Labour Economics 28 (2014) 58–69

Contents lists available at ScienceDirect

Labour Economics

journal homepage: www.elsevier.com/locate/labeco

The contribution of human capital variables to changes in the wage distribution function ☆,☆☆

Pallab Kumar Ghosh

Department of Economics, University of Oklahoma, 308 Cate Center Drive, Norman, OK 73019, United States

HIGHLIGHTS

- This paper proposes a modified wage decomposition method.
- I show the asymptotic properties of the proposed decomposition method.
- Empirical results of the proposed method are consistent with the previous studies.
- The proposed decomposition method can also be used in different applications.

ARTICLE INFO

Article history:
Received 1 December 2012
Received in revised form 6 March 2014
Accepted 14 March 2014
Available online 25 March 2014

JEL classification:
C14
C21
J31

Keywords:
Conditional quantile function
Box–Cox quantile regression
Oaxaca decomposition
Wage distribution function

ABSTRACT

This paper proposes a semi-parametric procedure to determine the contribution of human capital variables in explaining the changes of U.S. wage distribution function over the period of 1990–2000. The effects of these factors are estimated by using the Chamberlain's two stage Box–Cox quantile regression approach. One of the main contributions of this paper is that it relaxes the linearity assumption of the conditional quantile function to estimate the counterfactual wage distribution function consistently. This paper also shows that the proposed method provides better estimates of capturing the effects of human capital variables in the two tails of the wage distribution while the results of other parts of the distribution are comparable with the estimates of the previous study which used the linear quantile regression approach. To summarize, these results suggest that this proposed approach is more applicable in cases where the behavior of the tails of the distribution bear much importance.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Developing new decomposition methods to explain the changes in the wage distribution function has been an extensive research area for the last 20 years. The main reason for the continued interest in this topic is that wage inequality in several countries, specifically in the United States, has increased sharply since the early 1980s. Autor et al. (2008) show that the slowing of the growth of overall wage inequality in the 1990s hides a divergence in the paths of upper-tail (90/50) and lower-tail (50/10) inequality. The upper-tail inequality has increased steadily since 1980, even adjusting for changes in the labor force composition; lower-tail inequality rose sharply in the first half of the 1980s and contracted thereafter. In this paper, I focus on the role human capital variables play in causing changes in the U.S. wage distribution function during the period of 1990 to 2000.

This paper proposes an extension of the decomposition method proposed by Machada and Mata (2005). The aim is to estimate a counterfactual distribution function F, which is of the form

$$F(w) = \int G(w|x)dH(x).$$

Here $G(\cdot)$ is the conditional Cumulative Distribution Function (CDF) of wage (w) given the covariates (X) in period $t = 0$ and H is the unconditional CDF of X in period $t = 1$. Machada and Mata (2005) estimate the

☆ I am especially thankful to my advisor, Professor Thomas Kniesner, for his guidance and suggestions. I also thank Professor Jeffrey Kubik, Professor Badi Baltagi, Professor Jing Li, Dr. Alissa Dubnicki, and the seminar participants at the 24th Annual Conference of the European Association of Labour Economists, for their helpful comments and suggestions. Any errors or omissions are my own.
☆☆ I am very grateful to Professor David Autor for his comments.
E-mail address: pallab.ghosh07@gmail.com.

http://dx.doi.org/10.1016/j.labeco.2014.03.006
0927-5371/© 2014 Elsevier B.V. All rights reserved.

inverse of G through a linear quantile regression model and estimate the integral through a simulation method. Machada and Mata (2005) point out that one of the limitations of their approach is that the linearity assumption of the conditional quantile function may be violated in many scenarios when applying real life data. This paper shows that the linearity assumption does not hold when $G(\cdot)$ is a very skewed distribution function.

Barsky et al. (2002) show that the Oaxaca-Blinder (1973) decomposition, based on the linear regression model, yields biased estimates of the decomposing terms when the underlying conditional expectation of Y given X is nonlinear. The Barsky et al. (2002) results can be extended in the quantile regression framework. This extension implies that the Machada and Mata (2005) approach gives consistent results only where the conditional quantile function is approximately linear.

In this paper, I suggest a flexible nonlinear quantile regression model for G, namely the Chamberlain Buchninsky Box–Cox model. The main reason for proposing this model is that the shape parameter in the Box–Cox model adjusts with each quantile of the wage distribution function to estimate the conditional quantile function. I also develop the formal asymptotic properties of the proposed decomposition method and apply the method by using U.S. labor market data for the period of 1990–2000 from the American Community Survey.

Recently, there have been several papers that have studied decomposition methods, including Chernozhukov et al. (2013a), Rothe (2010a, 2012a, b), Firpo et al. (2011, 2007), Antonczyk et al. (2010), Altonji et al. (2008), Melly (2005), and DiNardo et al. (1996). (See also the extensive survey by Fortin et al. (2011) on the wage decomposition literature.) These papers show how the function G mentioned above can be estimated by a wide range of parametric, semi-parametric or non-parametric methods.

DiNardo et al. (1996) analyze the effect of a change in the conditional distribution of a single covariate given the remaining ones (see also Donald et al. (2000) and Gosling et al. (2000)). Hirano et al. (2002) and Firpo (2007, 2010) establish the efficiency of this estimation method. However, the DiNardo et al. (1996) method becomes practically infeasible when there are too many continuous variables. Moreover, reweighting can have some undesirable properties in small samples when there is a problem of common support. Frolich (2004) finds that reweighting estimators perform poorly in this context.

Rothe (2010a) proposes a fully non-parametric method to find the effect of an exogenous change in the distribution of covariates on the unconditional distribution of the outcome variable. In the wage decomposition literature, many researchers including Chernozhukov et al. (2013a), Firpo et al. (2011, 2007), and Machada and Mata (2005) propose parametric models. Using a correctly specified parametric model results in efficiency gains compared to the fully non-parametric method; however, such estimators are generally inconsistent when the parametric assumptions are violated.

Most recently, Rothe (2012a) measures the partial effect of a counterfactual change in the unconditional distribution of a single explanatory variable on some features of the unconditional distribution of Y in the context of a non-separable model. In a related paper, Firpo et al. (2009) estimate the impact of changing the distribution of explanatory variables X on the marginal quantiles of the dependent variable Y or other functionals of the marginal distribution of Y. By using either of the methods, one can directly calculate the distributional features. Firpo et al. (2009) identify their models only when X is exogenous. Rothe (2010b) estimates the unconditional partial effects in a nonseparable triangular model with endogenous regressors by using the control variable approach proposed by Imbens and Newy (2009).

Machada and Mata (2005) consider the effect of a change in the unconditional distribution of one of the covariates while holding constant the conditional distribution of the remaining covariates. Chernozhukov et al. (2013a) introduce the distributional regression model that corresponds to either changes in the distribution of the covariates, or changes in the conditional distribution of the outcome given covariates, or both. Chernozhukov et al. (2013a) argue that the distributional regression is an alternative to quantile regression. Thus, both the approaches provide similar results when the support of X is sufficiently rich.

Chernozhukov et al. (2013a) point out that Machada and Mata (2005) method works well if Y has a smooth conditional density and may provide a poor approximation of the conditional distribution function when Y is discrete or consists of mass points. In contrast, distributional regression does not require smoothness of the conditional density because the approximation is done pointwise in the threshold Y, and thus handles continuous, discrete, or mixed Y without any special adjustment.

This paper contributes to the previous literature by developing a flexible decomposition method to analyze the differences in the distribution of individuals' economic outcomes across two countries, time periods or subgroups of the population. In particular, the results can be used to make inference on the change of the distribution function as a whole, its moments, quantiles, and inequality measures such as the Lorenz curve or Gini coefficient. I also show that the estimated counterfactual distribution function of the proposed method converges to the true distribution function asymptotically. This asymptotical property of the proposed decomposition method establishes the validity of the inference procedure.

The empirical contribution of this paper is that the proposed decomposition method can precisely estimate the contribution of human capital variables especially in the two tails of the wage distribution function. In comparison with the proposed decomposition method results, I find that the Machada and Mata (2005) approach underestimates the price effects of labor market in the lower-tail and overestimates in the upper-tail of the U.S. wage distribution function for the period 1990–2000. Differences in the rest of the parts of the wage distribution function are not statistically significant.

The proposed decomposition method results show that there is a steady rise in upper-tail wage inequality and a flat or declining lower-tail wage inequality during the 1990s. The results are consistent with Autor et al. (2008). Autor et al. (2005) called this phenomena the 'polarization' of the U.S. labor market, with a particularly strong market for workers in the top parts of the skill distribution and, at the same time a deterioration in market conditions for workers in the middle, and reasonably steady market conditions for those near the bottom.

The structure of the paper is as follows. Section 2 describes the econometric model, the construction of the counterfactual wage distribution function, and the decomposition of the wage distribution function. In Section 3, I discuss the data, show the empirical results and also compare the results with the conventional wisdom of the wage inequality literature. I conclude in Section 4. All the proofs are shown in the Appendix.

2. Econometric model

The Mincer equation is the most widely used specification of the empirical earnings equation in the labor economics literature. This equation is typically specified in the linear functional form. A more general approach would be not to constrain the model to linear form but rather to let the data determine the exact transformation that results in a linear relationship between the dependent variable and the covariates. Box and Cox (1964) suggest such a model in a nonlinear setup.

