Decomposition of differences in distribution using quantile regression[+]

Blaise Melly
Swiss Institute for International Economics and Applied Economics Research (SIAW)
University of St. Gallen

Address for correspondence:

Swiss Institute for International Economics and Applied Economic Research (SIAW)
University of St. Gallen
Bodanstrasse 8, 9000 St. Gallen, Switzerland
Phone: +41 71 224 22 99
Fax: +41 71 224 22 98
E-mail: blaise.melly@unisg.ch
Internet: www.siaw.unisg.ch/lechner/melly

Abstract:

This paper proposes a semiparametric estimator of distribution functions in the presence of covariates. The method is based on the estimation of the conditional distribution by quantile regression. The conditional distribution is then integrated over the range of the covariates. Counterfactual distributions can be estimated, allowing the decomposition of changes in distribution into three factors: coefficients, covariates and residuals. Sources of changes in wage inequality in the USA between 1973 and 1989 are examined. Unlike most of the literature, we find that residuals account for only 20% of the explosion of inequality in the 80s.

Keywords: Wage Inequality, Quantile Regression, Oaxaca Decomposition.

JEL classification: J31.

[+] I am grateful for helpful comments to Markus Frölich, Michael Lechner, Benita von Lindeiner, Winfried Pohlmeier and seminar participants at University of Konstanz and Sankt Gallen, at the 2004 ESEM meeting in Madrid and at the 2004 EALE meeting in Lisbon. All remaining errors are mine.

1 Introduction

The pronounced increase in wage inequality in several countries and particularly in the United States since the early 80s has motivated the study of changes in the distribution of wages. The most common practice is to calculate, compare and decompose summary indices of inequality like the Gini coefficient. However, as is well known in the income distribution literature, different summary measures of inequality can yield different rankings of inequality, since they put different weights on different parts of the distribution. As a result, recent research has increasingly focused on more global methods for describing changes in the whole distribution of wages. There has been a surge of methodologies extending the Oaxaca (1973) and Blinder (1973) decomposition of differences at the mean to decomposition of the whole distribution.

Juhn, Murphy and Pierce (1993, JMP hereafter) have proposed a simple extension of the Oaxaca decomposition by taking account of the distribution of residuals. A problem of this decomposition is that it does not account for heteroscedasticity. In the original paper, JMP formally allow for the distribution of residuals to depend on the covariates but they do not explain how to do it empirically and give no details. Most other applications of this decomposition do not condition on the covariates. If the error term is really independent and normally distributed, this procedure is efficient. However, if the location model is inappropriate, this decomposition can produce misleading results. The whole conditional distribution of wages and not only the first two moments can depend on the covariates. This lack of flexibility have motivated new estimators which are less restrictive.

One of the most interesting approaches is the weighted-kernel estimator introduced by DiNardo, Fortin and Lemieux (1996, DFL hereafter). It is non-parametrically identified and the main advantage of this approach is the lack of restrictions on covariates effects and density shapes. However, if there are too many variables, in particular continuous variables, it becomes impossible to estimate counterfactual distributions non-parametrically. They must

estimate the probability of being in one period given the characteristics with a probit model. Naturally, probit (or logit as used by Lemieux (2002)) estimates are consistent only if the error term is homoscedastic, normally (respectively logistically) distributed and if the specification is accurate. However, these residuals cannot be interpreted economically: What does "the probability of being in 1989" mean for an observation in 1973? This model does not seem particularly intuitive and cannot be linked with any theory or with past research. Another disadvantage of this decomposition is that we can only identify the effect of changes in the distribution of characteristics and residual changes which comprise the effects of coefficients and residuals. If these residual factors are high, we don't know why and we have no clue to understand these changes.

In this study we propose a new estimator of distribution functions in the presence of covariates. The whole conditional wage distribution is estimated by quantile regression. Then, the conditional distribution is integrated over the range of covariates to obtain an estimate of the unconditional distribution. The approach can be qualified as semiparametric. The quantile regression framework does not need any distributional assumptions and allows the covariates to influence the whole conditional distribution. Naturally, we must assume that the conditional quantiles satisfy a parametric restriction but this assumption is traditional in the literature and there are many possibilities to render it less stringent. The values of the coefficients have a natural interpretation as rates of return to the different components of human capital. Finally, the estimation of the conditional distribution allows us to naturally integrate the results in order to obtain the unconditional distribution; a procedure that is not possible with the conditional mean. Of course, we are not the first to propose a decomposition procedure based on quantile regression. Machado and Mata (2004) and Gosling et. al. (2000) have proposed similar procedures. We extend their works by solving the problem of crossing of different quantile curves and by determining the asymptotic distribution of the estimator.

We use the proposed approach to reassess the sources of changes in the distribution of wages in the United States between 1973 and 1989 using hourly wage data from the May Current Population Survey (CPS) and from the outgoing rotation groups of the CPS. Unlike most others (JMP, Katz and Autor 2000, Acemoglu 2002, for instance), we find that residuals account only for about 20% of the total growth in wage inequality. Changes in characteristics, on the contrary, explain about half of the increase. The reason of the differences between our results and those obtained with others methodologies is that quantile regression accounts for heteroscedasticity while others, like the JMP decomposition, assume independent error terms. However, the variance of the residuals expands as a function of education and experience and is smaller within unionized workers or certain sectors (public administration, manufacturing). The fact that the population is getting more educated, less unionized and that employment in sectors with low variance declines puts more weight on groups with higher within-group inequality. This is a composition effect and not an increase in the price of unmeasured skills as concluded traditionally.

