Box–Cox power transformation unconditional quantile regressions with an application on wage inequality

Pallab Kumar Ghosh

JOURNAL OF APPLIED STATISTICS
https://doi.org/10.1080/02664763.2020.1795817

Taylor & Francis
Taylor & Francis Group

Box–Cox power transformation unconditional quantile regressions with an application on wage inequality

Pallab Kumar Ghosh

Department of Economics, University of Oklahoma, Norman, OK, USA

ABSTRACT

This study proposes a semi-parametric estimation method, Box–Cox power transformation unconditional quantile regression, to estimate the impact of changes in the distribution of the explanatory variables on the unconditional quantile of the outcome variable. The proposed method consists of running a nonlinear regression of the recentered influence function (RIF) of the outcome variable on the explanatory variables. We also show the asymptotic properties of the proposed estimator and apply the estimation method to address an existing puzzle in labor economics–why the 50th/10th percentile wage gap has been falling in the USA since the late 1980s. Our results show that declining unionization can explain approximately 10% of the decline in the 50/10 wage gap in 1990–2000 and 23% in 2000–2010.

ARTICLE HISTORY
Received 27 January 2019
Accepted 4 July 2020

KEYWORDS
Unconditional quantile regression; extended Box–Cox power transformation unconditional quantile regression; Yeo–Johnson RIF regression; Unionization; 50/10 and 90/50 percentile wage gaps

JEL CODE
C14; C21; J31

1. Introduction

Koenker and Bassett [41] introduce conditional quantile regression to estimate the impacts of explanatory variables (X) on the conditional quantiles of the outcome variable (Y). However, researchers applying this method are often interested in finding the effect of X on the unconditional distribution of Y. For example, the τth quantile regression estimates the union wage premium only for union workers. In contrast, [21] shows that in a general equilibrium frame work, a change in the unionization rate also affects earnings for nonunion workers. To capture the total effect of declining unionization for all workers at the τth quantile of the earnings distribution, we need to estimate the impact of unionization on the marginal distribution of income.

The impact of a change in the distribution of X on the marginal distribution of Y ($F_Y(y)$) or other function of $F_Y(y)$ can be estimated using a regression-based estimation method known as unconditional quantile regression developed by Firpo et al. [26]. This estimation method is also known as RIF (re-centered influence function) regression because it consists of running a regression of the RIF of Y on X.[1] To enable a direct comparison between

CONTACT Pallab Kumar Ghosh pallab.ghosh@ou.edu

Supplemental data for this article can be accessed here. https://doi.org/10.1080/02664763.2020.1795817

unconditional and conditional quantile regression estimates, [26] show that the partial effects of unconditional quantile regressions are simply weighted averages of the marginal effects of conditional quantile regressions. The weights are the ratio of the conditional to marginal density of Y at different quantiles.

One of the main reasons why the unconditional quantile regression approach has attracted a great deal of attention in the econometrics literature is that conditional quantiles do not average up to their unconditional population counterparts. In particular, $Q_\tau(y|X=x) = x'\beta_\tau$ does not imply that $Q_\tau(y) = Q_\tau(x')\beta_\tau$.[2] As a result, the estimates obtained by running a quantile regression cannot be used to estimate the impacts of X on the corresponding unconditional quantile of Y. For example, using quantile regression estimates, empirical researchers cannot determine the impact of a marginal increase in all workers' education on some features of the wage distribution, such as its moments, quantiles, Gini coefficient or other measures of wage inequality.

In a linear model, the unconditional quantile partial effect can be easily obtained by using an ordinary least squares regression.[3] However, the conceptual difficulties of defining unconditional quantile partial effects are caused by the possible nonlinear relationship between the distribution of Y and X. In the linear unconditional quantile regression model, we make the assumption that the change in the distribution of X simply shifts the location of the distribution of Y and leaves all other features of the distribution unchanged. In many empirical applications, this assumption may not be true. For example, [7,8] show that since the 1980s, the US income distribution conditional on labor market characteristics does not belong to a location family. The asymmetric changes in the US income distribution are known as the polarization of the US economy.

This study considers an alternative approach of Firpo *et al.* [26] by incorporating an extended Box–Cox power transformation model within the framework of RIF regression.[4] Gaudry and Laferriere [31] show that Box–Cox transformation is a power invariance transformation of the linear model. Thus the proposed Box–Cox power transformation RIF regression model is quite intuitive since a monotonic transformation preserves the order of the order statistics. Machado and Mata [44] argue that in a simple setup, Box–Cox transformation provides the needed flexibility, as both the transformation parameter and the coefficients of the linear function are allowed to vary freely at each point of the distribution. The additional feature of this transformation is that both linear and log-linear cases are special forms. In terms of estimation, both transformation parameter and coefficients can be estimated in a nonlinear regression framework.

Machado and Mata [45] 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. Machado and Mata [45] approach works under the assumption that conditional quantile function is linear in X. To address this concern, [32] uses the Box–Cox quantile regression model to construct a counterfactual wage distribution function and finds that both [45] method and Ghosh [32]'s simulation method based on Box–Cox quantile regression model provide similar results except the two-tail of the wage distribution function. Chernozhukov *et al.* [23] introduce the distributional regression model that corresponds to changes in the distribution of the covariates, changes in the conditional distribution of the outcome given covariates, or both. Both of these approaches yield similar results when the support of X is sufficiently rich.[5]

Using a linear unconditional quantile regression model, [25] construct a counterfactual wage distribution function to decompose the US wage distribution function. Barsky *et al.* [14] note that a linear regression-based decomposition method yields inconsistent estimates when the true conditional expectation function is not linear. To address this problem, [27] use the re-weighting procedure developed by DiNardo et al. [24]. However, the re-weighting procedure becomes practically infeasible when there are too many continuous variables. Moreover, re-weighting can have some undesirable properties in small samples when there is a problem of common support. Frolich [30] finds that re-weighting estimators perform poorly in this context.

To construct a counterfactual wage distribution, an alternative approach of Firpo *et al.* [27] would be an extended Box–Cox power transformation unconditional quantile regression model. In this context, [47] shows the use of the Box–Cox Transformation in limited dependent variable models.[6] In addition, [37] apply the Box–Cox power transformation to the odd ratio so that we can obtain the logistic model. In particular, the motivation of introducing an alternative RIF-regression approach is that unlike the linear RIF regression model, the proposed Box–Cox power transformation unconditional quantile regression model can be used in many empirical applications where a small change in an independent variable not only changes the location of the distribution of the outcome variable but also changes the other features of the distribution.

One such example we consider in the empirical application section in which we address an open question in labor economics as to why the 50/10 wage gap has been falling in the USA since the early 1980s. Autor *et al.* [8] show that returns from education, falling minimum wage, SBTC and other factors cannot explain why middle-income workers are becoming worse off than low-skilled workers in the left tail of the income distribution. This study provides a comprehensive analysis of the effects of declining unionization on US wage inequality over the last two decades. Figure 1 shows that the 50/10 wage gap decreased 5.7% from 1990 to 2000 and 2.6% from 2000 to 2010 and we find that declining unionization can explain approximately 10% and 23% of those changes in 50/10 wage gap.