2.1. Model setup

The econometric model I use in this paper specifies that the τth conditional quantile of w (wage), given x ($Quant_\tau(w|x)$), depends on a linear index, $x'\beta_\tau$, through a nonlinear function $g(\cdot)$, that is,

$$Quant_\tau(w|x) = g(x'\beta_\tau, \lambda_\tau) \tag{1}$$

where $g(\cdot)$ is strictly monotonically increasing in $x'\beta_\tau$ and also depends upon the parameter λ_τ. Moreover, $w > 0$, $x \in \mathbb{R}^k$ are observed, while the parameters $\beta_\tau \in \mathcal{B} \in \mathbb{R}^k$ and $\lambda_\tau \in \mathbb{R}$ are unknown and $\tau \in (0,1)$.

The specification of model (1) is quite flexible because it depends on the linear index $x'\beta_\tau$ and allows the whole transformation to change for each different quantile $\tau \in (0,1)$. Powell (1991) shows that the Box and Cox (1964) transformation can be estimated in a nonlinear quantile regression framework.

In the Box–Cox transformation, the inverse of $g(x'\beta_\tau, \lambda_\tau)$ is given by

$$w_\lambda = g^{-1}(w,\lambda) = \begin{cases} \left(w^\lambda - 1\right)/\lambda & if \lambda \neq 0 \\ log(w) & if \lambda = 0, \end{cases} \tag{2}$$

assuming $\lambda \in \mathbb{R}$ and $\mathbb{R} = [\underline{\lambda}, \overline{\lambda}]$ to be a finite closed interval. The advantage of using the Box–Cox transformation is that it preserves the ordering of the observation due to the equivariance property of quantiles with respect to the monotonically increasing w_λ. Therefore,

$$Quant_\tau\left(w_\lambda|x\right) = g^{-1}(Quant_\tau(w|x)). \tag{3}$$

Using the Box–Cox transformation as the inverse of $g(x'\beta_\tau, \lambda_\tau)$, Eq. (1) becomes

$$Quant_\tau\left(w|x\right) = \left(\lambda_\tau Quant_\tau(w_\lambda|x) + 1\right)^{1/\lambda_\tau} \tag{4}$$

Buchinsky (1995) shows that the Box–Cox transformation specified in Eq. (2) satisfies the property

$$Quant_\tau(w_\lambda|x) = x'\beta_\tau. \tag{5}$$

By using Eqs. (4) and (5), the conditional quantile function of w given x becomes

$$Quant_\tau\left(w|x\right) = \left(\lambda_\tau x'\beta_\tau + 1\right)^{1/\lambda_\tau}. \tag{6}$$

Note that the conditional quantile function, $Quant_\tau(w|x)$ becomes linear when $\lambda_\tau = 1$.

The estimation of β_τ and λ_τ is studied by Powell (1991), Chamberlain (1994), Buchinsky (1995) and Fitzenberger et al. (2010). Box–Cox quantile regression minimizes the distance function

$$\min_{\beta_\tau \in \mathcal{B}, \lambda_\tau \in \mathbb{R}} \frac{1}{n}\sum_{i=1}^{n} \rho_\tau\left(w_i - (\lambda_\tau x_i'\beta_\tau + 1)\right)^{1/\lambda_\tau}, \tag{7}$$

where the loss function is given by $\rho_\tau(u) = \tau|u|\mathbb{I}_{u\geq 0} + (1-\tau)|u|\mathbb{I}_{u<0}$ and $\mathbb{I}$ denotes the indicator function. Powell (1991) shows that these nonlinear estimators $\hat{\lambda}_\tau$ and $\hat{\beta}_\tau$ are consistent and asymptotically normal.

2.2. Chamberlain Buchinsky Two Step (CBTS) Estimation

Chamberlain (1994) and Buchinsky (1995) suggest the following two step procedure which exploits the equivariance property of quantiles:

Step 1 Estimate $\beta_\tau(\lambda_\tau)$ conditional on λ_τ by

$$\hat{\beta}_\tau(\lambda_\tau) = arg \min_{\beta_\tau \in \mathcal{B}} \frac{1}{n}\sum_{i=1}^{n} \rho_\tau\left(w_{\lambda_i} - x_i'\beta_\tau\right) \tag{8}$$

Step 2 Estimate λ_τ by solving

$$\min_{\lambda_\tau \in \mathbb{R}} \frac{1}{n}\sum_{i=1}^{n} \rho_\tau\left(w_i - \left(\lambda_\tau x_i'\hat{\beta}_\tau(\lambda_\tau) + 1\right)^{1/\lambda_\tau}\right) \tag{9}$$

Chamberlain (1994) shows the asymptotic distribution of the CBTS estimators. Buchinsky (1995) derives large sample properties of this estimator for discrete regressors when applying the minimum distance method.

Fitzenberger et al. (2010) address a numerical problem for implementing the CBTS estimation method. They argue that the estimation problem given in Eq. (9) can only be solved if

$$\lambda_\tau x_i'\hat{\beta}_\tau\left(\hat{\lambda}_\tau\right) + 1 > 0 \tag{10}$$

for all $i = 1, 2, \ldots, n$ and for all $\lambda_\tau \in \mathbb{R}$. Fitzenberger et al. (2010) suggest to solve this problem by incorporating an indicator function in the second step of the estimation. The indicator function takes a constant value whenever the condition in Eq. (10) fails to satisfy. This modification guarantees that the objective function is a well defined sum from 1 to n. Fitzenberger et al. (2010) show that the modified estimator has similar asymptotic properties to the original CBTS estimator.

2.3. Test for Box–Cox conditional distribution model

In this subsection, I test the correct specification of the conditional distribution model using the Generalized Conditional Cramer-von Mises (GCCM) test proposed by Rothe and Wied (2013). Let $\mathcal{F}$ be the class of all conditional distribution functions on the support of Y given X that satisfy all the regularity conditions shown by Rothe and Wied (2013). We consider a parametric family

$$\mathcal{F}^0 = \left\{F^{-1}(\tau|x) = g(x'\beta_\tau, \lambda_\tau) for \quad some \ \ \theta \in \mathcal{B}(\tau, \Theta)\right\} \in \mathcal{F} \quad \forall \tau \in (0,1)$$

The hypothesis that Rothe and Wied (2013) propose to test is that $\mathcal{F}$ coincides with an element of $\mathcal{F}^0$.

$$H_0 : F(y|x) = F(y|x, \theta_0) \ \ for \ \ all \ \ (y,x) \in \mathbb{R}^{k+1}$$
$$H_1 : F(y|x) \neq F(y|x, \theta_0) \ \ for \ \ some \ \ (y,x) \in \mathbb{R}^{k+1}$$

The test statistics T_n and the bootstrap α critical values $\hat{c}_n(\alpha)$ have been computed by following the procedure shown by Rothe and Wied (2013) where $\hat{c}_n(\alpha)$ is the smallest constant that satisfies

$$P_b\left(T_{b,n} \leq \hat{c}_n(\alpha)\right) \geq 1 - \alpha,$$

and P_b is the probability with respect to bootstrap sampling and $T_{b,n}$ is the bootstrap test statistics. Rothe and Wied (2013) test rejects H_0 if $T_n > \hat{c}_n(\alpha)$ for some pre-specified significance level $\alpha \in (0,1)$.

2.4. Counterfactual wage distribution function

I use CBTS estimators to construct the counterfactual wage distribution. The steps are as follows:

1. Draw random sample of size m from $\tau \in (0,1) : \tau_1, \ldots, \tau_m$.
2. For each $\tau \in (0,1)$, using the CBTS estimators, estimate the following conditional quantile function

$$\hat{Q}_{g,\tau} = g\left(x'\hat{\beta}_\tau, \hat{\lambda}_\tau\right) = \left(\hat{\lambda}_\tau x_t'\hat{\beta}_\tau + 1\right)^{1/\hat{\lambda}_\tau}$$

3. Use the estimated result to draw a random sample from $\{w^i_{t=1}\}^m_{i=1}$, $\{w^i_{t=0}\}^m_{i=1}$ and $\{w^{c,i}_{t=1}\}^m_{i=1}$ where

$$\left\{w^i_t\right\}^m_{i=1} = F^{-1}_{w_t|x_t}(\tau_i, x_t) = g\left(\left(x^i_t\right)'\hat{\beta}_\tau, \hat{\lambda}_\tau\right); \quad t = 0, 1$$
$$\left\{w^{c,i}_{t=1}\right\}^m_{i=1} = F^{-1}_{w_{t=1}|x_{t=1}}\left(F_{w_{t=0}|x_{t=0}}(w_{t=0}, x_{t=0}), x_{t=0}\right).$$

Here $w^{c,i}_{t=1}$ denotes the ith draw from the counterfactual wage distribution function. $F^{-1}_{w_t|x_t}(\tau_i, x_t)$ are the τ_ith conditional quantile functions for the periods $t = 0,1$. $F^{-1}_{w_{t=1}|x_{t=1}}(\cdot)$ is the τ_ith counterfactual conditional quantile function.

The counterfactual wage distribution function consists of two parts. The first operator $F^{-1}_{w_{t=1}|x_{t=1}}$ ensures that the market return of the covariates remains fixed at period $t = 1$ whereas the second part $\left(F_{w_{t=0}|x_{t=0}}(w_{t=0}, x_{t=0}), x_{t=0}\right)$ implies that the distribution of covariates remains constant at period $t = 0$.

2.5. Decomposition of changes in wage distribution

The main objective of constructing a counterfactual wage distribution function is to use it to decompose differences in distributions. I use the extended Oaxaca (1973) decomposition method like DiNardo et al. (1996), Machada and Mata (2005), and Firpo et al. (2007, 2011) to decompose the differences in the distribution of wages in the years 1990 and 2000.