The paper is organized as follows. In section 2 we present the methodology and show how it is possible to decompose differences in distribution into three factors: coefficients, covariates and residuals. Section 3 provides an application of the methodology to the distribution of wages in the United States between 1973 and 1989 and section 4 concludes.

2 Estimating distribution functions in the presence of covariates

2.1 Definition and motivation of the estimator

A good estimator of distribution functions in the presence of covariates must have some properties. It must be flexible in the way covariates affect the whole distribution, not only the first moment(s), of the dependent variable. A minimal number of assumptions should be imposed concerning the shape of the distribution function. The estimates must have a natural

economic interpretation and thus provide valuable information on the distribution of the variable in question. Finally, it must be estimable in the presence of a large number of possibly continuous covariates. Quantile regression is an excellent compromise on these requirements.

Let $\{y_i, x_i\}_{i=1}^N$ be an independent sample from some population where x_i is a $K \times 1$ vector of regressors. It is assumed that

$$F_{y|x}^{-1}(\tau|x_i) = x_i\beta(\tau), \forall \tau \in (0,1)$$

where $F_{y|x}^{-1}(\tau|x_i)$ is the τ^{th} quantile of y conditionally on x_i. A linear relationship is assumed between the quantiles of *y* and *x* similarly to OLS that assumes a linear relationship between the mean of *y* and *x*. Of course, this assumption is restrictive but can be relaxed by using dummy variables, polynomial expansions and interaction terms as it is done in the case of models that assume a linear functional form for the mean. In this application, the dependent variable is the logged wage and the covariates are human capital characteristics. Thus, the quantile regression coefficients can be interpreted as rates of return to the different characteristics at the specified quantile of the conditional distribution.

Koenker and Bassett (1978) show that $\beta(\tau)$ can be estimated by

$$\hat{\beta}(\tau) = \arg\min_{b \in \mathbb{R}^K} \frac{1}{N} \sum_{i=1}^{N} (y_i - x_i b)(\tau - 1(y_i \leq x_i b)).$$

where $1(\cdot)$ is the indicator function. $\beta(\tau)$ is estimated separately for each τ. Asymptotically, we could estimate an infinite number of quantile regressions. In finite samples, Portnoy (1991) shows that the number of numerically different quantile regressions is $O(N\log(N))$ and each prevails on an interval. Let $(\tau_0 = 0, \tau_1, ..., \tau_J = 1)$ be the points

where the solution changes. $\hat{\beta}(\tau_j)$ prevails from τ_{j-1} to τ_j for $j = 1,...,J$. Let $\hat{\beta}$ be the vector of all different quantile regression coefficients: $\hat{\beta} = \left(\hat{\beta}(\tau_1),...,\hat{\beta}(\tau_j),...,\hat{\beta}(\tau_J)\right)$.

This is a model for the conditional quantiles of *y*, but we want to estimate the unconditional quantiles of *y*. We therefore need to integrate the conditional distribution over the whole range of the distribution of the regressors. However, a problem with quantile regression is the potential lack of monotonicity, that is $\tau_j \le \tau_k \not\Rightarrow x_i\hat{\beta}(\tau_j) \le x_i\hat{\beta}(\tau_k)$. To overcome this problem, consider the following property of q_0, the population's θ^{th} quantile of *y*:

$$q_0 = F_Y^{-1}(\theta) \Leftrightarrow \int 1(y \le q_0)\, dF_y(y) = \theta \Leftrightarrow \int \left(\int 1(y \le q_0)\, f_{Y|X}(y|x)\, dy \right) dF_X(x) = \theta$$
$$\Leftrightarrow \int \left(\int_0^1 1\left(F_{Y|X}^{-1}(\tau|x) \le q_0\right) d\tau \right) dF_X = \theta$$

The last equivalence is obtained by changing the variable of integration and noting that $f_\tau(\tau_j) = 1,\ \forall \tau_j \in (0,1)$. Thus, replacing $F_{y|x}^{-1}(\tau_j|x_i)$ by its consistent estimate $x_i\hat{\beta}(\tau_j)$ and following the convention of taking the infimum of the set if the finite sample solution is not unique, the sample analog of q_0 is given by

$$\hat{q}(\hat{\beta}, x) = \inf\left\{ q : \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{J}(\tau_j - \tau_{j-1})\, 1\left(x_i\hat{\beta}(\tau_j) \le q\right) \ge \theta \right\}. \qquad (1)$$

Assuming traditional restrictions of the quantile regression model, one can prove that $\hat{q}$ is a consistent and asymptotically normally distributed estimator of q_0[1]. Given the difficulty in estimating the asymptotic variance, the statistical inference will be conducted with bootstrap procedures.

The formulas above and specially (1) may seem complicated to practitioners. However, what has to be done to estimate the θ^{th} quantile of *y* is very simple and consists of a straightforward 2-steps procedure:

1. Estimation of the whole quantile regression process $y = x\beta(\tau)$. In R^2, this can be done with a single line of command using the package *quantreg* written by Roger Koenker. Estimating the whole quantile regression process can take a very long time if the number of observations is big. However, the asymptotic results are also valid if the quantile regressions are estimated along a grid of τ-values whose mesh is sufficiently small (a mesh size of order $O\left(N^{-1/2-\varepsilon}\right)$ will work). If the estimation is based upon a big dataset (about $N > 3000$ depending on the available computer performance), a smaller number of quantile regressions should be estimated. We recommend to use the interior point algorithm written by Portnoy and Koenker (1997) in *R*.

