing, durables (20101-20520, 26110-37200), Wholesale trade (50101-51900), Retail trade (52111-52740), Transportation and warehousing (60100-64120), Information and communication (64201-64203), Financial intermediation (65110-67202), Real estate and renting (70110-71402), Administrative services, Entertainment, Accommodation and Food services, Other community, social and personal services (55102-55529, 74501-74874, 90010-95000), Health and social care (85111-85329), Professional services (72100-74409), Education (80101-80429), Public administration (75111-75300) with SNI2002 codes in parenthesis. Note that there was a change in industry classification in 2002. With one exception this had little consequence for the classification used in this paper. The exception was day-care workers working with 6 year old kids were prior to 2002 classified as working in Health and social care, but beginning in 2002 they were classified as working in Education. To keep consistency, all day-care workers that work at the same workplace in both 2001 and 2002 are classified as working in Health and social work.¹⁶

A.2 U.S. data

The results for the U.S. are derived using data from the 1990-2003 March CPS Annual Demographic Survey files provided by UNICON, from which we extract data for the years 1989-2002.¹⁷ Following Autor et al (2005a) I exclude (i) the self-employed, (ii) individuals with less than 40 weeks worked per year, (iii) individuals who work less than 35 hours per week, (iv) individuals with allocated income, (v) individuals whose weekly pay (after topcode adjustment has been made, see below) is less than \$67 in 1982 dollars (I have used CPI when transforming earnings into 1982 dollars), or whose hourly

¹⁵The choice of 17 industries is a results of aiming for consistency within each country and between countries. For Sweden I have experienced 29 industries following the broad industri categories in the SNI2002 classification. Results are very similar. As the samle for the U.S in smaller I have experimented with using a smaller set of industries (8), and results are again very similar.

¹⁶I have also excluded the year 2002, and focused on changes during the period 1990-2001. Results are very similar.

¹⁷See Autor et al (2005b) for a disucssion of the relative merits of using the CPS March data or the May/ORG data.

wage exceed 1/35th the topcoded value of weekly earnings. To give an indication of the sample size, the number of remaining observations are 141,211 for the years 1989-91 and 252,132 for the years 2000-2002.

Prior to March 1989, wage and salary income is reported in the CPS as a single variable, that is topcoded at between \$50,000 and \$99,999. Beginning in 1989, wage and salary income is reported in two separate variables, corresponding to primary and secondary earnings. Beginning in 1996, topcoded earners are assigned the mean of all topcoded earners (averages when individuals are grouped according to sex, race and worker status). I follow Autor et al. (2005a) and (i) topcode primary and secondary earnings separately before summing them, (ii) recensor primary and secondary earnings for the years 1996-2002 to their topcode values and (iii) multiply topcoded values by 1.5.

I use the following classification scheme for education; (i) less than 12 years of schooling, (ii) high-school graduate (completed 12 years of schooling), (iii) some college, and (iv) college graduate. College graduate include individuals who have completed 16 years of schooling. In 1992 there was a change in the recording of educational attainment. To keep consistency in classification I follow the suggestion in the UNICON documentation and classify individuals whose 13th year of schooling is not completed as 'with some college'.

I classify individuals in 17 industries using the CPS industry classification as follows. For the 2003 wave; Agriculture, Hunting, Forestry and Fishing (170-290), Mining (370-490), Utilities (570-690), Construction (770), Manufacturing, nondurables (1070-2390), Manufacturing, durables (2470-3990), Wholesale trade (4070-4590), Retail trade (4670-5790), Transportation and warehousing (6070-6390), Information and communication (6470-6780), Financial intermediation (6870-6990), Real estate and renting (7070-7190), Administrative services, Entertainment, Accommodation and Food services, Other community, social and personal services (7580-7790, 8560-8690), Health and social care (7970-8470), Professional services (7270-7590), Education (7860-7890), Public administration (9370-9590). For the 1990-2002 waves; Agriculture, Hunting, Forestry and Fishing (10-32), Mining (40-50), Utilities (450-472), Construction (60), Manufacturing, nondurables (100-229), Manufacturing, durables (230-392), Wholesale trade (500-571), Retail trade (580-691), Transportation and warehousing (400-432), Information and communication (440-442), Financial intermediation (700-711), Real estate and renting (712), Administrative services, Entertainment, Accom-

modation and Food services, Other community, social and personal services (721-810, 872-881), Health and social care (861-871), Professional services (882-893), Education (842-860), Public administration (900-932).

Table 1: Changes in Swedish earnings inequality: 1990 to 2002

	1990	2002	Total change	Observed prices	Observed quantities	Unobserved prices and quantities
Cross-sectional variances	0.1812	0.2102	0.0290	0.0035 (0.120)	0.0140 (0.483)	0.0116 (0.398)
P10-P1 gap	0.5659	0.5628	-0.0031	0.0281	0.0125	-0.0436
P50-P10 gap	0.6004	0.6171	0.0167	-0.0047	0.0134	0.0080
P90-P50 gap	0.4532	0.5057	0.0525	-0.0049	0.0346	0.0228
P99-P90 gap	0.5626	0.6662	0.1036	0.0358	0.0093	0.0585

Note: Share of total change within parenthesis.