The main analysis of changes in the distribution functions of wage between the period $t = 0$ (year 1990) to period $t = 1$ (year 2000) is as follows:

$$F(w_{t=1}) - F(w_{t=0}) = \underbrace{[F(w_{t=1}) - F(w_{t=1}; x_{t=0})]}_{\text{Composition effect}} + \underbrace{[F(w_{t=1}; x_{t=0}) - F(w_{t=0})]}_{\text{Wage structure effect}} + Residual$$

The first bracket represents the effect of changes in the distribution of covariates which is denoted as the composition effect. The second bracket represents the effect of changes in the market returns of the human capital variables which we call the wage structure effect, and the rest is called the residual effect. The advantage of this decomposition is that we can decompose all the standard summary measures such as Gini coefficients, Theil coefficients, skewness, etc., since we can estimate the whole counterfactual distribution function.

Similarly, we can also measure the impact of individual covariates $(\tilde{x}_t)$ by considering the following equation

$$\text{Impact of } \tilde{x}_t \equiv F(w_{t=1}) - F(w_{t=1}; x'_{t=1}, x_{t=0}),$$

where $x_t = [\tilde{x}_t | x'_t]$ and x'_t are all the other covariates except $\tilde{x}_t$. The counterfactual distribution function $F(w_{t=1}; x'_{t=1}, x_{t=0})$ can be recovered from $F(w_{t=1}; x_{t=0})$ by assuming that $\tilde{x}_t \in x_t$ follows independent distribution.

The independence assumption allows us to write the joint distribution of the covariates as the product of the marginal distributions. The assumption may not hold in different scenarios. However, we can relax this assumption and estimate $F(w_{t=1}; x'_{t=1}, x_{t=0})$ by using either the fully non-parametric or the spline method. The tradeoff is a relatively small marginal efficiency gain at the cost of computational complexity.

This decomposition is quite attractive since the measurement of the impact does not depend on the ordering, unlike the DiNardo et al. (1996) method. Autor et al. (2005) show that this decomposition provides a precise link between the 'full variance accounting' technique for analyzing inequality introduced by Juhn et al. (1993) and the kernel reweighting proposed by DiNardo et al. (1996). However, Rothe (2012b) points out that most of the previous decomposition methods, including Machada and Mata (2005), are not additive in the sense that the sum of the effects of all individual covariates is not equal to the effect of changing all covariates.

The preciseness of the estimated impact of human capital variables of the proposed method depends on the total number of quantiles at which we estimate the conditional quantile function. In this paper, I take 3000 quantiles between 0 and 1 to construct the unconditional wage distribution function. Theoretically, the more quantiles the better, but one of the drawbacks of the proposed decomposition method is that it is computational intensive and the computational time is greater compared to the simulation method introduced by Machada and Mata (2005).

2.6. Asymptotic properties of the decomposition method

I show the following two asymptotic properties of the proposed decomposition method. The proofs of the properties are shown in the appendix.

(i) Complete mappings Assume that the conditional cumulative distribution function of w_t given x_t, $F_{w_t|x_t}(w_t|x_t)$, is monotonically increasing over the domain of $w_t \in \mathcal{W}_t$ and $\mathcal{W}_t$ is closed and bounded in R^k for $t = 0$ and 1, where k is the row dimension of $\mathcal{W}_t$. Then we have

$$F(w_t) \equiv \left\{F^{-1}_{w_t|x_t}(\tau_i)\right\}_i^m as \quad m \to \infty,$$

where m is the number of draws.

(ii) Convergence in distribution Let $\hat{\theta}^c$ be the quantile estimator that minimizes the counterfactual conditional quantile function of w given x, $g^c(\theta^c)$. We define a random variable $Y = g(x'_{t=0}\beta^{t=1}_\tau, \lambda^{t=0}_\tau)$ and assume that the distribution function of Y, $F(x'_{t=0}\beta^{t=1}_\tau, \lambda_{t=0})$ exists. Let $\hat{Y}_n = g_n\left(x'_{t=0}\hat{\beta}^{t=1}_\tau, \hat{\lambda}^{t=0}_\tau\right)$ have a distribution function $Fn(\cdot)$ for each n. Then, under the regularity conditions as shown by Powell (1991), we obtain

$$F_n\left(\hat{Y}_n\right) d \to F(Y) as \quad n \to \infty.$$

3. Data and results

The data are drawn 1% self weighted 1990 and 2000 samples from the American Community Survey. The data consists of U.S.-born black and white men aged 40–49 with at least 5 years of education, positive annual earnings and hours worked in the year preceding the census, and a nonzero sampling weight. The wage measure used is a weekly wage measure computed by dividing the annual earnings by the number of weeks worked in the previous year. I have restricted my sample sizes to 30,000 for both the years 1990 and 2000 to reduce the computational time for the simulation method used in the empirical analysis. The regressors consist of years of schooling and other basic controls.[1] Potential experience for men has been calculated by subtracting 6 and number of schooling years from the age variable.

Table 1 shows between group changes in wages and labor force composition in the samples 1990–2000 for groups defined by education, potential experience and race. The third column indicates that there are about 1 to 2% more high school, some college and college graduates and 2 to 3% less high school dropout and post college graduate workers in the year 2000 compared to 1990. These differences imply that although the average education level remains almost the same over the period 1990 to 2000, the relative fraction of the low skilled workers has decreased. The sixth column indicates that the average racial wage gap has fallen to approximately 2% in our samples.

In Table 2, I report the empirical rejection probabilities from n replications of the simulation process discussed by Rothe and Wied (2013), where the values of n are 50, 100, and 200. We see that for the year 2000, the linear quantile regression model leads to an adequate fit of the conditional male wage distribution data because we reject the null hypothesis of linear specification. We also note that the linear quantile regression model rejection rates are always higher than the Box–Cox quantile regression models at both 10% and 5% critical levels. This implies that the linear quantile regression model specification used in

[1] In the regressors, I have excluded various determinants of earnings such as tenure and occupation dummies because these variables are potentially endogenous and determined by education itself (Angrist and Pischke (2009)). Annual income is expressed in 1989 dollars using the Personal Consumption Expenditures Price Index.

Table 1
Descriptive statistics (percentage and average wage) for the period 1990–2000 by education, experience and race.

	Percentage			Mean wage		
	2000	1990	Changes	2010	2000	Changes
Education						
HS dropout	6.243	9.253	−3.010	6.055	6.025	0.030
HS grad	24.820	23.643	1.177	6.265	6.272	−0.007
Some college	34.420	32.403	2.017	6.464	6.437	0.026
College grad	20.177	18.623	1.553	6.860	6.715	0.090
Post college grad	14.340	16.077	−1.737	7.011	6.869	0.141
Experience						
Exp 0–10 years	14.290	19.350	−5.060	6.919	6.737	0.181
Exp 11–20 years	82.537	75.243	7.293	6.484	6.439	0.045
Exp 20–35 years	3.173	5.407	−2.233	6.179	6.157	0.022
Race						
White	92.177	92.123	0.053	6.562	6.509	0.053
Black	7.823	7.877	−0.053	6.234	6.161	0.073

the empirical analysis is not the best model specification for our sample data.

Rothe and Wied (2013) point out that rejection of the null hypothesis does not directly imply that such specifications result in misleading conclusions. However, for large samples Rothe and Wied (2013) argue that the Generalized Conditional Cramer-von Mises (GCCM) test is able to pick up deviations from the null hypothesis even if they are not economically significant in magnitude. Thus, by using the Box–Cox Quantile Regression model we can obtain more precise results at those quantiles of the wage distribution function for which the linearity assumption in parameters is a poor approximation of the conditional quantile function. Moreover, the linear and log-linear Quantile Regression models are special cases of the Box–Cox Quantile Regression model.

3.1. Overall changes in wage distribution

The top part of Fig. 1 shows how the probability density function changes from 1990 to 2000 and the bottom panel shows the same for the U.S. wage distribution function. We see that the area under the wage density curve of the upper-tail in 2000 is greater in comparison to 1990. This difference implies that the employment opportunities for the high-skill workers have gone up compared to the moderately and low-skill workers during the 1990s in the U.S. labor market. The result is in line with the Autor et al. (1998) findings on relative demand growth for the high-skill workers during the 1990s.[2]

The bottom part of Fig. 1 represents the fraction of workers' wage less than a given wage. By comparing the cumulative distribution functions of both the years, we see that employment shares of the very low and high-skill workers increased while employment shares for the moderately skilled workers contracted during the 1990s. Firpo et al. (2011) argue that the recent changes in task-based occupational structure play a very important role to explain this phenomenon. Acemoglu and Autor (2010) propose a task-based framework for analyzing the allocation of skills to task for studying the effect of new technologies on the labor market and their impact on the distribution of earnings.

Table 3 shows the changes in the estimated unconditional wage distribution by using both the quantile regression and the Box–Cox quantile regression simulation methods. In the seventh row, the term 'Scale' refers to the percentage change in the wage gap between the 75th quantile and 25th quantile. The random samples from the wage distribution functions have been generated through the steps described in Section 2, with the number of replications equal to 3000. The numbers in the parentheses indicate the 95% bootstrap confidence interval which has been generated by 1000 replications.

[2] Card and DiNardo (2002), Autor et al. (2003) emphasize more on the role of Skill Biased Technological Changes (SBTC) in the US wage structure.

Table 2
Conditional distribution model specification test for year 2000*.