This paper contributes to the previous literature by introducing a new semi-parametric estimation method, Yeo-Johnson unconditional quantile regression which is based on an extended Box-Cox power transformation, to evaluate the impact of changes in the distribution of X on the unconditional quantile of Y. By construction of the Yeo-Johnson transformation, the linear unconditional quantile regression model is a special case of the proposed method. We show that the asymptotic properties of the proposed Yeo-Johnson unconditional quantile partial effect estimator. The asymptotic properties establish the validity of the inference procedure. In addition, we empirically show that declining unionization can explain approximately 10% to 23% of the decline in the 50/10 wage gap in the period 1990–2000 and 2000–2010 respectively. To the best of our knowledge no other study in wage inequality literature has been able to explain the decline in 50/10 wage gap.

The structure of the paper is as follows. In Section 2, we present the Yeo-Johnson RIF regression model, discuss the estimation, show the asymptotic properties for the proposed Yeo-Johnson unconditional quantile partial effect estimator, and establish the relationship between proposed and existing estimators of the unconditional quantile partial effect. In Section 3, we discuss an empirical application of the Yeo-Johnson unconditional quantile regression model. We conclude the paper in Section 4. All derivations are shown in appendix.

2. Econometric model

Our parameter of interest, the Unconditional Quantile Partial Effect (UQPE) is defined as

$$\int \partial_x E\Big(RIF\big(Y, q_\tau\big)|X = x\Big)\,\mathrm{d}F(x) \tag{1}$$

where RIF (Recentered Influence Function) is $q_\tau + \tau - \mathcal{I}\{Y \leq q_\tau\}/f_Y(q_\tau)$; q_τ is the τth population quantile function and $\mathcal{I}(\cdot)$ is a binary variable which takes the value 1 when the outcome variable is less than or equal to q_τ and $f_Y(q_\tau)$ is the density function of Y evaluated at q_τ. Firpo *et al.* [26] discuss the estimation of $E(RIF(Y, q_\tau)|X = x)$ by nonparametric method and as well as some ad-hoc parametric specification such as RIF-OLS and RIF-Logit models.

The RIF-OLS model produces an estimate of $E(RIF(Y, q_\tau)|X = x)$ by running a linear regression of a sample analogue of $RIF(Y, q_\tau)$ on X. The estimated coefficients from this regression are used to estimate the average derivative vector as shown in Equation (1). The linear specification of $E(RIF(Y, q_\tau)|X = x)$ implies that a small change t in covariate X_j simply shifts the location of the distribution of Y by $\beta_j.t$ and leaves all other features of the distribution unchanged. The linear specification has some undesirable properties in many empirical applications. For example, if there is an increase in the variance but not the mean of a covariate then the linear RIF regression finds no effect in any quantile of the outcome variable.

It is also questionable whether a linear specification performs well in practice given that $E(RIF(Y, q_\tau)|X = x)$ can be highly nonlinear even in simple settings. An alternative approach would involve not constraining the model to be linear in probability or logistic but rather allowing the data to determine the exact transformation that results in a linear relationship between Y and X.[7] Box and Cox [18] introduce one such transformation which has gained popularity in economics because it contains linear and log-linear cases as special forms. For nonnegative values of Y, the Box–Cox transformation is defined as

$$Y_{\lambda_\tau} = \begin{cases} \big(Y^{\lambda_\tau} - 1\big)/\lambda_\tau & \text{if } \lambda_\tau \neq 0 \\ \log(\mathrm{Y}) & \text{if } \lambda_\tau = 0, \end{cases} \tag{2}$$

where Y_{λ_τ} is the transformed variable and λ_τ is the Box–Cox transformation shape parameter. The asymptotic properties of Box–Cox estimators are studied by Amemiya and Powell [4] and Bickel and Doksum [15].

In general, we use the Box–Cox transformation in the settings where the dependent variable is highly skewed. Although in a quantile context $RIF(\cdot)$ for a given value of $Y = y$ is binary, the Box–Cox transformation is still applicable because the highly skewed distribution of Y makes the conditional expectation function of $RIF(\cdot)$ also highly nonlinear.[8] Therefore, in such scenarios the Box–Cox model is more suitable than a linear specified model. The Box–Cox RIF regression specification has also an advantage over RIF–Logit model because the Box–Cox transformation does not impose any distributional assumption to specify $E(RIF(Y, q_\tau)|X = x)$. In this setup, we can think of only one extreme case where X consists of a set of saturated dummies, the RIF function will be linear, regardless of how the distribution of Y varies with X.

Wooldridge [52] studies the Box–Cox transformation of the mean regression model. Powell [48] shows that the Box–Cox transformation model can be estimated and interpreted in the framework of quantile regression. Chamberlain [22] and Buchinsky [19] develop the two-step CBTS (Chamberlain Buchisnsky Two-Step) estimators by exploiting the equivariance property of quantiles.[9] Machado and Mata [44] discuss the asymptotic properties of CBTS estimators and argue that Box–Cox transformation is especially applicable when the heterogeneity of the conditional distribution of outcome variables is not captured solely by location shifts.

Although Box–Cox transformation is widely used, one of the major limitations of this transformation is that it is valid only for positive values of Y. To circumvent this problem, [15] introduce signed power transformation to include negative values of Y. Atkinson [5] shows the standard asymptotic results of the maximum likelihood method are not applicable for such signed power transformation because the range of the distribution is determined by the unknown shape parameter.

Yeo and Johnson [54] have proposed a new family of distributions that can be used without restrictions on Y that have many of the good properties of the Box–Cox family. These transformations are defined by

$$Y_{\lambda_\tau} = \begin{cases} \left((Y+1)^{\lambda_\tau} - 1\right)/\lambda_\tau & \text{if } \lambda_\tau \neq 0, Y \geq 0 \\ \log(\mathrm{Y}+1) & \text{if } \lambda_\tau = 0, Y \geq 0 \\ -\left[(-Y+1)^{2-\lambda} - 1\right]/(2-\lambda) & \text{if } \lambda_\tau \neq 2, Y < 0 \\ -\log(-Y+1) & \text{if } \lambda_\tau = 2, Y < 0 \end{cases} \tag{3}$$

As shown in Equation (3) when Y is strictly positive, the Yeo–Johnson transformation is the same as the Box–Cox power transformation of $(Y+1)$. Also, when Y is strictly negative, Yeo–Johnson transformation is the Box–Cox power transformation of $(-Y+1)$ with power $2-\lambda$. Therefore, similar like Box–Cox transformation, $Y_{\lambda_\tau} \geq 0$ for positive values of Y and $Y_{\lambda_\tau} < 0$ for strictly negative values of Y.[10]