2. Estimation of the θ^{th} quantile of the sample $\left\{\left\{x_i\hat{\beta}(\tau_j)\right\}_{j=1}^{J}\right\}_{i=1}^{N}$ by weighting each "observation" by $(\tau_j - \tau_{j-1})$. The weights are not necessary if a regular grid of quantiles has been used.

2.2 Decomposition of differences in distribution

The utility of estimating the unconditional distribution of a variable by using quantile regression as done in (1) is pretty small since the sample quantiles are in any case consistent (Glivenko-Cantelli theorem) and are simpler to estimate. The main interest in this estimator is the possibility of simulating counterfactual distributions that can be used to decompose differences in distribution. We use the same framework as JMP to decompose the differences in wage distributions between 1973 and 1989. Taking the median as a measure of the central tendency of a distribution, we can write a simple wage equation for each year

$$y_i^t = x_i^t\beta^t(0.5) + u_i^t,\ t = 73,89$$

---

[1] A formal proof and the asymptotic variance can be found in Melly (2004).

[2] *R* is an open-source programming environment for conducting statistical analysis and graphics. The software can be downloaded at no cost from the site www.r-project.org. See R Development Core Team (2003) for details.

where $\beta^t(0.5)$ is the coefficient vector of the median regression in year *t*. We can now isolate the effects of changes in characteristics *x*, coefficients $\beta(0.5)$ and residuals *u*. We estimate first the counterfactual distribution of wages that would have prevailed in 1973 if the distribution of individual attributes had been as it is in 1989 by minimizing (1) over the distribution of *x* in 1989 and using the coefficients estimated in 1973. Formally,

$$\hat{q}\left(\hat{\beta}^{73}, x^{89}\right) = \inf\left\{q : \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{J}\left(\tau_j - \tau_{j-1}\right)1\left(x_i^{89}\hat{\beta}^{73}\left(\tau_j\right) \le q\right) \ge \theta\right\}$$

is the θ^{th} quantile of this counterfactual distribution of wages. Thus, the difference between $\hat{q}\left(\hat{\beta}^{73}, x^{89}\right)$ and $\hat{q}\left(\hat{\beta}^{73}, x^{73}\right)$ is explained by changes in characteristics. This decomposition is less restrictive than the JMP decomposition because the characteristics are allowed to influence the whole conditional distribution of *y*.

To separate the effects of coefficients from the effects of residuals, note that the τ^{th} quantile of the residuals distribution conditionally on *x* is consistently estimated by $x\left(\hat{\beta}(\tau) - \hat{\beta}(0.5)\right)$. We define the $J \times 1$ vector $\hat{\beta}^{m89,r73}$ where its j^{th} element is given by $\hat{\beta}^{m89,r73}\left(\tau_j\right) = \left(\hat{\beta}^{89}(0.5) + \hat{\beta}^{73}\left(\tau_j\right) - \hat{\beta}^{73}(0.5)\right)$. Thus, we estimate the distribution that would have prevailed if the median return to characteristics had been the same as in 1989 but the residuals had been distributed as in 1973 by $\hat{q}\left(\hat{\beta}^{m89,r73}, x^{89}\right)$. Therefore, the difference between $\hat{q}\left(\hat{\beta}^{m89,r73}, x^{89}\right)$ and $\hat{q}\left(\hat{\beta}^{73}, x^{89}\right)$ is due to changes in coefficients since characteristics and residuals are kept at the same level. Finally, the difference between $\hat{q}\left(\hat{\beta}^{89}, x^{89}\right)$ and $\hat{q}\left(\hat{\beta}^{m89,r73}, x^{89}\right)$ is due to residuals.

The final decomposition is the following

$$\hat{q}\left(\hat{\beta}^{89}, x^{89}\right) - \hat{q}\left(\hat{\beta}^{73}, x^{73}\right) =$$
$$\left(\hat{q}\left(\hat{\beta}^{89}, x^{89}\right) - \hat{q}\left(\hat{\beta}^{m89,r73}, x^{89}\right)\right) + \left(\hat{q}\left(\hat{\beta}^{m89,r73}, x^{89}\right) - \hat{q}\left(\hat{\beta}^{73}, x^{89}\right)\right) + \left(\hat{q}\left(\hat{\beta}^{73}, x^{89}\right) - \hat{q}\left(\hat{\beta}^{73}, x^{73}\right)\right) \quad (2)$$

where the first bracket represents the effect of changes in residuals, the second the effects of changes in (median) coefficients and the third the effects of changes in the distribution of the covariates. Note that we can decompose all statistics (variance, difference between the 9th and the 1st deciles, Gini coefficient, coefficient of variation,…) since we can estimate the whole counterfactual distribution.