	GCCM-QR		GCCM-BCQR	
No of replications	10%	5%	10%	5%
$n = 50$	0.303	0.157	0.080	0.067
$n = 100$	0.214	0.147	0.067	0.037
$n = 200$	0.113	0.085	0.045	0.017

* By using the Generalized Conditional Cramer-von Mises (GCCM) test proposed by Rothe and Wied (2013), we test two functional forms of the conditional quantile function for the 5 and 10% critical values. The numbers in the table represent the rejection probabilities from n replications of the simulation process proposed by Rothe and Wied (2013).

By comparing the third and sixth columns of Table 3, we note that the linear quantile regression and the Box–Cox quantile regression models give us statistically different results in the upper-tail of the male wage distribution. The proposed decomposition method estimates the changes in wages as 9 to 23% in the upper-tail of the U.S. wage distribution function while the Machada and Mata (2005) method estimates are 1 to 7%. However, both Machada and Mata (2005) decomposition method and the proposed approach show that there are marginal changes in wages in the lower-tail. This difference at the upper-tail is because the linearity assumption in the parameters of the conditional quantile function is unable to capture the precise changes in the upper-tail of a very skewed wage distribution function.

Our results on U.S. wage inequality are qualitatively similar to Autor et al. (2008), Goos and Manning (2003) and Autor et al. (2003). This pattern of changes in the U.S. wage distribution in the 1990s differs sharply compared to the 1980s, with more rapid growth of real wages at the top of the wage distribution relative to the middle and the bottom of the wage distribution. Goos and Manning (2003) and Autor et al. (2005) characterize this pattern as the polarization of the U.S. labor market. This pattern suggests a central role for labor demand shifts in explaining changes in the U.S. wage distribution function during the 1990s.

3.2. Decomposition method results

In Table 4, the wage distribution function has been decomposed into three parts, namely composition effect, wage structure effect and residual effect. Row by row, I report the estimates of these three effects for different quantiles and various inequality measures. The proposed decomposition method results show that the wage structure effects are roughly 1.5 to 2 times larger than the composition effects. These results suggest that during the 1990s, the primary source of changes in the wage distribution is the wage structure effect. The composition effect plays a very minor role.

The residual effect is defined as the unexplained part of the change in the wage distribution as shown in Table 3. By comparing the third and sixth columns of Table 4, we note that the residuals from the proposed method are smaller than those from the Machada and Mata (2005) method for all parts of the distribution except the region around the median. The results in Table 4 are consistent with the properties of the quantile regression estimators. The linear quantile regression provides the best estimators when the conditional quantile function is linear. Angrist et al. (2006) show that quantile regression gives the minimum mean-squared error linear approximation to the conditional quantile function even when the linear model is misspecified. Buchinsky (1995) gets similar results like ours by using Minimum Distance (MD) estimators for the Box–Cox quantile regression model.

The results of Tables 3 and 4 have been summarized in Fig. 2. The top left panel of Fig. 2 shows that the gaps between the estimated changes in wages for the linear and nonlinear quantile regression approaches increase in the higher quantiles of the wage distribution function. The residual effects have been shown in the bottom right panel. As shown, the absolute values of the residuals of the nonlinear quantile regression approach are significantly in the two tails of the wage distribution

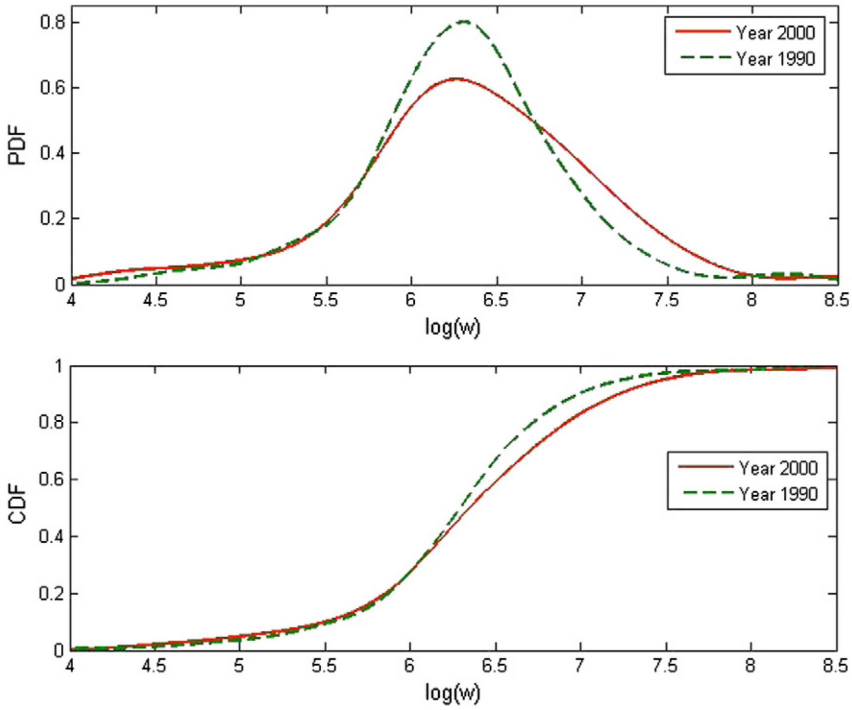

Fig. 1. Male weekly log wage density and distribution function for the years 1990 and 2000.

Table 3
Estimated changes in the U.S. male unconditional wage distribution over the period 1990–2000 by using the quantile regression and Box–Cox quantile regression*.

	Linear QR			Box–Cox QR		
Index	2000	1990	Changes	2000	1990	Changes
10th quant	5.663	5.669	−0.005	5.859	5.873	−0.014
	(5.660, 5.665)	(5.666, 5.671)		(5.857, 5.861)	(5.872, 5.875)	
25th quant	6.074	6.116	−0.042	6.164	6.183	−0.018
	(6.072, 6.076)	(6.114, 6.118)		(6.163, 6.165)	(6.181, 6.184)	
50th quant	6.467	6.495	−0.027	6.574	6.539	0.034
	(6.465, 6.468)	(6.493, 6.496)		(6.573, 6.575)	(6.538, 6.540)	
75th quant	6.843	6.830	0.013	6.940	6.844	0.095
	(6.842, 6.845)	(6.828, 6.831)		(6.938, 6.942)	(6.843, 6.845)	
90th quant	7.291	7.213	0.077	7.378	7.146	0.232
	(7.288, 7.294)	(7.210, 1.215)		(7.376, 7.381)	(7.144, 7.148)	
Skewness	−1.883	−1.099	−0.783	0.575	0.231	0.344
	(−1.904, −1.862)	(−1.35, 1.064)		(0.570, 0.581)	(0.226, 0.236)	
Scale	0.115	0.107	0.008	0.115	0.103	0.012
	(0.115, 0.116)	(0.106, 0.107)		(0.115, 0.116)	(0.102, 0.104)	
Gini coeff	0.081	0.066	0.015	0.052	0.045	0.006
	(0.081, 0.082)	(0.066, 0.067)		(0.052, 0.053)	(0.045, 0.046)	

* The numbers in the table represent the estimated weekly log wages, therefore their differences indicate the percentage changes in wages. The numbers in the parenthesis show the 95% bootstrap confidence interval which has been generated by 1000 replications. To construct the unconditional wage distribution functions, I have used 3000 quantiles between 0 and 1 from uniform distribution.

compared to the residuals from the linear quantile regression approach introduced by Machada and Mata (2005).

This can be explained by using the other three subplots of Fig. 2. The top right and bottom left panels show that the composition effects are similar in both methods but the absolute values of wage structure effects from the linear quantile regression approach are higher compared to the proposed Box–Cox quantile regression method at the two tails. Since the linear quantile regression method underestimates the changes in wages in the upper-tail and marginally overestimates in the lower tail of the distribution, the absolute values of the residuals are higher at the two tails of the distribution compared to the proposed wage decomposition method.

The next part of the decomposition measures the impact of individual workers' characteristics such as education, experience and race. This extended Oaxaca decomposition method takes care of the unobserved heterogeneity in returns to education (Henderson et al. 2011) by assuming the distribution of unobserved skills remains the same over the two time period we consider. Thus by using this type of decomposition method, we can only compare similar types of workers in two time periods. Alternatively one can use the Instrumental Variable Quantile Regression (IVQR) approach proposed by Chernozhukov and Hansen (2006, 2008, 2013b) to address the unobserved heterogeneity issue. Although the IVQR method has been applied in many research fields, it is rather new to the human capital variables literature.[3]

The second column of Table 5 shows that had the distribution of education remain the same as in 1990, then the 90/50 percentile wage gap would have increased by 4.2%. The inequality increasing effect of education is directly driven by the increase of the estimated effects for

[3] Wang (2013) and Arabsheibani and Staneva (2012) used IVQR method to find out heterogeneous returns to education.

Table 4
Decomposition of wage distribution function for the period 1990–2000 by using the quantile regression and Box–Cox Quantile regression*.