2.1. Identification

Our aim is to study the effect of a marginal change in X on some features of the distribution of Y while holding everything else constant. Thus we consider the following RIF regression model introduced by Firpo *et al.* [26]:

$$RIF\big(Y, q_\tau\big) = E\Big(RIF(Y, q_\tau)|\; X = x\Big) + \varepsilon_\tau \tag{4}$$

where $E(RIF(Y, q_\tau)|X = x)$ is the conditional expectation of $RIF(\cdot)$ and ε_τ is the error term. Following Firpo *et al.* [26], we also assume that X and ε_τ are independent and also $E(\varepsilon_\tau(q_\tau)|X = x) = 0$.[11]

This strong strict exogeneity assumption may not hold in different scenarios especially when the dependent variable is binary. However, [33] shows that in this setup when X is potentially an endogenous regressor, we can still consistently estimate the UQPE by using the control function approach introduced by Imbens and Newey [38]. For the [54] transformation, the conditional expectation of $RIF(Y, q_\tau)$ given $X = x$ can be written as

$$E\Big(RIF(Y, q_\tau)|X = x\Big) = \begin{cases} \Big(\big(x^T\beta_\tau\big)^{\lambda_\tau} - 1\Big)/\lambda_\tau & \text{if } \lambda_\tau \neq 0,\ x^T\beta_\tau \geq 0 \\ \log\big(x^T\beta_\tau\big) & \text{if } \lambda_\tau = 0,\ x^T\beta_\tau \geq 0 \\ \Big(\big(-x^T\beta_\tau\big)^{2-\lambda_\tau} - 1\Big)/(2-\lambda_\tau) & \text{if } \lambda_\tau \neq 2,\ x^T\beta_\tau < 0 \\ -\log\big(-x^T\beta_\tau\big) & \text{if } \lambda_\tau = 2,\ x^T\beta_\tau < 0 \end{cases} \tag{5}$$

where λ_τ and β_τ are the true parameters and x contains the constant term. Note that when $\lambda_\tau = 1$, Equation (5) yields the linear RIF regression model studied by Firpo *et al.* [26].

Since in this study, we are only interested about the distributional statistic quantile, the expression for the unconditional partial effects in Equation (1) simplifies to

$$\begin{aligned} &\int \partial_x E\big(RIF(Y, q_\tau)|X = x\big)\, \mathrm{d}F(x) \\ &= \frac{1}{f_Y(q_\tau)} \int \partial_x\Big(Pr\big[Y > q_\tau|X = x\big]\Big)\, \mathrm{d}F(x) \end{aligned} \tag{6}$$

The integral in Equation (6) is the average marginal effects of the covariates in a probability response model $Pr(Y > q_\tau|X = x)$. Firpo *et al.* [26] show that both *UQPE* and *CQPE* (conditional quantile partial effect) are equal to the parameter β_τ of the structural form for any quantile in a linear RIF regression model. However, this relation does not hold beyond the linear specification of $E(RIF(Y, q_\tau)|X = x)$. The below result shows the relationship between *UQPE* and the structural parameter β_τ for the Yeo–Johnson RIF regression model.

Proposition 2.1: *Assume that $E(RIF(Y, q_\tau)|X = x)$ is well defined for all values of $X = x$ and X is independent of the unobserved error term ε_τ. Then $UQPE(\tau)$ can be expressed as follows:*

$$UQPE(\tau) = \begin{cases} \beta_\tau\Big(x^T\beta_\tau\Big)^{\lambda_\tau - 1} & \textit{if } \lambda_\tau \neq 0,\ x^T\beta_\tau \geq 0 \\ \beta_\tau\Big(x^T\beta_\tau\Big)^{-1} & \textit{if } \lambda_\tau = 0,\ x^T\beta_\tau \geq 0 \\ -\beta_\tau\Big(-x^T\beta_\tau\Big)^{1-\lambda_\tau} & \textit{if } \lambda_\tau \neq 2,\ x^T\beta_\tau < 0 \\ -\beta_\tau\Big(x^T\beta_\tau\Big)^{-1} & \textit{if } \lambda_\tau = 2,\ x^T\beta_\tau < 0 \end{cases} \tag{7}$$

Proposition 2.1 implies that *UQPE* depends on τ and both the parameters β_τ and λ_τ except when λ_τ equals to 0 and 2. However, it does not depend on the distribution of ε_τ. As expected, when $\lambda_\tau = 1$ and $x^T\beta_\tau \geq 0$, *UQPE* equals the structural parameter β_τ because the linear model is a special case of the extended Box–Cox power transformation proposed by Yeo and Johnson [54].[12]

2.2. Estimation

Following [26], we first estimate the following re-centered influence function:

$$\widehat{RIF}\left(Y,\widehat{q}_\tau\right)=\widehat{q}_\tau+\frac{\left(\tau-\mathcal{I}\{Y-\widehat{q}_\tau\leq 0\}\right)}{\widehat{f}_Y\left(\widehat{q}_\tau\right)} \tag{8}$$

where $\widehat{q}_\tau$ is the sample quantile and $\widehat{f}_Y(\widehat{q}_\tau)$ is the estimated density function of Y evaluated at $\widehat{q}_\tau$. We need to estimate two unknown quantities $\widehat{q}_\tau$ and $\widehat{f}_Y(\widehat{q}_\tau)$ to estimate the $\widehat{RIF}(Y,\widehat{q}_\tau)$. The usual τth sample quantile can be represented as

$$\widehat{q}_\tau=\underset{q_\tau}{\operatorname{argmin}}\sum_{i=1}^{N}\left(\tau-\mathcal{I}\left\{Y_i-q_\tau\leq 0\right\}\right)\left(Y_i-q_\tau\right) \tag{9}$$

The second unknown quantity $\widehat{f}_Y(\cdot)$ can be estimated using the following kernel density estimator:

$$\widehat{f}_Y\left(\widehat{q}_\tau\right)=\frac{1}{h_b}\sum_{i=1}^{N}\mathcal{K}_Y\left(\frac{Y_i-\widehat{q}_\tau}{h_b}\right) \tag{10}$$

where $\mathcal{K}_Y(\cdot)$ is a kernel function and h_b is the bandwidth parameter. Finally, under some regularity conditions as shown in [53] and also stated in the proof of Proposition 2.2, the parameter vector $\theta_\tau=(\beta_\tau\lambda_\tau)^T$ can be consistently estimated by using the nonlinear least squares method:

$$\hat{\theta}_\tau=\arg\min_{\theta_\tau}\begin{cases}\sum_{i=1}^{N}\left(\widehat{RIF}\left(Y,\hat{q}_\tau\right)-\left(\left(x_i^T\beta_\tau\right)^{\lambda_\tau}-1\right)/\lambda_\tau\right)^2 & \text{for } \lambda_\tau\neq 0,\ x^T\beta_\tau\geq 0\\ \sum_{i=1}^{N}\left(\widehat{RIF}\left(Y,\hat{q}_\tau\right)-\log\left(x_i^T\beta_\tau\right)\right)^2 & \text{for } \lambda_\tau=0,\ x^T\beta_\tau\geq 0\\ \sum_{i=1}^{N}\left(\widehat{RIF}\left(Y,\hat{q}_\tau\right)\right. & \\ \left.+\left(\left(-x_i^T\beta_\tau\right)^{2-\lambda_\tau}-1\right)/(2-\lambda_\tau)\right)^2 & \text{for } \lambda_\tau\neq 2,\ x^T\beta_\tau<0\\ \sum_{i=1}^{N}\left(\widehat{RIF}\left(Y,\hat{q}_\tau\right)+\log\left(-x_i^T\beta_\tau\right)\right)^2 & \text{for } \lambda_\tau=2,\ x^T\beta_\tau<0\end{cases} \tag{11}$$