Table 2: Composition effects and the residual variance

	$var(u_{it})$	Controlling for the distribution in		
		gender &	education &	age & industry
Change 1990 to 2002	0.0197 (1.000)	0.0197 (1.000)	0.0169 (0.858)	0.0169 (0.587)

Note: Share of total change in $var(u_{it})$ within parenthesis.

Table 3: Changes in the cross-sectional variance of male and female earnings: 1990 to 2002

	Total change	Observed prices	Observed quantities	Unobserved prices and quantities	Share of sample	
					1990	2002
All	0.0290	0.0042	0.0140	0.0116 (1.000)		
Men						
All	0.0439	0.0110	0.0127	0.0203 (0.890)	0.508	0.507
< High school	0.0243	0.0033	0.0062	0.0147 (0.210)	0.145	0.100
High school	0.0111	0.0004	-0.0005	0.0112 (0.270)	0.245	0.253
Some college	0.0636	0.0460	-0.0312	0.0489 (0.270)	0.056	0.078
College grad.	0.0462	0.0071	0.0109	0.0282 (0.171)	0.062	0.077
Women						
All	0.0291	0.0088	0.0177	0.0026 (0.110)	0.492	0.493
< High school	-0.0077	0.0040	-0.0040	-0.0077 (-0.094)	0.125	0.100
High school	0.0219	0.0125	0.0079	0.0015 (0.035)	0.244	0.211
Some college	0.0671	0.0405	0.0041	0.0225 (0.152)	0.069	0.088
College grad.	0.0182	0.0047	0.0159	-0.0025 (-0.013)	0.055	0.094

Note: Share of change in overall residual dispersion within parenthesis.

Table 4: Changes in male and female earnings inequality: 1990 to 2002

Change in	Men					Women				
	All	< High school	High school	Some college	College grad.	All	< High school	High school	Some college	College grad.
P10-P1 gap	-0.0477	-0.0895	-0.0450	-0.1679	-0.0766	-0.0104	-0.0145	-0.0436	-0.1079	-0.0702
effect of prices	0.0132	0.0121	0.0053	0.0313	-0.0397	0.0267	0.0201	0.0193	0.0632	-0.0239
quantities	0.0197	-0.0213	0.0243	0.0557	0.0041	0.0088	-0.0065	-0.0100	0.0165	-0.0072
unobservables	-0.0806	-0.0803	-0.0746	-0.2549	-0.0410	-0.0460	-0.0259	-0.0529	-0.1876	-0.0391
P50-P10 gap	0.0595	0.1075	0.0112	0.1866	0.0034	0.0474	0.0246	0.0523	0.1942	-0.0339
effect of prices	0.0127	0.0098	-0.0005	0.0964	-0.0052	-0.0066	-0.0066	-0.0068	0.0492	-0.0267
quantities	-0.0067	0.0344	-0.0248	-0.0974	0.0248	0.0362	-0.0102	0.0242	0.0356	0.0286
unobservables	0.0535	0.0633	0.0365	0.1877	-0.0163	0.0178	0.0413	0.0349	0.1093	-0.0358
P90-P50 gap	0.0810	0.0215	0.0173	0.0752	0.0896	0.0373	-0.0458	0.0386	0.0773	0.1265
effect of prices	0.0192	0.0059	-0.0011	0.0416	0.0320	0.0179	0.0122	0.0521	0.0607	0.0628
quantities	0.0444	0.0046	0.0114	-0.0068	0.0030	0.0163	-0.0087	0.0093	-0.0130	0.0402
unobservables	0.0174	0.0110	0.0069	0.0405	0.0546	0.0031	-0.0493	-0.0229	0.0296	0.0235
P99-P90 gap	0.1226	0.0495	0.0863	0.1142	0.2446	0.1518	0.0329	0.0941	0.1003	0.0765
effect of prices	0.0346	0.0019	0.0031	0.0186	0.0166	0.0399	0.0098	0.0237	0.0299	0.0320
quantities	0.0007	0.0099	0.0033	-0.0034	0.0254	0.0651	-0.0084	0.0027	-0.0029	-0.0066
unobservables	0.0873	0.0377	0.0799	0.0990	0.2025	0.0468	0.0315	0.0677	0.0733	0.0510

Table 5: The 8 industries contributing most to overall change in residual dispersion: 1990-2002

	Total change in cross-sectional variance	Observed prices	Observed quantities	Unobserved prices and quantities	Share of sample 1990	Share of sample 2002
All industries	0.0291	0.0042	0.0140	0.0116 (1.000)		
Retail trade	0.0252	-0.0015	0.0182	0.0084 (0.039)	0.051	0.050
Manufacturing, durable goods	0.0143	-0.0002	0.0098	0.0047 (0.061)	0.141	0.131
Construction	0.0131	0.0001	-0.0026	0.0156 (0.097)	0.069	0.053
Wholesale trade	0.0227	0.0044	0.0018	0.0166 (0.098)	0.065	0.064
Services	0.0237	0.0019	0.0061	0.0157 (0.109)	0.076	0.095
Professional services	0.0227	0.0096	-0.0202	0.0332 (0.132)	0.044	0.065
Financial intermediation	0.0979	0.0156	0.0142	0.0681 (0.144)	0.023	0.024
Information and communication	0.0468	0.0014	0.0056	0.0398 (0.230)	0.064	0.061

Note: Share of change in overall residual dispersion within parenthesis.