3 Changes in US wage inequality between 1973 and 1989

3.1 Introduction

A large and growing literature documents the changes in the US wage structure during the past three decades. Many researchers who used a variety of measures and datasets have found that wage inequality increased substantially during the 80's. An important finding is that residual inequality accounts for most of the growth in wage inequality. Katz and Autor (2000) survey this literature. They present a between- and within-group decomposition of the growth of the variance and find that residual inequality accounts for about 60% of the increase in inequality. Acemoglu (2002) summarizes four salient facts from the post-war US economy, one of which states: "Overall wage inequality rose sharply beginning in the early 1970s. Increases in within-group (residual) inequality account for much of this rise." JMP have implemented their decomposition presented in the introduction and found that 56% of the rise of the 90-10 wage differential from 1964 to 1988 is explained by residuals. On the other hand, in all these studies, the effect of changes in characteristics onto the rise in inequality is negligible, less than 10% in JMP for instance. Acemoglu (2002), Aghion (2002) and others use these results as building blocks for models of technical changes and economic growth.

However, this literature does not account for the fact that residuals are not necessarily independent of characteristics. As suggested by Mincer's (1974) famous human capital earnings model, residual wage dispersion should increase with experience and education. The literature on union and public sector wage effects shows also that the union membership and the public sector status reduce the variance of the unexplained component of earnings. Thus, changes in characteristics do not only affect the level of wages but also higher moments of the distribution. A part of the increase in the variance of residuals found in the literature is maybe due to changes in the composition of the workforce and not to higher returns to unobservable skills. The decomposition presented in section 2 can distinguish between both causes.

3.2 Data

We use hourly wage data from the May 1973 Current Population Survey (CPS) and from the 1989 outgoing rotation group files of the CPS. The samples used are broadly similar to those of DFL, Lemieux (2002) and Card and DiNardo (2002). The period has been chosen principally for three reasons. First, a major theme in the discussion on the widening of the wage distribution is the effect of the minimum wage. However, the minimum wage induces non-linearity of the wage function. JMP and Lemieux (2002) estimate linear regressions but, as recognized by Lemieux, this can only be an inadequate approximation in years where the real value of the minimum wage is high. To avoid this problem, we have simply chosen two years where the minimum wage was quite low and only few observations were at or below the minimum wage (0.33% in 1973, 0.28% in 1989). The differences in the distribution of earnings between these years cannot be caused by changes in the real value of the minimum wage. Secondly, several studies (such as Card and DiNardo, 2002) find that 1988 or 1989 is a turning point in the evolution of wage inequality. Wage inequality appears to have stabilized and no noticeable change can be seen between 1989 and 2001. Finally, the period 1973-1989 offers the possibility of comparing the results of the proposed estimator with the numerous empirical works covering this period.

The measure of wage we use is the hourly wage of those workers that are paid on an hour-basis and usually weekly earnings divided by usually hours of work for the others. Allocated earnings are excluded because they can bias the results, in particular for distributional analysis since the mean value of the wage given the covariates is imputed. Unfortunately, flags indicating which observations are allocated are not available in 1989. However, Hirsch and Schumacher (2004) explain how it is possible to identify allocated earners by using unedited weekly earnings. We use their method to exclude allocated earnings. We use a broad sample of male workers but weight observations by the product of the CPS sample weight with usual hours of work. This yields a better representation of the distribution of wages for each and every hour worked in the USA in the given years. We deflate wages to 1982-84 dollars using the CPI-U. Only men of age 16 to 65 and reporting an hourly wage above 1$ are kept in the sample. All observations with at least one missing for one of the variables are excluded. The sample sizes are approximately 23'000 in 1973 and 75'000 in 1989.

The vector of regressors x consists of a quartic in potential experience (defined as $\max(0, age-5-years\ in\ school)$), 11 education dummies, 6 interaction terms between education and experience[3], a part-time dummy, union status, 5 race dummies, 3 region dummies, and 17 industry dummies. The mean of all covariates in both periods can be found in the first two columns of table 1. The level of potential experience decreased between 1973 and 1989 because of the entry of the baby-boom generation into the labor market and because of longer education. Educational attainment increased clearly over the period. For instance, the percentage of workers with a college degree increased by 60%. As is well known, de-unionization was impressing with a 11.4% fall in union members.

[3] Only interaction terms that were significant in at least one of both periods were kept as regressors.

3.3 First step quantile regression results

Since some of the earnings are top coded, the conditional distributions have been estimated with censored quantile regressions. We have used the 3-step censored quantile regression algorithm suggested by Chernozhukov and Hong (2002). Their estimator requires a separation restriction on the censoring probability that costs a small reduction in generality but preserves the plausible semiparametric, distribution-free and heteroscedastic features of the model. It has the advantage of being easily computable. Because of the number of observations, it is simply not possible to estimate the whole quantile regression process. Therefore, we have estimated 200 different quantile regressions uniformly distributed between 0 and 1.

We consider first the coefficients of the median regressions in the third and fourth columns of table 1. Points and stars indicate significant differences from zero, with standard errors estimated by bootstrapping the results 100 times. The coefficients have generally the expected signs and are conform to previous studies. The negative public sector wage differential is surprising but it is compensated by the high positive coefficient on the public administration sector, where about 40% of public sector employees work. Between 1973 and 1989, we note that the wage-experience profile changed. The linear and cubic terms decreased but the quadratic and quartic terms increased. Therefore, the return to experience decreased for levels of experience below 10 years and increased above. Returns to education, particularly for persons with a degree higher than high school, increased.