Suppose we denote the UQPE parameter as $\alpha_\tau(q_\tau)$ and $\widehat{\alpha}_\tau(\widehat{q}_\tau)$ is the corresponding estimator obtained from the extended Box–Cox power transformation RIF regression model. Our next result shows that the UQPE estimator of the extended Box–Cox power transformation RIF regression model is asymptotically normal under some regularity conditions. We formally state this result in the following proposition.

Proposition 2.2: *Suppose that the Assumptions A1–A6 hold.*[13] *Then, for all* $\tau\in(0,1)$,

$$\sqrt{Nh_b}\left(\widehat{\alpha}_\tau\left(\widehat{q}_\tau\right)-\alpha_\tau\left(q_\tau\right)\right)\xrightarrow{D}N\left(0,V_{bc}\left(q_\tau,\mathcal{K}\right)\right)$$

where $V_{bc}(q_\tau,\mathcal{K})=\lim_{h_b\to 0}((1/h_b)E[\psi_{bc}(q_\tau,h_b,\mathcal{K})\psi_{bc}(q_\tau,h_b,\mathcal{K})^T])$, $\mathcal{K}(\cdot)$ *is a symmetric real-valued kernel function,* h_b *is the bandwidth parameter, and* $\psi_{bc}(\cdot)$ *is defined in the proof.*

The proof of Proposition 2.2 is shown in the appendix. This asymptotic property of $\widehat{\alpha}_\tau(\widehat{q}_\tau)$ can be used for conducting the inference.

3. Empirical application

In this section, we study one of the most important and relevant issues in the US labor market: the trend of growing income inequality over the last three decades.

3.1. Hollowing out of the middle class

A large body of literature, including [1–3,8,10,11,24,39,40,43], and [17], investigates the impacts of different labor market factors contributing to these changes in US income inequality. In this section, we focus on the open question as to why the income gap between the middle class and the poor has been falling since the early 1980s. This phenomenon in labor economics is also known as the *hollowing out of the middle class*. Autor *et al.* [7] characterize this pattern as the polarization of the US labor market because median-income earners are pushed toward the left tail of the US income distribution. Goos and Manning [36] show that the UK has been also experiencing a similar pattern of changes in income inequality.

3.1.1. Wage gap puzzle

Katz and Murphy [40] and Autor *et al.* [9] propose an explanation based on relative demand growth for high-skilled workers and fluctuations in relative skill supplies based on a canonical supply–demand model. Autor *et al.* [8] argue that explanation based on returns to education and demand for skilled labor cannot explain why middle-class workers are worse off than low-skilled workers because middle-class workers have higher levels of human capital.[14] For the same reason, demand for skilled workers and skill-biased technological changes also cannot explain the 8% fall in the 50/10 wage gap.

Goldin and Katz [35] show that skill-biased technological changes (SBTC) continually increase the demand for high-skilled workers. The excess demand for high-skilled labor increases the earnings gap between the top and bottom tails of the income distribution.[15] An explanation based on SBTC and excess demand for skilled workers can explain why the 90/50 and 90/10 wage gaps are increasing but not why the 50/10 wage gap has been falling over time. In this application, we investigate the role of declining unionization on the fall in the 50/10 wage gap since the early 1990s.

3.1.2. Data and hypothesis

We use Current Population Survey (CPS) data, which contain income and demographic information for individual workers. The data are drawn from 5 percent self-weighted Outgoing Rotation Group (ORG) Supplements of the 1980–2010 samples. The data consist of US-born black and white men, aged 18–65, with positive annual earnings and hours worked in the year preceding the census and a nonzero sampling weight. The wage measure used is an hourly wage measure computed by dividing annual earnings by hours of work for workers not paid hourly. For workers paid hourly, we use a direct measure of the hourly wage rate. Annual income is expressed in 2010 US dollars using the Personal Consumption Expenditures price index.

Table 1. Summary statistics (mean and standard deviation) of the current population survey from 1990 to 2010.

	Period 1990 to 2000				Period 2000 to 2010			
	Combined	Year 1990	Year 2000	Difference (*t*-stat)	Combined	Year 2000	Year 2010	Difference (*t*-stat)
No. of observations	67,654	36,645	31,009		74,272	31,009	43,263	
Log wage (hourly)	2.903	2.899	2.907	−0.008	2.928	2.907	2.942	−0.035
	(0.691)	(0.676)	(0.708)	(−1.508)	(0.729)	(0.708)	(0.743)	(−6.443)
Union	0.206	0.225	0.182	0.043	0.165	0.182	0.148	0.034
	(0.404)	(0.418)	(0.386)	(6.418)	(0.372)	(0.386)	(0.355)	(5.217)
Education	13.093	13.025	13.175	−0.150	13.375	13.175	13.518	−0.344
	(3.060)	(3.060)	(3.059)	(−6.351)	(2.985)	(3.059)	(2.922)	(−15.513)
Experience	18.820	18.172	19.585	−1.413	20.306	19.585	20.823	−1.237
	(12.183)	(12.351)	(11.936)	(−15.059)	(12.093)	(11.936)	(12.179)	(−13.770)
Black	0.084	0.082	0.086	−0.004	0.097	0.086	0.104	−0.019
	(0.277)	(0.275)	(0.280)	(−1.701)	(0.295)	(0.280)	(0.097)	(−8.432)
Married	0.644	0.654	0.633	0.020	0.638	0.633	0.642	−0.009
	(0.479)	(0.476)	(0.482)	(5.528)	(0.480)	(0.482)	(0.479)	(−2.472)
Non-urban	0.261	0.272	0.247	0.025	0.246	0.247	0.246	0.001
	(0.439)	(0.445)	(0.431)	(7.244)	(0.431)	(0.431)	(0.431)	(0.436)
Skilled occupations	0.244	0.233	0.257	−0.024	0.265	0.257	0.271	−0.014
	(0.429)	(0.423)	(0.437)	(−7.182)	(0.441)	(0.437)	(0.444)	(−4.203)

The summary statistics table reports means and standard deviations (in parentheses) for the Current Population Survey from 1990 to 2010. The data consist of US-born black and white men aged 18–65. Hourly wage is expressed in 2010 US dollars using the personal consumption expenditure price index. The 'Difference' column reports the *t*-statistic (in parentheses) for the null hypothesis of equality in means. Union, black, married, non-urban and skilled occupation are dummy variables which take value equal to 1 when an individual is respectively a union member, African American, married, lives in a non-metropolitan area, and works in a skilled occupation. Note that the above-reported summary statistics for these five dummy variables can be used to find out the fraction of non-union, white, unmarried, urban and unskilled occupation because a dummy variable takes only two values.