	Linear QR			Box–Cox QR		
Index	Composition	Wage structure	Residual	Composition	Wage structure	Residual
10th quant	−0.099 (−0.100, −0.098)	−0.224 (−0.224, −0.225)	−0.341	−0.088 (−0.088, −0.089)	−0.135 (−0.135, −0.134)	0.208
25th quant	−0.047 (−0.047, −0.048)	−0.100 (−0.101, −0.100)	0.141	−0.041 (−0.041, −0.040)	−0.067 (−0.067, −0.066)	0.090
50th quant	−0.000 (−0.001, −0.000)	−0.002 (−0.002, −0.001)	0.001	0.001 (0.001, 0.002)	0.003 (0.003, 0.004)	0.0294
75th quant	0.40 (0.040, 0.041)	0.089 (0.089, 0.090)	0.128	0.039 (0.039, 0.040)	0.070 (0.070, 0.071)	−0.016
90th quant	0.076 (0.076, 0.077)	0.181 (0.181, 0.182)	−0.247	0.079 (0.079, 0.080)	0.134 (0.134, 0.135)	0.016
Skewness	0.129 (0.125. 0.133)	−0.129 (−0.134, −0.125)	0.259	0.105 (0.101, 0.110)	0.033 (0.028, 0.038)	0.072
Scale	−1.886 (−1.893, −1.878)	−0.575 (−0.576,−0.573)	−1.311	0.486 (0.485, 0.487)	0.253 (0.253, 0.254)	0.232
Gini coeff	0.388 (0.387, 0.389)	0.247 (0.247, 0.248)	0.140	0.267 (0.266, 0.268)	0.123 (0.123, 0.124)	0.143

* Composition effect measures the impact due to the changes in labor force composition and the wage structure effect measures the market returns of the workers' characteristics. The unexplained part is called the residual effect.

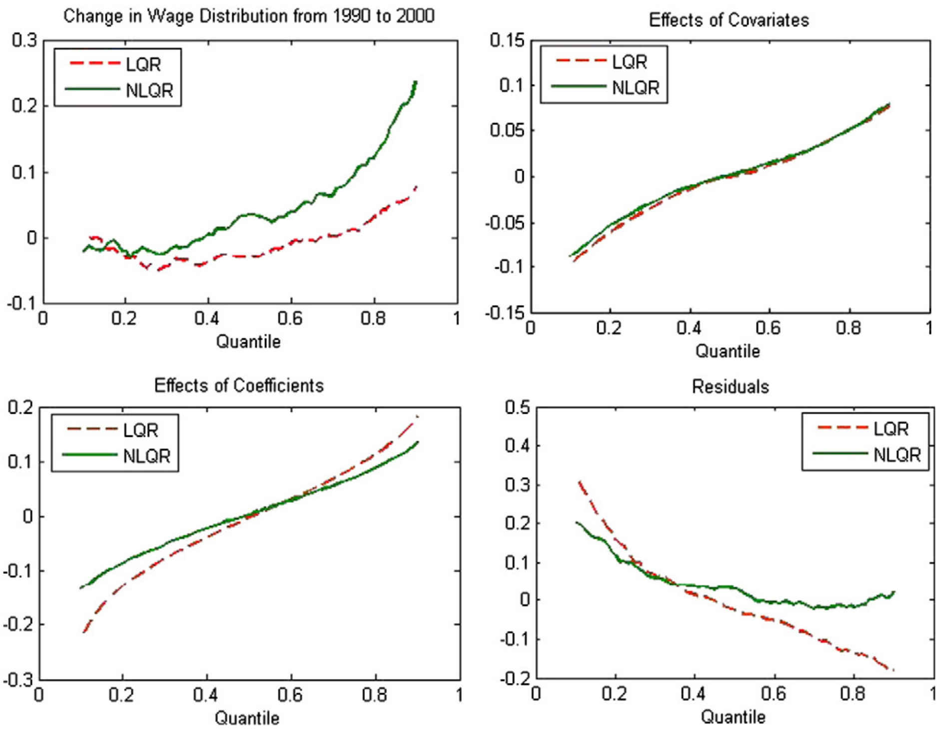

Fig. 2. Decomposition results by linear and non-linear quantile regression methods.

Table 5
Impacts of human capital variables to changes in the male wage distribution function over the period 1990–2000*.

Index	Education	Experience	Race	Residual
10th quant	−0.048 (−0.049, −0.047)	−0.073 (−0.073, −0.072)	−0.027 (−0.028, −0.027)	−0.134
25th quant	−0.030 (−0.030, −0.029)	−0.047 (−0.048, −0.047)	−0.022 (−0.022, −0.021)	−0.081
50th quant	−0.000 (−0.001, 0.000)	−0.007 (−0.008, −0.007)	−0.010 (−0.011, −0.010)	0.016
75th quant	0.019 (0.019, 0.020)	0.017 (0.17, 0.018)	−0.009 (−0.010, −0.009)	0.067
90th quant	0.042 (0.041, 0.043)	0.054 (0.054, 0.055	0.002 (0.002, 0.003)	0.133

* Impacts of human capital variables are estimated by: $F(w_{t=1}) - F(w_{t=1}; x'_{t=1}, x_{t=0})$ for $\widetilde{x}_t$ where $x_t$represents the set of all the other variables except $\widetilde{x}_t$.

education in the upper tail of the wage distribution. The impact of race does not have much contribution to the changes in wage distribution function. The fifth column shows that the residuals contribute approximately 8 to 13% of the total changes in the U.S. wage distribution during the 1990s. Melly (2005) and Lemieux (2006) both find that residual factors account for about 20% of the increase in wage inequality during the period 1973 to 1989.

3.3. Discussion

It is important to see how the results of our proposed decomposition method compare with the conventional wisdom of the existing wage inequality literature. Our results suggest that the dominating factor of the recent changes in the U.S. wage inequality is the wage structure effect. Machada and Mata (2005), Autor et al. (2005, 2008) and Firpo et al.

Table 6
(Extension of Autor et al. (2008) Table 2) — Regression and quantile regression models for the college high school log wage gap, 1963–2005.

	OLS	10th Q	50th Q	90th Q	OLS	10th Q	50th Q	90th Q
	(1)				(2)			
CLG/HS relative supply	−0.609	−0.931	−0.590	0.549	−0.728	−0.459	−0.689	−0.707
	(0.102)	(0.068)	(0.107)	(0.010)	(0.155)	(0.061)	(0.206)	(0.062)
Log real minimum wage					−0.049	−0.150	−0.011	−0.017
					(0.051)	(0.010)	(0.060)	(0.016)
Male prime-age unemp rate					0.04	−0.007	0.003	0.004
					(0.004)	(0.001)	(0.005)	(0.002)
Time	0.021	0.033	0.023	0.016	0.028	0.015	0.0282	0.022
	(0.006)	(0.002)	(0.006)	(0.006)	(0.007)	(0.002)	(0.010)	(0.003)
$Time^2/100$	0.030	0.0267	0.018	0.043	0.017	0.023	0.010	0.038
	(0.015)	(0.010)	(0.016)	(0.015)	(0.017)	(0.006)	(0.021)	(0.008)
$Time^3/1000$	−0.006	−0.007	−0.005	−0.008	−0.005	−0.004	−0.004	−0.007
	(0.002)	(−0.433)	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)
Observations	43	43	43	43	43	43	43	43
R-squared/Pseudo R-squared	0.952	0.702	0.813	0.838	0.955	0.734	0.814	0.850
	(3)				(4)			
CLG/HS relative supply	−0.403	−0.319	−0.399	−0.600				
	(0.067)	(0.015)	(0.105)	(0.045)				
Log real minimum wage	−0.117	−0.234	−0.065	−0.089	−0.144	−0.173	−0.004	0.0121
	(0.047)	(0.018)	(0.070)	(0.35)	(0.065)	(0.064)	(0.097)	(0.028)
Male prime-age unemp rate	−0.001	−0.006	−0.001	0.04	−0.018	−0.10	−0.018	−0.008
	(0.04)	(0.001)	(0.005)	(0.002)	(0.003)	(0.001)	(0.004)	(0.001)
Time	0.017	0.014	0.017	0.023	0.006	0.008	−0.006	0.006
	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.003)	(0.001)	(0.001)
Observations	43	43	43	43	43	43	43	43
R-squared/Pseudo R-squared	0.944	0.725	0.788	0.838	0.891	0.543	0.712	0.745

(2011) apply different decomposition methods to different data sets but obtain similar results. Another finding of this paper is that had the distribution of labor force remained the same as in 1990, the 90/50 wage gap would have increased roughly by 16% [4] and our estimated effects of returns to education can explain only one fourth of this change. Buchinsky (1994) shows that the differences of returns to education have increased from 0.73 to 2.32% over the period 1978 to 1987. Katz and Murphy (1992) find that college wage premium compared to the high school graduate has increased approximately 2.6% per annum during the period 1963–1987.

For a direct comparison with the results of Autor et al. (2008), I have used the same data to explain the trends in U.S. wage inequality over the period 1963–2005.[5] Table 6 represents regression models for the overall college high school wage gaps following the specifications proposed by Autor et al. (2008). In addition, I add quantile regression estimates for the 10th, 50th and 90th quantiles.[6] Models 3 and 4 consider linear time trends along with the college high school relative supply, real minimum wage and male prime-age unemployment rate whereas models 1 and 2 allow more flexible time trends by incorporating quadratic and cubic functions. By comparing the first and third columns of the four different specifications, we note that OLS estimates are very similar to the median regression.

The quantile regression estimates from the four specifications suggest that the range of the college wage premium estimates is 1.5 to 3.3% per annum from the lower-tail to the upper-tail of the wage distribution. The OLS estimates of the relative supply measures are significantly different from the quantile regression estimates at the two tails of the wage distribution. The Quantile Regression estimated elasticity of substitution between high school and college graduate workers varies over the range 1.07(1/0.93) to −1.82(1/−0.54) from the lower-tail to upper-tail of the wage distribution. This estimated elasticity of substitution between high school and college graduate workers cannot explain the changes in the upper-tail wage inequality and this is consistent with Autor et al. (2008) since they also argue that the two-factor CES model fails to provide an adequate explanation of the evolution between wage inequality since the early 1990s.