Coefficients of the median regression indicate how the level of wages depends on covariates. To analyze the effects of characteristics on the dispersion of earnings, the fifth and sixth columns of table 1 illustrate the difference between the quantile regression coefficients at the 9^{th} decile and the coefficients at the 1^{st} decile. If the error term is independent of a characteristic, the coefficient on this variable does not vary with the quantile and thus $\hat{\beta}(0.9)-\hat{\beta}(0.1)$ should not be significantly different from zero. If the difference between the

9^{th} and 1^{st} decile coefficient on a covariate is positive (negative), a higher value of this variable increases (decreases) within-group inequality. The results show that for more than half of the variables the interdecile difference is significantly different from zero. Thus, heteroscedastic inconsistent methods will yield biased results. Consistent with Mincer (1974), the within-group inequality grows as a function of experience[4] and the interdecile range even increased between both periods. The within-group inequality also tends to grow as a function of education with low or negative interdecile differences at low educational levels and significant positive interdecile differences at the highest levels. As is well-known in the literature, the variance of wages is significantly lower for union members and this difference increased between 1973 and 1989. Finally, we observe that the within-group inequality differs between sectors of employment.

Given that within-group inequality depends on characteristics, changes in characteristics will affect overall inequality. A straightforward example helps to understand the importance of this composition effect. In order to keep the illustration simple, we consider first only $\overline{x}^{73}$ and $\overline{x}^{89}$, the means of the characteristics distributions in 1973 and 1989. In the quantile regression framework, the interdecile range in 1973 evaluated at $\overline{x}^{73}$ can be estimated by $\overline{x}^{73}\left(\hat{\beta}^{73}(0.9)-\hat{\beta}^{73}(0.1)\right)$. Using the results of table 1, we get a value of 0.883 for the interdecile range. Now, if we evaluate the interdecile range in 1973 at $\overline{x}^{89}$, we obtain a value of 0.934. This increase cannot possibly be due to changes in residuals or coefficients since they were kept unchanged. It has to arise from changes in characteristics such as the increase in the percentage of college and post college degrees, the de-unionization and the fall of employment in the manufacturing sector. Nevertheless, if we keep the mean characteristics at $\overline{x}^{73}$ but replace $\left(\hat{\beta}^{73}(0.9)-\hat{\beta}^{73}(0.1)\right)$ by their value in 1989, the estimated wage interdecile

[4] However, his model predicts first declining and then increasing residual variance as a function of experience. We do not find evidence for the first prediction.

range attains a level of 0.936. Thus, a part of the increase in within-group inequality can be attributed to residuals but this effect will be strongly overestimated if we do not correct for the composition effect.

3.4 Decomposition results

These first results indicate that the composition effect is potentially important. To consider also the effects of changes in coefficients and changes in the distribution of characteristics, the decomposition (2) proposed in the second section has been estimated. Figure 1 plots the decomposition results at each of the 999 per mill. Table 2 presents decomposition results for the median and for various measures of wage dispersion: the standard deviation and the 90-10, 90-50, 50-10, 75-25 and 95-5 gaps of the log wage distribution. Standard errors computed by bootstrapping the results 100 times are given in parentheses. For each statistic, the relative importance of each component of the decomposition is given in italic in the second row.

As documented in many other studies, there is a clear widening in the unconditional wage distribution over this time period. Real wages for workers at the 10^{th} percentile declined by about 20%, while they rose by about 2% for workers at the 90^{th} percentile. As a result, the standard deviation of log wages and all other measures of wage inequality such as the interquartile and the interdecile ranges increased significantly. Perhaps more surprising but totally consistent with the findings of previous studies is that the mean and the median real wage were lower in 1989 than in 1973.

The positive effect of characteristics on the median indicates that if workers' attributes had been rewarded the same in 1989 as in 1973, wages should have risen, not fallen, in 1989. The lower level of wages is explained by changes in coefficients, that is how workers characteristics are rewarded. This is mainly the consequence of a lower constant and not of lower return to human capital characteristics. Naturally, the effect of the residuals on the median is not significantly different from zero.

These results could have been obtained by the traditional Oaxaca / Blinder decomposition but much more interesting is the decomposition of indices of inequality. We note first that the six indices of inequality yield similar results about the relative importance of each component. Residuals account for about 20% (between 15 and 25%, depending on the statistic) of the increase of inequality. This is much less than what is widely accepted in the literature and was not subject of controversy until now[5]. The literature briefly surveyed in section 3.1 finds that residuals account for most of the increase in inequality. Coefficients account for about 30% (between 25 and 35%) of the growth in overall wage inequality. This is a standard result that is explained by the rise in the returns to education. It is not surprising that our estimates are almost the same as those of JMP because the methodologies are principally similar for what concerns the effects of coefficients. Finally, the most important part of the growth in inequality is explained by changes in the distribution of characteristics. This stands again in contradiction to the literature quoted above and is explained by the composition effect of characteristics on residuals. We have seen in table 1 that wage dispersion increases with experience and education but is smaller for union member and in the construction sector. Thus, the increase in education, the de-unionization and the fall of the occupation in the manufacturing sector do not only affect the level of wages but also increase the within-group inequality. Methods that do not account for the dependence between residuals and characteristics overstate the effects of residuals and understate the effects of characteristics (see also the discussions in Lemieux, 2002 and 2004, and Machado and Mata, 2004).