The summary statistics are shown in Table 1. The sample sizes are 69,651, 59,750 and 84,239 for the years 1990, 2000 and 2010, respectively. The set of demographic variables consists of years of schooling, potential experience and experience square, union coverage, dummies for marital status, race, statistical metropolitan area and unskilled sector occupations. The last columns in both the right and left panels of Table 1 show the mean differences of these variables in two time periods, and t-statistics are shown in the parentheses. The mean differences are statistically significant for most of the variables in the two panels.

The top left panel of Figure 1 shows the scatter plot of 50/10 wage gaps from 1976 to 2012, and the dotted red line is the best linear fit. The fitted line shows a steady downward trend for the 50/10 wage gap since the early 1980s. This result is consistent with the studies of Autor *et al.* [11], Autor and Dorn [6], Firpo *et al.* [27], and Autor *et al.* [7], who show that the change in real wage follows a U-shaped pattern across the quantiles of income distribution since the 1980s. In contrast, the bottom left panel illustrates that the 90/50 wage gap has been increasing since the mid-1970s. These findings imply that the income gap between the rich and middle class is increasing, whereas the income gap between the middle class and the poor has been falling over time. This complex pattern of changes in which upper-tail wage inequality has increased and lower-tail wage inequality has fallen over time is known as the polarization of the US economy because the US labor market has been polarized in the direction of the two extremes.

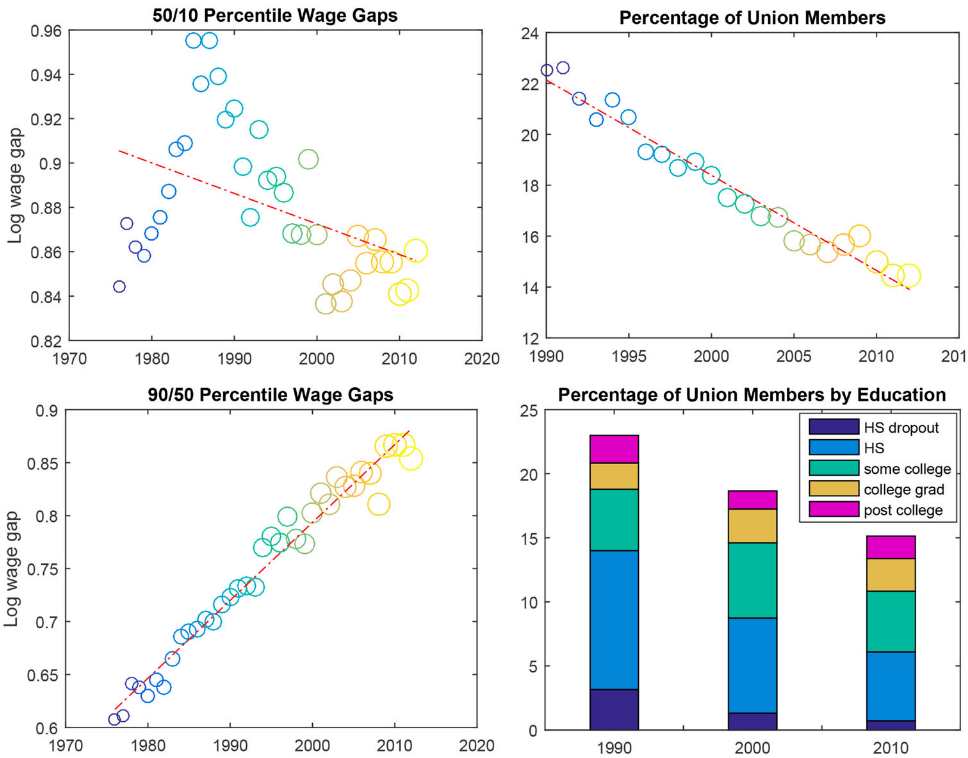

Figure 1. Changes in US wage inequality from 1976 to 2012 and the unionization rate from 1990 to 2010 by education.

Note: The 5 percent outgoing rotation groups from the Current Population Survey from 1976 to 2012 were used to estimate changes in US wage inequality. The two left panels show a scatter plot of the 50/10 and 90/50 wage gaps, and the red line indicates the best linear fit. The top right panel shows the scatter plot of the unionization rate from 1990 to 2010, and the bottom panel shows the decomposition of union members by five educational categories.

The top right panel of Figure 1 shows that the unionization rate for men has declined approximately 7% during the period from 1990 to 2010.[16] In the bottom right panel, we show the decomposition of this change by five education groups.[17] As shown, the relative proportions of union members with some college, those who graduated from college and those who pursued postgraduate studies remain almost the same in the last two decades. However, the share of high school graduate union workers decreased approximately 5% out of the total 7% decline in the unionization rate. Therefore, union workers who were high school graduates represent approximately 70% of the decline in the unionization rate over the period from 1990 to 2010.[18] In general, high school graduate union workers belong to the middle of the US income distribution. Because these workers are out of unions, their income has fallen, which may lead to a decrease in overall income in the middle of the income distribution. In this paper, we test the above hypothesis using the proposed Yeo–Johnson RIF regression model.

3.1.3. *Identification*

The RIF regression estimates of unionization are inconsistent because union and nonunion workers differ in their unobserved heterogeneity. Because there is no valid instrument for unionization, we use the distribution function based [46] decomposition method as proposed by DiNardo *et al.* [24] to address the potential endogeneity issue.[19] We note that over the last two decades apart from DiNardo *et al.* [24], many studies including [13,23,25–27,32,34,45,49–51] have extended the literature on distribution function based [46] decomposition method.[20]

Following the above-mentioned studies on distribution function based decomposition method, we use the following expression to estimate the effects of an individual covariate X^* at different quantiles of the US income distribution:

$$\text{Effects of } X^* = \nu\left(F_W^C\right) - \nu\left(F_W^t\right) \tag{12}$$

where ν is any distributional statistics such as mean, quantile, quantile difference. F_W^t is the wage distribution function in period $t \in \{0, 1\}$ and F_W^C is the counterfactual wage distribution which is defined as

$$F_W^C(w) = F\left(w; X_{t=0}^*, \tilde{X}_{t=1}\right), \tag{13}$$

where the set of covariates $X_t = [X_t^* \tilde{X}_t]$, and $\tilde{X}_t$ denotes all the other covariates except X_t^*.