Services include Administrative services, Entertainment, Accomodation and Food services, Other community, social and personal services

Table 6: Industry distribution by percentiles 1990-2002

	Percentile ¹		
	0-10	45-55	90-100
Agriculture, Hunting, Forestry and Fishing	3.41	1.86	1.04
Mining	0.09	0.19	0.31
Utilities	0.23	0.48	1.62
Construction	3.11	5.85	5.67
Manufacturing, nondurables	4.97	6.65	8.38
Manufacturing, durables	7.92	16.75	14.44
Wholesale trade	4.40	7.00	9.18
Retail trade	9.37	4.25	1.82
Transportation and warehousing	0.12	0.29	0.49
Information and communication	5.03	6.90	5.85
Services	16.05	7.52	5.30
Financial intermediation	1.32	1.99	4.83
Real estate and renting	1.70	2.02	1.81
Health and social care	26.13	21.96	8.32
Professional services	3.13	3.38	14.30
Education	9.36	6.79	7.65
Public administration	3.67	6.14	8.97

Note:¹ Percent of workers in the percentile range that are employed in industry j .

Table 7: Changes in U.S. earnings inequality: 1989-91 to 2000-02

	1990-91	2000-02	Total change	Observed prices	Observed quantities	Unobserved prices and quantities
Cross-sectional variances	0.3640	0.4101	0.0460	0.0043 (0.093)	0.0145 (0.315)	0.0272 (0.592)
P50-P10 gap	0.7802	0.7870	0.0068	-0.0067	0.0153	-0.0017
P90-P50 gap	0.7208	0.8141	0.0933	0.0170	0.0111	0.0652

Note: Share of total change within parenthesis.

Table 8: Composition effects and the residual variance in U.S. earnings

	$var(u_{it})$	Controlling for the distribution in			
		gender	& education	& age	& race & industry
Change 1989-91 to 2000-02	0.0377 (1.000)	0.0384 (1.019)	0.0358 (0.951)	0.0331 (0.878)	0.0272 (0.723)

Note: Share of total change in $var(u_{it})$ within parenthesis.

Table 9: Changes in the cross-sectional variance of male and female earnings in the U.S.: 1989-91 to 2000-02

		Total change	Observed prices	Observed quantities	Unobserved prices and quantities	Share of sample 1989-91	Share of sample 2000-02
All		0.0460	0.0043	0.0145	0.0272 (1.000)		
Men	All	0.0589	0.0099	0.0199	0.0291 (0.622)	0.582	0.570
	< High school	-0.0179	-0.0212	-0.0020	0.0053 (0.016)	0.075	0.055
	High school	0.0149	-0.0170	0.0057	0.0263 (0.210)	0.207	0.184
	Some college	0.0203	-0.0178	0.0089	0.0292 (0.155)	0.137	0.152
	College grad.	0.0569	-0.0028	0.0202	0.0394 (0.249)	0.163	0.178
Women	All	0.0497	0.0114	0.0136	0.0247 (0.378)	0.418	0.430
	< High school	-0.0213	-0.0097	-0.0015	-0.0100 (-0.014)	0.037	0.028
	High school	0.0186	-0.0045	0.0043	0.0188 (0.116)	0.159	0.136
	Some college	0.0210	-0.0032	0.0006	0.0236 (0.102)	0.111	0.134
	College grad.	0.0498	0.0001	0.0107	0.0390 (0.166)	0.110	0.133

Note: Share of change in overall residual dispersion within parenthesis.

Table 10: The 8 industries contributing most to overall change in residual dispersion in the U.S.: 1989-91 to 2000-02

	Total change in cross-sectional variance	Observed prices	Observed quantities	Unobserved prices and quantities	Share of sample 1989-91	Share of sample 2000-02
All industries	0.0460	0.0043	0.0145	0.0272 (1.000)		
Transportation and warehousing	0.0340	0.0055	-0.0010	0.0295 (0.051)	0.048	0.049
Manufacturing, non-durable goods	0.0334	0.0095	0.0066	0.0173 (0.057)	0.092	0.065
Retail trade	0.0299	-0.0041	0.0183	0.0157 (0.071)	0.125	0.120
Services	0.0474	0.0171	0.0096	0.0207 (0.071)	0.095	0.120
Public administration	0.0390	0.0094	-0.0051	0.0348 (0.080)	0.064	0.059
Financial intermediation	0.0462	-0.0038	0.0074	0.0425 (0.093)	0.060	0.057
Education	0.0285	0.0010	-0.0053	0.0328 (0.096)	0.081	0.085
Manufacturing, durable goods	0.0597	0.0172	0.0029	0.0395 (0.185)	0.129	0.102

Note: Share of change in overall residual dispersion within parenthesis.

Services include Administrative services, Entertainment, Accomodation and Food services, Other community, social and personal services