We note that the estimates are fairly precise if we take into account that we control for 48 independent variables and that it is more difficult to estimate changes in inequality than levels. In order to strengthen the conclusions, different robustness checks were performed. They are not given here in details but are available on the internet page of the author. First, the order of the decomposition has been changed. Since there are 3 components, we can

[5] However, recently Lemieux (2004) finds results very similar to ours with a different methodology.

imagine 6 different orders. All these 6 decompositions give similar results. Second, the number of quantile regressions in the first step of the estimation was set to 10, 100 and 400. The results are almost the same and absolutely not sensitive to this change.

Finally, we have estimated the decomposition over sub-periods: 1973-1979, 1979-1984 and 1984-1989. The overall inequality first decreased between 1973 and 1979, then increased a lot between 1979 and 1984 and increased also but slightly less fast in the third period. The effect of changes in characteristics onto inequality is positive over the whole period. It is logical that changes in characteristics are more continuous than other changes since it takes time to increase the level of education or experience, for instance. Changes in coefficients reduce inequality in the first period, they account for most of the increase in inequality during the second period and they are almost insignificant in the third period. Residuals have a moderate positive effect on inequality in all sub-periods. We observe that the effects of residuals and of coefficients seem to go into opposite directions during the first period, which would contradict the prediction of a single-index model of skills. It is also interesting to note that the effects of coefficients seem to be connected with the level of the minimum wage. As a matter of fact, over the 1973-1979 period, both the real value of the minimum wage and coverage rose substantially. On the contrary, the real value of the minimum wage decreased a lot from 1979 to 1989.

4 Conclusion

In this paper, we have proposed and implemented a flexible, intuitive and semiparametric estimator of distribution functions in the presence of covariates. The conditional wage distribution is estimated by quantile regression. Then, the conditional distribution is integrated over the range of the covariates to obtain estimates of the unconditional distribution. Counterfactual distributions can be estimated, allowing the decomposition of changes in distribution into three factors: changes in regression coefficients, changes in the distribution

of covariates and residuals changes. This decomposition is in the spirit of the JMP decomposition but it allows the covariates to influence the whole distribution of the dependent variable.

We have applied this methodology to US data for the period 1973-1989, a period during which earnings inequality increased quite dramatically. We find that about half of the increase in inequality can be explained by changes in the distribution of characteristics. Increases in the return to skills, particularly education, account also for a substantial proportion of the increase in inequality. On the contrary, changes in residuals account only for about 20% in the growth of inequality, suggesting that there was only a moderate increase in the price of unmeasured skills. These results, which are different from those of most other studies, show how important it is to allow the covariates to affect the whole residuals distribution and not only the first moment(s).

References

Acemoglu, D., 2002, Technical change, inequality, and the labor market, Journal of Economic Literature 49, 7-72.

Aghion, P., 2002, Schumpeterian growth theory and the dynamics of income inequality, Econometrica 70, 855-882

Blinder, A., 1973, Wage discrimination: reduced form and structural estimates, Journal of Human Resources 8, 436-455.

Card D. and J.E. DiNardo, 2002, Skill-Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles, Journal of Labor Economics 20, 733-783.

Chernozhukov V. and H. Hong, 2002, Three-Step Censored Quantile Regression and Extramarital Affairs, Journal of the American Statistical Association 97, 872-882.

DiNardo J.E., Fortin N.M. and T. Lemieux, 1996, Labor market institutions and the distribution of wages, 1973-1992: a semiparametric approach, Econometrica 65, 1001-1046.

Gosling A., Machin S. and C. Meghir, 2000, The changing distribution of male wages in the UK, Review of Economic Studies 67, 635-686.

Hirsch B.T. and E.J. Schumacher, 2004, Match Bias in Wage Gap Estimates Due to Earnings Imputation, Journal of Labor Economics 22, 689-722.

Juhn C., Murphy K.M. and B. Pierce, 1993, Wage inequality and the rise in returns to skill, Journal of Political Economy 101, 410-442.

Katz L. and D. Autor, 2000, Changes in the wage structure and earnings inequality, in: O. Ashenfelter and D. Card, eds., Handbook of labor economics, Vol. 3a (Elsevier, Amsterdam) 1463-1555.

Koenker R. and G. Bassett, 1978, Regression Quantiles, Econometrica 46, 33-50.

Lemieux T., 2002, Decomposing changes in wage distributions: a unified approach, Canadian Journal of Economics 35, 646-688.

Lemieux T., 2004, Increasing residual wage inequality: composition effects, noisy data, or rising demand for skill?, working paper, University of British Columbia

Machado J. and J. Mata, 2004, Counterfactual decompositions of changes in wage distributions using quantile regression, Journal of Applied Econometrics, forthcoming.

Melly B., 2004, Decomposition of differences in distribution using quantile regression, mimeo, downloadable from www.siaw.unisg.ch/lechner/melly.

Mincer J., 1974, Schooling, Experience, and Earnings (NBER, New York).

Oaxaca R., 1973, Male-female wage differentials in urban labor markets, International Economic Review 14, 693-709.

Portnoy S., 1991, Asymptotic behavior of the number of regression quantile breakpoints, SIAM Journal of Scientific and Statistical Computing 12, 867-883.

Portnoy S. and R. Koenker, 1997, The Gaussian Hare and the Laplacian Tortoise: Computability of Squared-Error versus Absolute-Error Estimators, Statistical Science 12, 279-300.

R Development Core Team, 2003, R: A language and environment for statistical computing (R Foundation for Statistical Computing, Vienna).