In our application, X^* is the unionization variable. We denote $t = 0$ as our base period and $t = 1$ as the last period. Thus the counterfactual distribution function $F_W^C(w)$ corresponds to the distribution of wage that would have observed in $t = 1$ if the union coverage variable was distributed as in $t = 0$ and everything else remain the same as in 2010. To estimate the $\nu(F_W^t)$ and $\nu(F_W^C)$, we use the following result from Firpo *et al.* [26],

$$\nu(F_W) = E_X\Big(E\big(RIF(w, \nu)|X = x\big)\Big).$$

Therefore, we use both RIF-OLS and RIF–Yeo–Johnson regression methods to obtain the estimates of $E[RIF(w, \nu)|X = x]$ and also the sample analog for the expectation function. The estimation details of RIF–Yeo–Johnson method are discussed in Section 2.2. The estimated impacts from Equation (12) can be treated as causal effects under some regularity conditions. In the survey of decomposition methods, [29] discuss in detail under what regularity conditions [46] decomposition method estimates can be interpreted as causal effect.[21]

3.1.4. Results

The top panel of Table 2 shows the changes in overall US wage inequality, summarized by the 90/10 log wage differential; changes in wage inequality in the upper and lower halves of the distribution, summarized by the 90/50 and 50/10 log wage gaps (which are referred to as upper-tail and lower-tail wage inequality in the labor economics literature); changes in the 75/25 log wage gaps and changes in Gini coefficients. All these changes are decomposed in two time periods, 1990 to 2000 and 2000 to 2010. The estimation results are shown in the lower panel of Table 2.[22]

As shown in the upper panel of Table 2, the 50/10 wage gap decreased by 5.7% from 1990 to 2000 and by 2.6% from 2000 to 2010. Consistent with [6,11], and [7], we also find that wages remain almost the same in the lower tail of the income distribution. Therefore, the 50/10 wage gap fell by 5.7% in the 1990s because middle-class workers' real wage decreased significantly during that period. Alvaredo *et al.* [3] show that most of the benefits

Table 2. Effects of declining unionization on the change in US wage inequality from 1990 to 2010.

I. Wage inequality	Period 1990 to 2000					Period 2000 to 2010				
Wage gap	90/10	90/50	50/10	75/25	Gini	90/10	90/50	50/10	75/25	Gini
Δ	2.280	7.983	−5.697	0.227	−1.378	3.740	6.374	−2.641	1.242	−1.218
	(1.582)	(1.493)	(1.417)	(1.132)	(0.097)	(1.649)	(1.375)	(1.488)	(1.070)	(0.088)
II. Impacts of Unions										
Proposed YJUQR										
Estimated effects	−0.306	−0.887	0.580	−0.600	0.000	−0.400	−1.020	0.620	−0.597	−0.001
Standard errors	(0.002)	(0.001)	(0.001)	(0.003)	(0.000)	(0.004)	(0.002)	(0.002)	(0.001)	(0.000)
% of Δ	[0.134]	[0.111]	[0.102]	[2.645]	[0.000]	[0.107]	[0.160]	[0.235]	[0.480]	[0.001]
Linear UQR										
Estimated effects	0.028	−0.007	0.035	0.002	0.000	−0.012	−0.013	0.000	−0.066	−0.002
Standard errors	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)
% of Δ	[0.012]	[0.001]	[0.006]	[0.008]	[0.000]	[0.003]	[0.002]	[0.000]	[0.053]	[0.001]
Logit UQR										
Estimated effects	−0.825	−1.063	0.238	−0.761	−0.001	−0.735	−1.117	0.383	−0.828	−0.002
Standard errors	(0.002)	(0.001)	(0.004)	(0.002)	(0.000)	(0.001)	(0.003)	(0.004)	(0.003)	(0.000)
% of Δ	[0.361]	[0.133]	[0.042]	[3.355]	[0.000]	[0.197]	[0.175]	[0.145]	[0.667]	[0.002]
Nonparametric UQR										
Estimated effects	−1.265	−1.489	0.224	−0.933	0.000	−1.295	−1.560	0.265	−1.021	0.003
Standard errors	(0.004)	(0.004)	(0.003)	(0.004)	(0.000)	(0.001)	(0.005)	(0.002)	(0.003)	(0.000)
% of Δ	[0.553]	[0.187]	[0.039]	[4.113]	[0.000]	[0.347]	[0.245]	[0.100]	[0.822]	[0.003]

The table reports the estimation results from a distribution function based decomposition method. Following the recent studies related to distribution function-based decomposition method mentioned in Section 3.1.3, we use $\nu_\tau(F(w_t)) - \nu_\tau(F_c(w))$, to estimate the effects of unions on different features of the wage distribution function. $F(w_t)$ is the wage distribution function at period $t \in \{0, 1\}$, and $F_c(w)$ is the counterfactual wage distribution function. The counterfactual wage distribution $F_c(w)$ is defined as the only unionization variable that remains the same as in period $t = 0$, and all other variables are distributed as in $t = 1$ that is $F_c(w) = F(w; X^*_{t=0}, X'_{t=1})$. The naive bootstrapped standard errors (200 replications) for linear, logit, and nonparametric UQR and the proposed Yeo–Johnson unconditional quantile regression (YJUQR) are given in parentheses.

of the increase in US income have been captured by the top 10% since the early 1980s. These results suggest that the driving force of changing upper-tail wage inequality is a steep increase in income in the upper tail of the wage distribution.

The effects of unions on wage inequality can be decomposed into two parts. First, unions raise the wage of unionized low-skilled workers more than that of high-skilled workers and thus reduce wage inequality across different income groups. This is known as the inequality-decreasing 'within' effect.[23] Second, while unions reduce wage inequality within the union sector, they increase the wages of their members relative to the wages of non-unionized workers.[24] This is the inequality-increasing 'between' effect. The UQPE represents the sum of within and between effects of unionization, while the quantile regression measures only the within effect.[25]

The distribution function-based decomposition method results from the proposed Yeo–Johnson unconditional quantile regressions and also the [26] models are reported in bottom panels of Table 2.[26] In the left panel, the Yeo–Johnson unconditional quantile regression (YJUQR) estimate of 0.58 can be interpreted as follows: if the distribution of unionization remained the same in 1990, then the 50/10 wage gap would have increased by 0.58%. This implies that declining unionization can explain approximately 10% ($\approx$ 0.58/5.7) of the decline in the 50/10 wage gap from 1990 to 2000. Similarly, the fall in unionization rate can also explain approximately 23%($\approx$ 0.62/2.64) of the decline in the

50/10 wage gap from 2000 to 2010. Therefore, we conclude that declining unionization can explain a significant fraction of the downward trend of the 50/10 wage gap since the early 1990s.