Table 7 shows the regression and quantile regression models for the college high school wage gaps by overall and four different experience groups. The inverse of the aggregate supply coefficient represents the aggregate elasticity of substitution; similarly, the inverse of the coefficients for the first row shows the partial elasticity of substitution between different experience groups within the same education group. Consistent with Autor et al. (2008), the quantile regression estimates also suggest that there are substantial effects of both own group and aggregate supplies on the evaluation of the college wage premium by experience groups.

Autor et al. (2008) find that aggregate elasticity of substitution in Table 7 models are very similar to the aggregate models in Table 6 which is also true in the upper-tail of the wage distribution, shown in the 4th and 8th column. However, in the lower tail, the values of the aggregate elasticity of substitution are significantly different in Table 6 and Table 7. This implies that the return to college for the low-experienced workers is different compared to the medium-experienced (20–29 years) and older workers (30–39 years). The partial elasticity of substitution between experience groups is very consistent across the wage distribution and similar to the OLS estimate of 3.55 as shown by Autor et al. (2008). The lower panel shows that the OLS and median regression estimates are very similar for the four experience groups.

Table 8 summarizes the overall and residual inequality over the period 1963–2005 by using the proposed counterfactual wage decomposition method. For the counterfactual analysis, Autor et al. (2008) use a model which consists of 95 dummy variables: full set of age dummies, dummies for nine discrete schooling categories, and a full set of interactions among the schooling dummies and a quartic in age and no continuous variable. I do not use the same covariates as used by Autor et al. (2008) to avoid the computational complexity of the interior point algorithms for quantile regression as discussed by Koenker (2005)[7]; rather I follow the same specification used in the main empirical analysis for

[4] The difference between the fifth and third rows of wage structure effect at Table 4.

[5] See Table 1 in Autor et al. (2008) for descriptive statistics and appendix for data preparation details.

[6] I do not include nonlinear quantile regression estimates because the nonlinear quantile regression estimates from the statistical software 'R' are infeasible for 43 observations.

[7] See Roger Koenker (2005) 'Quantile Regression' for a detailed discussion. Standard statistical softwares fail to provide quantile regression estimates for the models which have too many dummy variables.

Table 7
(Extension of Autor et al. (2008) Table 3) — Regression and quantile regression models for the college high school log wage gap by potential experience group 1963–2005, males and females pooled.

All experience groups								
	OLS	10th Q	50th Q	90th Q	OLS	10th Q	50th Q	90th Q
Own supply minus agg supply	−0.282	−0.238	−0.296	−0.284	−0.281	−0.212	−0.321	−0.341
	(0.027)	(0.054)	(0.046)	(0.036)	(0.026)	(0.038)	(0.039)	(0.050)
Aggregate supply	−0.600	−0.738	−0.561	−0.507	−0.705	−1.012	−0.732	−0.521
	(0.087)	(0.149)	(0.150)	(0.144)	(0.130)	(0.193)	(0.190)	0.209
Log real minimum wage					−0.074	−0.126	−0.071	−0.119
					(0.037)	(0.067)	(0.055)	(0.058)
Male prime-age unemp rate					0.004	0.009	0.004	−0.001
					(0.003)	(0.004)	(0.005)	(0.006)
Time	0.027	0.034	0.024	0.023	0.031	0.004	0.030	0.023
	(0.004)	(0.008)	(0.007)	(0.006)	(0.006)	(0.009)	(0.009)	(0.009)
$Time^2$/1000	−0.009	−0.017	−0.005	−0.005	−0.012	−0.026	−0.010	−0.007
	(0.005)	−0.433)	(0.008)	(0.007)	(0.006)	(0.010)	(0.008)	(0.008)
Observations	172	172	172	172	172	172	172	172
R-squared/Pseudo R-squared	0.863	0.576	0.641	0.682	0.868	0.594	0.647	0.687
Potential experience groups	0–9 years		10–19 years		20–29 years		30–39 years	
	OLS	Median	OLS	Median	OLS	Median	OLS	Median
Own supply minus agg supply	−0.169	−0.173	−0.325	−0.344	0.101	0.052	0.002	0.038
	(0.130)	(0.101)	(0.084)	(0.201)	(0.084)	(0.176)	(0.119)	(0.168)
Aggregate supply	−0.854	−0.955	−0.474	−0.349	−0.398	−0.382	−0.544	−0.596
	(0.262)	(0.207)	(0.182)	(0.429)	(0.224)	(0.448)	(0.190)	0.292
Log real minimum wage	−0.340	−0.655	−0.145	−0.130	0.098	0.125	0.028	0.043
	(0.076)	(0.062)	(0.49)	(0.119)	(.54)	(0.111)	(0.067)	(0.087)
Male prime-age unemp rate	0.005	0.011	0.002	−0.001	0.003	0.004	000	0.06
	(0.007)	(0.005)	(0.004)	(0.010)	(0.005)	(0.009)	(0.006)	(0.008)
Time	0.040	0.040	0.015	0.010	0.016	0.013	0.028	0.026
	(0.012)	(0.009)	(0.009)	(0.020)	(0.011)	(0.021)	(0.011)	(0.013)
$Time^2$/1000	−0.025	−0.021	0.010	0.012	0.000	−0.026	−0.021	−0.016
	(0.012)	(−0.008)	(0.009)	(0.020)	(0.010)	(0.020)	(0.012)	(0.015)
Observations	43	43	43	43	43	43	43	43
R-squared/Pseudo R-squared	0.926	0.787	0.969	0.839	0.868	0.666	0.647	0.687

consistency. To construct the counterfactual wage distribution function by using the simulation method as described in Section 2, I use 5000 replications. The numbers in Table 4 represent the effects of labor market prices on the overall earnings inequality while holding labor force composition at the base years 1973, 1989 and 2005.

For comparison with Autor et al. (2008), we begin by discussing the overall wage inequality. The upper panel of Table 4 shows that the male 90/50 percentile wage gap (upper-tail wage inequality) rose during both halves of the sample by 10.3% from 1973 to 1989 and 15.4% from 1989 to 2005. We note that the price effects can explain almost all changes in the upper-tail wage inequality. For females, the 90/50 percentile wage gap rose by 11% between 1973 and 1989 and 12.9% between 1989 and 2005. Again holding labor force composition constant at its 1973, 1989 and 2005 levels gives us results similar to the men. The middle and bottom panels of Table 8 show the price effects for 50/10 and 90/10 percentile wage gaps for both men and women. These results suggest that the basic message, price effect dominates the composition effect, does not change for both men and women over the period 1973–2005.

Autor et al. (2008) conclude that "Changes in labor force composition do not substantially contribute to an explanation for the diverging path of upper and lower-tail inequality, either overall or residual, over the past three decades". As shown in Table 8, our findings are qualitatively similar with the conclusions of Autor et al. (2008) although we use a simpler model specification and a different counterfactual decomposition method to obtain the price effects and composition effects. For residual inequality, I find that for the period 1989–2005, in most of the cases price effects are the dominating factors for changing in residual wage inequality which is consistent with Autor et al. (2008). However for the period 1973–1989, I find that although price effect plays a significant role to changes in residual wage inequality, it is not the dominating factor as suggested by Autor et al. (2008).

To conclude, Autor et al. (2008) use a structural model and college wage premium compared to high school graduate as a measure of returns to education whereas this paper uses a canonical Mincer earnings equation and constructs a counterfactual wage distribution function to estimate the returns from human capital variables. However, the results from the main empirical application and comparison with Autor et al. (2008) demonstrate that although the relative magnitudes of the price effects from the proposed method are higher in some cases, especially for the wage gap between the two tails of the distribution, these results are qualitatively similar. Our results find even stronger price effects for the overall wage inequality compared to Autor et al. (2008) especially for the 90/10 percentile wage gap; however, the main conclusion remains the same. Autor et al. (2008) point out that their counterfactual analysis depends on the partial equilibrium assumption that prices and quantities are independent which is precisely to the opposite spirit of the supply–demand analysis; the proposed method does not require such a strong economically unappealing assumption.

4. Conclusion

This paper proposes a modified wage decomposition method based on a very popular approach proposed by Machada and Mata (2005). First, I show that the linearity assumption of the conditional quantile function does not hold for the U.S. male wage distribution data at the year 2000 by using the Generalized Conditional Cramer-von Mises (GCCM) test proposed by Rothe and Wied (2013). To overcome this linearity issue, I use the Chamberlain Buchinsky Two Step Box–Cox quantile regression method. The advantage of this method over the linear quantile regression method is that the shape parameter of the Box–Cox transformation adjusts with each quantile of the wage

Table 8
100 × Observed and composition constant changes in overall and residual wage inequality measures by using counterfactual decomposition method.