Table 1: Mean of the covariates, median regression coefficients and interdecile ranges

Variable	Mean		$\hat{\beta}(0.5)$		$\hat{\beta}(0.9)-\hat{\beta}(0.1)$	
	1973	1989	1973	1989	1973	1989
Constant	1	1	0.888**	0.784**	0.956**	0.793**
Experience	19.98362	18.02282	0.072**	0.065**	0.028**	0.045**
Experience squared	588.6353	462.8457	-0.003**	-0.002**	-0.001·	-0.003**
Experience to the power of 3	20414.70	14303.18	5.17E-5**	4.05E-5**	3.6E-5·	6.90E-5**
Experience to the power of 4	774768.1	495823.1	-3.54E-7**	-2.84E-7**	-3.54E-7·	-6.84E-7**
Grade1to4	1.78%	0.83%	0.094	0.053	0.009	-0.064
Grade5to6	2.45%	1.52%	0.032	0.039	0.036	-0.199
Grade7to8	7.98%	2.49%	0.125	0.123*	0.112	-0.098
Grade9	3.92%	2.02%	0.210*	0.189**	0.179	-0.059
Grade10	6.44%	3.50%	0.336**	0.298**	0.000	-0.009
Grade11	6.02%	3.90%	0.352**	0.396**	0.044	-0.054
Grade12	2.63%	1.99%	0.421**	0.445**	0.062	-0.032
High school	34.17%	34.76%	0.470**	0.468**	0.098	0.075
Some college	17.99%	23.34%	0.593**	0.626**	0.186	0.146
College	9.03%	14.44%	0.837**	0.942**	0.239	0.216·
Post-college	7.28%	10.94%	1.021**	1.116**	0.360·	0.339**
Experience*grade5to6	0.849151	0.419983	0.004*	0.000	-0.004	0.004
Experience*grade7to8	2.665141	0.787266	0.003**	0.002	-0.003	0.003
Experience*grade9	1.024756	0.470180	0.002·	0.002	-0.005·	0.001
Experience*grade11	1.161559	0.653538	0.001	-0.002**	-0.001	0.002
Experience*grade12	0.346608	0.247193	-0.001	-0.003**	-0.001	0.004
Experience*college	1.369628	2.245673	0.005**	0.001	0.001	0.003·
Part-time	5.16%	5.97%	-0.126**	-0.174**	0.090**	0.050*
Union member	31.61%	20.21%	0.148**	0.179**	-0.109**	-0.164**
Black non-Hispanic	8.31%	9.04%	-0.138**	-0.139**	0.013	0.000
Mexican	2.86%	5.85%	-0.145**	-0.138**	-0.031	0.010
Other Hispanic	1.87%	2.91%	-0.118**	-0.128**	-0.017	0.070*
Other non-white	1.23%	2.86%	-0.109**	-0.128**	0.244·	0.067*
Northeast	22.21%	19.55%	0.089**	0.125**	-0.048*	-0.023
Midwest	28.03%	25.40%	0.103**	0.021**	-0.058**	-0.017
West	18.35%	20.89%	0.125**	0.128**	-0.080**	0.020
Public sector employee	15.82%	14.64%	-0.038·	-0.071**	-0.004	-0.113**
Construction	9.95%	9.62%	0.453**	0.305**	-0.264**	-0.210**
Manufacturing durable goods	21.67%	17.25%	0.266**	0.250**	-0.484**	-0.304**
Other manufacturing	11.27%	9.30%	0.227**	0.201**	-0.447**	-0.229**
Transport	5.35%	6.47%	0.362**	0.245**	-0.356**	-0.155**
Utilities services	3.72%	3.71%	0.336**	0.352**	-0.478**	-0.297**
Wholesale trade	4.95%	5.40%	0.258**	0.128**	-0.371**	-0.203**
Retail trade	12.75%	13.68%	0.073*	-0.018	-0.344**	-0.167**
Finance	3.66%	4.61%	0.283**	0.264**	-0.284**	-0.071
Business services	3.09%	5.76%	0.167**	0.112**	-0.251**	-0.059
Personal services	1.02%	1.54%	0.004	-0.062·	-0.292**	-0.072
Entertainment services	0.77%	1.04%	0.111·	-0.027	-0.015	-0.110·
Health services	0.63%	0.96%	0.229**	0.062·	-0.333**	-0.047
Hospitals	1.73%	1.97%	0.121*	0.071*	-0.362**	-0.098
Educational services	5.49%	5.28%	0.081*	-0.017	-0.494**	-0.195**
Social services	1.11%	0.66%	-0.487**	-0.129**	0.229	-0.172*
Other professional services	1.59%	3.07%	0.348**	0.164**	-0.425**	0.111·
Public administration	7.01%	5.86%	0.365**	0.246**	-0.376**	-0.192**