In contrast, the linear unconditional quantile regression estimates suggest that declining unionization has almost no effect in explaining the 8% decline in the 50/10 wage gap from 1990 to 2010. Consistent with the Yeo–Johnson unconditional quantile regression (YJUQR) estimates, both the logit and nonparametric UQR estimates are significantly greater than zero although marginally lower than the estimates of our proposed approach. These results suggest that in this application the linear unconditional quantile regression model does not perform well most likely because the true conditional expectation function of $RIF(\cdot)$ is not linear.

4. Conclusion

In this paper, we introduce a semi-parametric estimation method to estimate the effect of changes in the distribution of X on the unconditional quantile of Y. The proposed Yeo–Johnson unconditional quantile regression consists of running a nonlinear least square regression of the RIF of Y on X. We show that the proposed Yeo–Johnson unconditional quantile partial effect estimator is consistent. By definition, in Yeo–Johnson transformation, [26]'s linear unconditional quantile regression is a special case of the proposed method. The main advantage of the proposed method over [26]'s RIF nonparametric approach is that we do not need to estimate a large number of parameters to make a polynomial approximation of $E[RIF(y, q_\tau)|X = x]$; instead, we estimate one additional Yeo–Johnson transformation parameter λ_τ using the nonlinear least square method.

In the empirical application, we investigate an open question in labor economics as to why the income gap between the middle class and the poor has been decreasing since the early 1980s. This phenomenon in labor economics is known as the *disappearing middle class*. Using the proposed Yeo–Johnson RIF-regression model, first we construct a counterfactual wage distribution function and then subtract that from the actual wage distribution function to estimate the impact of declining unionization on US wage inequality. We find that declining unionization can explain approximately 10–23% of this downward trend of the 50/10 wage gap since the early 1990s.

Notes

1. The RIF is defined in Section 2.
2. $Q_\tau(\cdot)$ is the conditional quantile function, β_τ is the regression parameter and τ denotes a quantile.
3. The unconditional quantile partial effect is defined in Section 2.
4. We use the extended Box–Cox power transformation proposed by Yeo and Johnson [54].
5. The support of X sufficiently rich implies that there exists a huge variation in X so that the choice of link function is not important to approximate the conditional distribution function $F_{Y|X}(y|x)$ arbitrarily well.
6. In this context implies that we refer to the situation where our dependent variables $RIF(\cdot)$ are binary for a given value of the outcome variable when we consider the distributional statistic quantile. $RIF(\cdot)$ for quantile is defined in Equation (1) and the detailed explanation are provided in Section 2.

7. As we mentioned earlier, the Firpo et al. [26] nonparametric model is a power approximation of logistic model specification.
8. $RIF(\cdot)$ is defined in Equation (1). As shown, for a given value of $Y = y$, $RIF(\cdot)$ takes value $q_\tau + (\tau - 1)/f_Y$ for $y \leq q_\tau$ and $q_\tau + \tau/f_Y$ otherwise.
9. Koenker and Park [42] develop an algorithm to implement the CBTS estimator. Fitzenberger et al. [28] suggest a modified version of the CBTS estimator.
10. Yeo and Johnson [54] show that the two transformation differ in their behavior the values of Y are close to zero, with Box–Cox transformation providing a much larger change for small values than does the Yeo–Johnson transformation.
11. Note that this model setup is similar to Firpo et al. [26] in which the authors also estimate the RIF regression model parameters using the strict exogeneity assumption.
12. Similarly, when $\lambda_\tau = 1$ and $x^T\beta_\tau < 0$, *UQPE* equals to -β_τ.
13. The assumptions A1–A6 are shown in the proof of Proposition 2.2.
14. In this context, skill refers to years of schooling.
15. Technological changes are measured by changes in the market price of cognitive and non-cognitive skills.
16. Unionization includes both union members and workers covered by unions.
17. The dark and light blue colors represent the high school dropouts and high school graduates; the green, yellow and purple colors represent union members who attended college, graduated from college and pursued postgraduate education.
18. During the period from 1990 to 2010, the unionization rate has declined 7.7% ($=$ 22.5% – 14.8%), and the proportion of high school graduate union workers declined 5.47% ($=$ 10.84% – 5.37%). Thus the contribution of high school graduate union workers is $\approx$ (5.47/7.7) $\times$ 100 $=$ 71%.
19. Over the last five decades of unionization literature, we still do not have a valid instrument.
20. Fortin et al. [29] provide a detail discussion on the methodological development of decomposition method since [46] and [16].
21. In this empirical setup, one of those assumptions implies that the distribution of workers' unobserved skills remains the same over the two periods. Thus taking the difference in wages at the same quantile of the wage distribution in two time periods yields a consistent estimate of the impact of unionization.
22. We report the Yeo–Johnson RIF regression estimates of β_τ and λ_τ for years 1990, 2000 and 2010 in Appendix Table 2. Because $\hat{\lambda}_\tau$ do not vary much across those three years for a given τ, we could apply the same Yeo–Johnson RIF regression model to construct counterfactual distributions for different years.
23. See [20] for details.
24. Card [21] shows that the wages of nonunion workers fall because of the increased supply of workers squeezed out from union jobs.
25. See [12] for a detailed discussion.
26. The third row of the bottom panel shows the contribution of unionization in terms of percentages to explain the change in the wage gap as reported in the top panel.

Acknowledgments

I am thankful to Christoph Rothe, Sergio Firpo, Hugo Jales, Aureo De Paula, Thomas Lemieux, David Autor, Anil Bera, Badi Baltagi, Thomas Kniesner, Le Wang, Carlos Lamarche, and Jeffrey Kubik as well as seminar participants at the North American Econometric Society Meeting, Royal Economic Society, Canadian Economic Association, European Association of Labour Economists Annual Conference, Syracuse University and Oklahoma University for their helpful comments and suggestions. Any errors or omissions are my own.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

[1] D. Acemoglu and D. Autor, *Skills, tasks, and technologies: implications for employment and earnings*, in *Handbook of Economics* IV(B), O. Ashenfelter and D. Card, eds., Amsterdam: North-Holland, 2011, pp. 1043–1172.

[2] D. Acemoglu and R. Pascual, *Low-skill and high-skill automation*, J. Hum. Cap. 12 (2018), pp. 204–232.

[3] F. Alvaredo, A. Atkinson, T. Piketty, and E. Saez, *The top 1 percent in international and historical perspective*, J. Econ. Persp. 27 (2013), pp. 3–20.

[4] T. Amemiya and J. Powell, *A comparison of the Box–Cox maximum likelihood estimator and the non-linear two stage least squares estimator*, J. Econom. 17 (1981), pp. 351–381.

[5] C. Atkinson, *Plots, Transformation and Regression*, Oxford University Press, Oxford, 1985.