		Residual Inequality			Overall Inequality		
Δ 90/50							
		1973–1989	1989–2005	1973–2005	1973–1989	1989–2005	1973–2005
Males	Observed	5.40	4.23	9.63	10.31	15.40	25.71
	1973 X's	0.64	2.16	2.81	9.51	8.96	18.47
	1989 X's	1.00	2.50	3.49	10.92	10.78	21.71
	2005 X's	1.69	2.22	3.90	9.30	9.19	18.49
Females	Observed	7.60	4.11	11.71	11.06	12.97	24.03
	1973 X's	3.60	0.88	4.48	10.82	7.92	18.74
	1989 X's	3.55	2.61	6.16	11.28	10.07	21.35
	2005 X's	4.38	4.04	8.42	10.90	10.62	21.53
Δ 50/10							
		1973–1989	1989–2005	1973–2005	1973–1989	1989–2005	1973–2005
Males	Observed	7.39	−1.67	5.71	7.59	−2.43	5.61
	1973 X's	2.43	−0.37	2.06	10.77	3.93	14.70
	1989 X's	2.72	−0.05	2.67	10.84	3.27	14.11
	2005 X's	3.17	0.50	3.66	11.83	4.30	16.13
Females	Observed	8.84	0.02	8.86	13.28	0.39	13.66
	1973 X's	2.15	−1.80	0.35	10.55	5.10	15.65
	1989 X's	2.86	−1.44	1.43	9.81	5.60	15.41
	2005 X's	2.85	−0.22	2.64	9.77	7.11	16.88
Δ 90/10							
		1973–1989	1989–2005	1973–2005	1973–1989	1989–2005	1973–2005
Males	Observed	12.79	2.56	15.34	17.90	12.97	30.87
	1973 X's	30.8	1.79	4.87	20.28	12.89	33.17
	1989 X's	3.72	2.45	6.16	21.76	14.05	35.18
	2005 X's	4.85	2.71	7.57	21.13	13.49	34.62
Females	Observed	16.43	4.14	20.57	24.23	13.35	37.69
	1973 X's	5.75	−0.91	4.84	21.37	13.02	34.38
	1989 X's	6.41	1.18	7.59	21.08	15.57	36.75
	2005 X's	7.23	3.82	11.05	20.67	17.74	38.41

distribution to capture the nonlinear shape of the conditional quantile function.

This paper also shows that the estimated counterfactual distribution function by using the Box–Cox quantile regression method converges to the true distribution function asymptotically. This asymptotic property ensures the validity of the inferences based on the decomposition method results. By using the re-sampling procedure, I also show that theoretically we can fully recover the unconditional wage distribution function. This implies that we need a large number of draws to construct a counterfactual wage distribution function. I have used 3000 draws for the empirical analysis and 5000 draws for the comparison results with Autor et al. (2008).

This paper also implements the proposed decomposition method to determine the impact of human capital variables that contribute to changes in the male wage distribution function over the period 1990–2000. The proposed decomposition method results show that changes in the labor force composition do not contribute an explanation for the steep rise in wages at the upper-tail of the U.S. wage distribution during the 1990s. This result is consistent with Autor et al. (2008) who used a variant of DiNardo et al. (1996) wage decomposition method for the period 1963–2005. In addition, this paper also shows that returns from education can explain only a small fraction of the total changes in wages in the upper tail of the wage distribution and do not explain why the lower tail wage inequality fell during the 1990s. Thus, return from schooling is not the primary source of recent divergence of upper-tail and lower-tail wage inequality in the U.S. labor market.

The sources of the asymmetric changes in the U.S. wage distribution with a steady rise in the upper-tail and a marginal fall in the lower-tail during the 1990s still remain unresolved. Autor et al. (2008) show that hypotheses such as falling minimum wages, relative demand growth for more skilled workers and skill biased technological change fail to fully explain the upper-tail inequality paired with declining lower-tail inequality. Firpo et al. (2011) use unconditional quantile regression method to explain the changes in the U.S. wage structure over the last three decades by looking at the roles of skilled biased technological changes and job offshorability. Katz and Murphy (1992) and Autor et al. (2008) propose an explanation based on relative demand growth for high-skill workers and fluctuations in relative skill supplies based on canonical supply demand models. However, this argument is unable to explain the slowdown in relative demand growth during the early 1990s.

There is no unique explanation for the complex pattern of changes in the U.S. wage distribution function during the 1990s. Having argued that labor market institutional factors, skill biased technological change, relative demand for skilled labor, effects of international trade and their dynamic interactions play very important roles in explaining changes to the wage distribution function; this paper provides an explanation of these asymmetric changes in the U.S. wage distribution function by considering the role of human capital variables in the canonical Mincer earnings framework and finds that changes in the composition of the labor force play a secondary role, whereas the primary source of the asymmetric changes in the U.S. wage distribution during the 1990s is the changing labor market prices. This paper also shows that the steep increase of the upper-tail wage inequality in the U.S. labor market during the 1990s is not caused primarily by the heterogeneous returns from education.

Appendix A

Proof of property 1: Complete mappings

By assumption, $F_{w_t|x_t}(.)$ is monotonically increasing over the domain of x_t so the inverse $F^{-1}_{w_t|x_t}(\tau)$ exists and is well defined. Since $\tau\in(0,1)$, by applying the probability integral transformation theorem we set

$$\tau = F_{w_t|x_t}(w_t)$$
$$\Rightarrow F^{-1}_{w_t|x_t}(\tau) = w_t.$$

So for each value of $\tau\in(0,1)$ we get a unique draw from $F_{w_t|x_t}(.)$. We know that the number of values τ can take in (0,1) is countable infinite. Again, by assumption, $\mathcal{W}_t$ is closed and bounded in $\mathbb{R}^k$ so for each value of $w_t\in\mathcal{W}_t$, we can find a unique $\tau\in(0,1)$. This implies that as the number of draws (m) goes to infinity we can recover every possible values from $F_{w_t|x_t}(w_t)$.

Proof of property 2: Convergence in distribution

Under some regularity conditions, Powell (1991) shows in Theorem 2, the joint distribution of $\theta^t = \begin{pmatrix}\lambda_\tau^t\\ \beta_\tau^t\end{pmatrix}$ for $t = 0$ and 1 is,

$$\sqrt{n}\left(\hat{\theta}^t(\tau)-\theta_0(\tau)\right)d\to\mathcal{N}\left(0, H_n^t(\theta)^{-1}V_n^t(\theta)H_n^t(\theta)^{-1}\right). \quad (11)$$

The Hessian matrix H_n is

$$H_n^t\left(\theta^t\right) = \frac{1}{n}\sum_{i=1}^{n} E\left[f_w\left(g\left(x_i^t,\theta^t\right)\middle|x_i^t\right)\dot{g}_i^t\left(\theta_0^t\right)\left(\dot{g}_i^t\left(\theta_0^t\right)\right)'\right] \quad (12)$$

where $f_w(g(\cdot)|x_i^t)$ is the conditional density of w_i given x_i,$\dot{g}_i = \delta g\left(x_i^t,\theta^t\right)/\delta\theta^t$ and

$$V_n^t\left(\theta^t\right) = \frac{1}{n}\sum_{i=1}^{n}\psi_i\left(\theta^t\right)\left[\psi_i\left(\theta^t\right)\right]' \quad (13)$$

where $\psi_i\left(\theta^t\right) = \left[\tau - \mathbb{I}\left(w_i < g\left(x_i^t, \theta^c\right)\right)\right] t_i\left(\theta^t\right)$ and

$$t_i\left(\theta^t\right) = min\left[\frac{\delta^- s_i\left[h\left(x_i^t, \theta^t\right)\right]}{\delta h}, \frac{\delta^+ s_i\left[h\left(x_i^t, \theta^t\right)\right]}{\delta h}\right]. \tag{14}$$

The counterfactual conditional quantile function obtained from the simulation method described in the Section 2 has the form

$$\begin{aligned} w_i^c &= q\left(f(x_i, \theta^c) + u_i^c\right) \\ &= g^c(x_i, \theta^c, u_i^c) \end{aligned}$$

where u_i^c is the error term. Note that the superscript c denotes the counterfactual. The above function can be rewritten in the form of a nonlinear regression model with additive error,

$$w_i^c = g(x_i, \theta^c) + v_i^c \quad \text{where} \quad v_i^c = g(x_i, \theta^c, u_i^c) - g(x_i, \theta^c)$$

Assumptions:

A1. The transformation $g^c(x_i, \theta^c, u_i)$ is of the form

$$g^c(x_i, \theta^c, u_i^c) = s^c\left[h^c(x_i, \theta^c, u_i^c), x_i\right] \equiv s_i^c\left[h^c(x_i, \theta^c, u_i^c)\right] \tag{15}$$

where $h(x_i^t, \theta, u_i)$ is scalar valued and continuously differentiable with respect to θ and $s_i(.)$ is (right and left) differentiable and satisfies the other regularity conditions described by Powell (1991).

A2. The parameter vector $\theta_0^c = ((\lambda_0^c)', (\beta_0^c)')'$ is an interior point of the parameter space Θ.

By assumption $\hat{\theta}^c$ is the quantile estimator that minimizes $g^c(.)$ and, also by using the above two assumptions, we obtain the following F.O.C.

$$\Psi_n^c\left(\hat{\theta}^c\right) = \frac{1}{n}\sum_{i=1}^{n}\left[\tau - \mathbb{I}\left(w_i < g\left(x_i^t, \theta^c\right)\right)\right] t_i(\theta^c) = 0 \tag{16}$$

where

$$t_i(\theta^c) = min\left\{\frac{\delta^- s_i\left[h\left(x_i^t, \theta^t\right)\right]}{\delta h}, \frac{\delta^+ s_i\left[h\left(x_i^t, \theta^t\right)\right]}{\delta h}\right\} \frac{\delta h(x_i, \theta)}{\delta \theta}.$$

Denoting $\psi_i(\theta^c) = \left[\tau - \mathbb{I}(w_i < g(x_i^t, \theta^c))\right] t_i(\theta^c)$ we have,

$$\Psi_n^c\left(\hat{\theta}^c\right) = \frac{1}{n}\sum_{i=1}^{n} \psi_i^c\left(\hat{\theta}^c\right) = o_p\left(\frac{1}{\sqrt{n}}\right). \tag{17}$$

Now the relation in Eq. (17) has been established, the conditions of Theorem 2 of Powell (1991) have been satisfied. By applying Powell's result shown above, we obtain the asymptotic normality of the estimator $\hat{\theta}^c$

$$\sqrt{n}\left(\hat{\theta}^c(\tau) - \theta_0^c(\tau)\right) \xrightarrow{d} \mathcal{N}\left(0, H_n^c(\theta)^{-1} V_n^c(\theta) H_n^c(\theta)^{-1}\right) \tag{18}$$

where $H_n^c(\theta^c) = \frac{1}{n}\sum_{i=1}^{n} E\left[f_w\left(g(x_i^c, \theta^c)|x_i^c\right)\dot{g}_i^c(\theta_0^c)\left(\dot{g}_i^c(\theta_0^c)\right)'\right]$, $V_n^c(\theta^c) = \frac{1}{n}\sum_{i=1}^{n} \psi_i(\theta^c)\left[\psi_i(\theta^c)\right]'$, and $f_w(g^c(.)|x_i^t)$ is the conditional density of w_i^c given x_i, $\dot{g}_i^c = \delta g^c\left(x_i^t, \theta^c\right)/\delta\theta^c$.