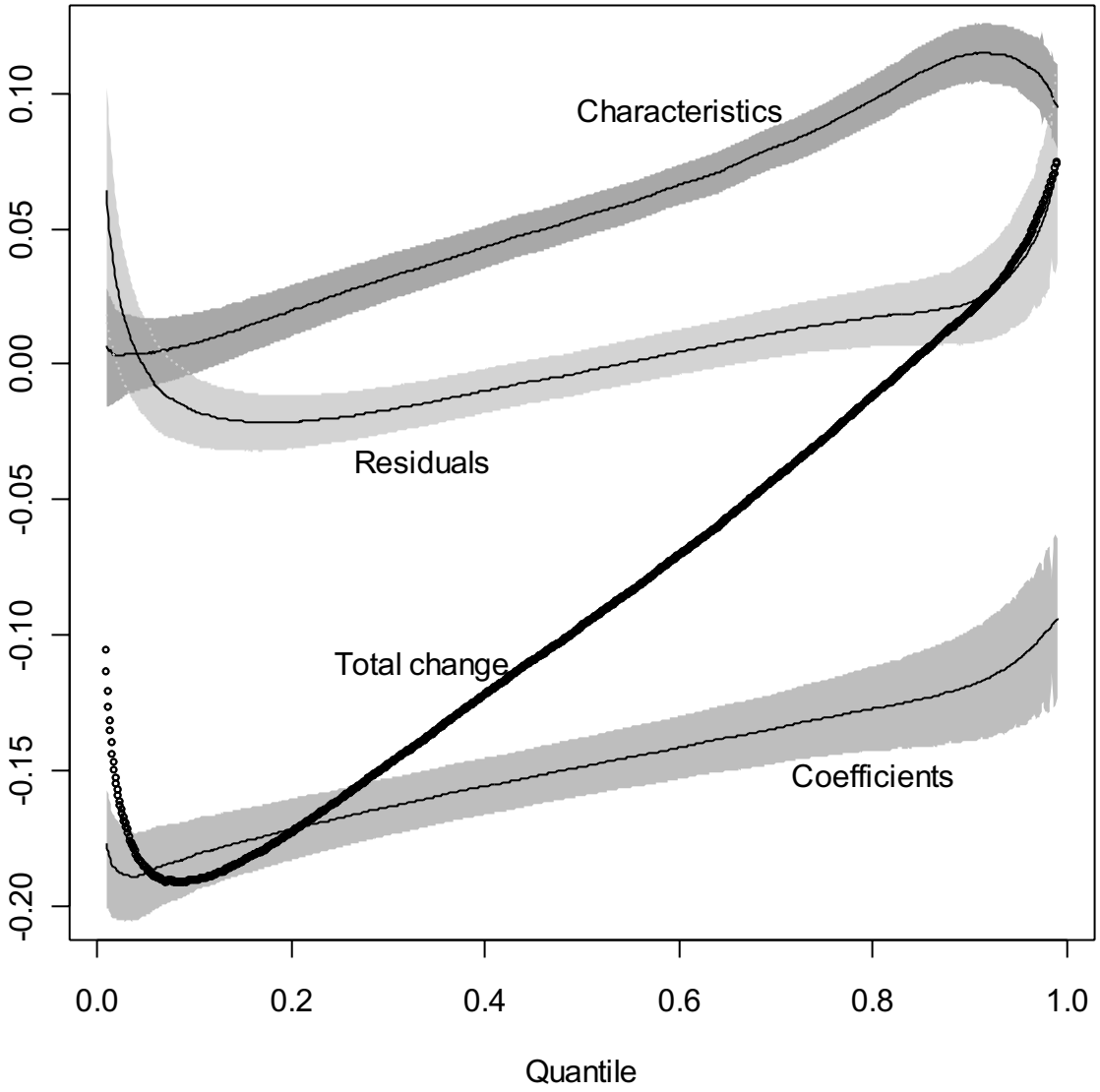
Bootstrap with 100 replications. ·: significant at the 5%, *: significant at the 1%, **: significant at the 0.1%.

Table 2: Decomposition of changes in measures of wage dispersion using quantile regression

Statistic	Total change	*Effects of:*		
		residuals	coefficients	characteristics
Median	-9.71 (0.45) *100%*	-0.3 (0.31) *3.08% (3.22)*	-14.85 (0.42) *152.87% (6.23)*	5.43 (0.29) *-55.94% (4.92)*
Standard deviation	6.76 (0.36) *100%*	1.2 (0.31) *17.79% (4.43)*	2.22 (0.31) *32.95% (4.24)*	3.33 (0.18) *49.26% (2.81)*
90-10	20.94 (0.9) *100%*	4.01 (0.71) *19.14% (3.23)*	6.26 (0.89) *29.89% (3.6)*	10.68 (0.52) *50.97% (2.62)*
50-10	9.36 (0.55) *100%*	1.42 (0.46) *15.14% (4.37)*	3.28 (0.48) *35.1% (4.1)*	4.66 (0.31) *49.77% (3.24)*
90-50	11.59 (0.64) *100%*	2.59 (0.53) *22.37% (4.34)*	2.98 (0.54) *25.68% (4.47)*	6.02 (0.36) *51.94% (3.81)*
75-25	13.32 (0.48) *100%*	3.42 (0.39) *25.69% (2.68)*	3.7 (0.43) *27.76% (2.98)*	6.2 (0.31) *46.55% (2.41)*
95-5	22.55 (1.19) *100%*	3.81 (1.08) *16.91% (4.64)*	7.91 (1.16) *35.06 (4.34)*	10.83 (0.63) *48.03% (2.83)*

Note: all numbers have been multiplied by 100. Bootstrap standard errors with 100 replications in parentheses.

Figure 1: Decomposition of differences in distribution using quantile regression



Note: Decomposition results obtained by applying formula (2) at each of the 999 per mills. Bootstrap standard errors with 100 replications.