The next step is to show that Y_n is a continuous function of $\theta^c \in \Theta$ where Θ is the parameter space of θ^c. We can write Y_n as a composite function of h_1 and h_2 in the following way:

$$\begin{aligned} Y &= g^c\left(x'_{t=0}\beta_\tau^{t=1}, \lambda_\tau^{t=0}\right) \\ &= \left(\lambda_\tau^{t=0} x'_{t=0}\beta_\tau^{t=1} + 1\right)^{1/\lambda_\tau^{t=0}} \\ &= h_1(h_2(\cdot)) \end{aligned}$$

where $h_1 = 1/\lambda_\tau^{t=0}$ and $h_2 = \lambda_\tau^{t=0} x'_{t=0}\beta_\tau^{t=1} + 1$. We note that h_1 and h_2 both are continuous function in the parameter space Θ. By applying the theorem (Walter Rudin, Theorem 4.10, page 87) which states a continuous function of a continuous function yields a continuous function, we conclude that Y_n is continuous in θ^c. Now by applying the Continuous mapping theorem we obtain the following result,

$$F_n\left(\hat{Y}_n\right) \xrightarrow{d} F(Y) as \quad n \to \infty$$

References

Acemoglu, D., Autor, D., 2010. Skills, tasks and technologies: implications for employment and earnings. NBER Working Paper No 16082.

Altonji, J., Bharadwaj, P., Lange, F., 2008. Changes in the characteristics of American youth: implications for adult outcomes. NBER Working Paper.

Angrist, J., Pischke, S., 2009. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press, Princeton.

Angrist, J., Chernozhukov, V., Fernandez-Val, I., 2006. Quantile regression under misspecification, with an application to the U.S. wage structure. Econometrica 20, 445–465.

Antonczyk, D., Fitzenberger, B., Sommerfeld, K., 2010. Rising wage inequality, the decline of collective bargaining, and the gender gap. IZA Working Paper No 4911.

Arabsheibani, R., Staneva, A., 2012. Returns to education in Russia: where there is risky sexual behavior there is also an instrument. IZA Working Paper No 6726.

Autor, D., Katz, L., Krueger, A., 1998. Computing inequality: have computers changed the labor market? Q. J. Econ. 113, 1169–1213.

Autor, D., Levy, F., Murname, R., 2003. The skill content of technological change: an empirical investigation. Q. J. Econ. 118, 1279–1333.

Autor, D., Katz, L., Kearney, M., 2005. Rising wage inequality: the role of composition and prices. NBER Working Paper 11628.

Autor, D., Katz, L., Kearney, M., 2008. Trends in the U.S. wage inequality: revising the revisionists. Rev. Econ. Stat. 97, 300–323.

Barsky, R., Bound, J., Charles, K., Lupton, J., 2002. Accounting for the black–white wealth gap: a nonparametric approach. J. Am. Stat. Assoc. 97, 663–673.

Box, G., Cox, D., 1964. An analysis of transformation. J. R. Stat. Soc. 26, 211–252.

Buchinsky, M., 1994. Estimating the asymptotic covariance matrix for quantile regression models A Monte Carlo study. J. Econ. 68, 303–338.

Buchinsky, M., 1995. Quantile regression, Box–Cox transformation model, and the U.S. wage structure, 1963–1987. J. Econ. 65, 109–154.

Card, D., Dinardo, J., 2002. Skill-biased technological change and rising wage inequality: some problems and puzzles. J. Labor Econ. 20, 733–783.

Chamberlain, G., 1994. Quantile regression, censoring and the structure of wage. In: Sims, C., Laffont, J.J. (Eds.), Proceedings of the Sixth World Congress of the Econometric Society. Barcelona, Spain. Cambridge University Press, New York, NY.

Chernozhukov, V., Hansen, C., 2006. Instrumental quantile regression inference for structural and treatment effect models. J. Econ. 132, 491–525.

Chernozhukov, V., Hansen, C., 2008. Instrumental quantile regression: a robust inference approach. J. Econ. 142, 379–398.

Chernozhukov, V., Hansen, C., 2013. Quantile models with endogeneity. Ann. Rev. Econ. 5, 57–81.

Chernozhukov, V., Fernandez-Val, I., Melly, B., 2013. Inference on counterfactual distributions. Econometrica 81 (6), 2205–2268.

DiNardo, J., Fortin, N., Lemieux, T., 1996. Labor market institution and the distribution of wages, 1973–1992: a semiparametric approach. Econometrica 64, 1001–1044.

Donald, S., Green, D., Paarsch, H., 2000. Differences in wage distributions between Canada and the United States: an application of a flexible estimator of distribution functions in the presence of covariates. Rev. Econ. Stud. 67, 609–633.

Firpo, S., 2007. Efficient semi-parametric estimation of quantile treatment effects. Econometrica 75, 259–276.

Firpo S., 2010. Identification and Estimation of Distributional Impacts of Interventions Using Changes. EESP-FGV. mimeo.

Firpo, S., Fortin, N., Lemieux, T., 2007. Decomposing wage distributions using re-centered influence function regressions. NBER working paper.

Firpo, S., Fortin, N., Lemieux, T., 2009. Unconditional quantile regressions. Econometrica 77, 953–973.

Firpo, S., Fortin, N., Lemieux, T., 2011. Occupational tasks and changes in the wage structure. IZA working paper no 5542.

Fitzenberger, B., Wilke, A., Zhang, X., 2010. Implementing Box–Cox quantile regression. Econ. Rev. 29, 158–181.

Fortin, N., Lemieux, T., Firpo, S., 2011. Decomposition methods in economics. Handb. Labor Econ. 4, 1–102.

Frolich, M., 2004. Finite-sample properties of propensity-score matching and weighting estimators. Rev. Econ. Stat. 86, 77–90.
Goos M., and Manning A., 2003. Lousy and Lovely Jobs: The Rising Polarization of Work in Britain. Mimeo, London School of Economics.
Gosling, A., Machin, S., Meghir, C., 2000. The changing distribution of male wages in the U.K. Rev. Econ. Stud. 65, 635–666.
Henderson, J., Polachek, W., Wang, L., 2011. Heterogeneity in schooling rates of return. Econ. Educ. Rev. 30, 1202–1214.
Hirano, K., Imbens, G., Ridder, G., 2002. Identification and estimation of triangular simultaneous equations models without additivity. Econometrica 77 (5), 1481–1512.
Imbens, G., Newy, W., 2009. Identification and estimation of triangular simultaneous equations models without additivity. Econometrica 77 (5), 1481–1512.
Juhn, C., Murphy, M., Pierce, B., 1993. Wage inequality and the rise in return to skill. J. Polit. Econ. 101, 410–442.
Katz, L., Murphy, K., 1992. Changes in relative wages, 1963–1987: supply and demand factors. Q. J. Econ. 107, 35–78.
Koenker, R., 2005. Quantile Regression. Cambridge University Press.
Lemieux, T., 2006. Increasing residual wage inequality: composition effects, noisy data, or rising demand for skill? Am. Econ. Rev. 96, 461–498.
Machada, A.F., Mata, J., 2005. Counterfactual decomposition of changes in wage distributions using quantile regression. J. Appl. Econ. 20, 445–465.
Melly, B., 2005. Decomposition of differences in distribution using quantile regression. J. Labor Econ. 12, 577–590.
Oaxaca, R., 1973. Male–female wage differentials in urban labor markets. Int. Econ. Rev. 14, 693–709.
Powell, J., 1991. Estimation of monotonic regression models under quantile restrictions. In: Barnett, W.A., Powell, J., Tauchen, G. (Eds.), Nonparametric and semiparametric methods in econometrics and statistics, Proceeding to the Fifth International Symposium in Economic Theory and Econometrics. Cambridge University Press, New York, NY, pp. 357–384.
Rothe, C., 2010a. Nonparametric estimation of distributional policy effects. J. Econom. 155, 56–70.
Rothe, C., 2010b. Identification of unconditional partial effects in nonseparable models. Econ. Lett. 109, 171–174.
Rothe, C., 2012a. Partial distributional policy effects. Econometrica 80, 2269–2301.
Rothe, C., 2012b. Decomposing the composition effect. IZA working paper no 6397.
Rothe, C., Wied, D., 2013. Misspecification testing in a class of conditional distributional models. J. Am. Stat. Assoc. 108 (501), 314–324.
Wang, L., 2013. How does education affect the earnings distribution in urban China 1995–2002? Oxf. Bull. Econ. Stat. 75, 435–454.