[6] D. Autor and D. Dorn, *The growth of low-skill service jobs and the polarization of the us labor market*, Amer. Econ. Rev. 103 (2013), pp. 1553–1597.

[7] D. Autor, L. Katz, and M. Kearney, *The polarization of the U.S. labor market*, Rev. Econ. Stat. 96 (2006), pp. 189–194.

[8] D. Autor, L. Katz, and M. Kearney, *Trends in the U.S. wage inequality: revising the revisionists*, Rev. Econ. Stat. 90 (2008), pp. 300–323.

[9] D. Autor, L. Katz, and A. Krueger, *The polarization of the U.S. labor market*, Q. J. Econ. 113 (1998), pp. 1169–1213.

[10] D. Autor, F. Levy, and R. Murname, *The skill content of technological change: an empirical investigation*, Q. J. Econ. 118 (2003), pp. 1169–1213.

[11] D. Autor, A. Manning, and C. Smith, *The contribution of the minimum wage to us wage inequality over three decades: a reassessment*, Amer. Econ. J. Appl. Econ. 8 (2016), pp. 58–99.

[12] B. Baltagi and P. Ghosh, *Replication of unconditional quantile regressions*, J. Appl. Econ. 32 (1) (2017), pp. 218–223.

[13] B. Baltagi and P. Ghosh, *Partial distributional policy effects under endogeneity*, Sankhya B 81 (2019), pp. 123–145.

[14] R. Barsky, J. Bound, K. Charles, and J. Lupton, *Accounting for the black–white wealth gap: a nonparametric approach*, J. Am. Stat. Assoc. 97 (2002), pp. 663–673.

[15] P. Bickel and K. Doksum, *An analysis of transformations revisited*, J. Am. Stat. Assoc. 76 (1981), pp. 296–311.

[16] A. Blinder, *Wage discrimination: reduced form and structural estimates*, J. Human Resour. 8 (1973), pp. 436–455.

[17] J. Bound and G. Johnson, *Changes in the structure of wages in the 1980s: an evaluation of alternative explanations*, Am. Econ. Rev. 82 (1992), pp. 371–392.

[18] G. Box and D. Cox, *An analysis of transformation*, J. Royal Stat. Soc. 26 (1964), pp. 211–252.

[19] M. Buchinsky, *Quantile regression, Box–Cox transformation model, and the u.s. wage structure 1963–1987*, J. Econom. 65 (1995), pp. 109–154.

[20] D. Card, *The effect of unions on the structure of wages: a longitudinal analysis*, Econometrica 64 (4) (1996), pp. 957–979.

[21] D. Card, *Falling union membership and rising wage inequality: whats the connection?*, *NBER Working Paper No. 6520*, 1998.

[22] G. Chamberlain, *Quantile regression, censoring and the structure of wages*, in *Advances of econometrics*, C. Sims, ed., Cambridge University press, New York, 1994, pp. 171–209.

[23] V. Chernozhukov, I. Fernandez-Val, and B. Melly, *Inference on counterfactual distributions*, Econometrica 81 (2013), pp. 2205–2268.

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

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

[26] S. Firpo, N. Fortin, and T. Lemieux, *Unconditional quantile regressions*, Econometrica 77 (2009), pp. 953–973.

[27] S. Firpo, N. Fortin, and T. Lemieux, *Occupational tasks and changes in the wage structure*, *IZA working paper no. 5542*, 2011.
[28] B. Fitzenberger, R. Wilke, and X. Zhang, *Implementing Box–Cox quantile regression*, Econ. Rev. 29 (2) (2009), pp. 158–181.
[29] N. Fortin, T. Lemieux, and S. Firpo, *Decomposition methods in economics*, Handbook Labor Econ 4 (2011), pp. 1–102.
[30] M. Frolich, *Finite-sample properties of propensity-score matching and weighting estimators*, Rev. Econ. Stat. 86 (2004), pp. 77–90.
[31] M. Gaudry and R. Laferriere, *The Box–Cox transformation: power invariance and a new interpretation*, Econ. Lett. 30 (1989), pp. 27–29.
[32] P. Ghosh, *The contribution of human capital variables to changes in the wage distribution function*, Labour. Econ. 28 (2014), pp. 58–69.
[33] P. Ghosh, *Unconditional quantile regressions under endogeneity*, *Working Paper, University of Oklahoma*, 2016.
[34] P. Ghosh and J. Lee, *Decomposition of changes in Korean wage inequality, 1998–2007*, J. Labor. Res. 37 (2016), pp. 1–28.
[35] C. Goldin and L. Katz, *The Race Between Education and Technology*, Harvard University Press, Cambridge, 2008.
[36] M. Goos and A. Manning, *Lousy and lovely jobs: the rising polarization of work in Britain*, Rev. Econ. Stat. 89 (2007), pp. 118–133.
[37] V. Guerrero and R. Johnson, *Use of the Box–Cox transformation with binary response models*, Amer. Stat. Assoc. 69 (1982), pp. 309–314.
[38] G. Imbens and W. Newey, *Identification and estimation of triangular simultaneous equations models without additivity*, Econometrica 77 (2009), pp. 1481–1512.
[39] L. Katz and D. Autor, *Changes in the wage structure and earning inequality*, in *Handbook of Labor Economics*, 3, O. Ashenfelter and D. Card, eds., 1999, pp. 1463–1555.
[40] L. Katz and K. Murphy, *Changes in relative wages, 1963–1987: supply and demand factors*, Quater. J. Econ. 107 (1992), pp. 35–78.
[41] R. Koenker and G. Bassett, *Regression quantiles*, Econometrica 46 (1978), pp. 33–50.
[42] R. Koenker and B. Park, *An interior fixed point algorithm for quantile regressions*, J. Econ. 71 (1996), pp. 265–283.
[43] T. Lemieux, *Increasing residual wage inequality: composition effects, noisy data, or rising demand for skill?*, Am. Econ. Rev. 96 (2006), pp. 461–498.
[44] J. Machado and J, Mata, *Box–Cox quantile regression and the distribution of firm sizes*, J. Appl. Econ. 15 (2000), pp. 253–274.
[45] J. Machado and J. Mata, *Counterfactual decomposition of changes in wage distributions using quantile regression*, J. Appl. Econ. 20 (2005), pp. 445–465.
[46] R. Oaxaca, *Male-female wage differentials in urban labor markets*, Int. Econ. Rev. (Philadelphia) 14 (1973), pp. 693–709.
[47] D. Poirier, *The use of the Box–Cox transformation in limited dependent variable models*, Amer. Stat. Assoc. 73 (1978), pp. 284–287.
[48] J. Powell, *Estimation of monotonic regression models under quantile regressions*, in *Nonparametric and Semiparametric Methods in Econometrics and Statistics: Proceedings of the Fifth International Symposium on Economic Theory and Econometrics*, W. Barnett, J. Powell and G. Tauchen, eds., Cambridge University Press, New York, 1991, pp. 357–384.
[49] C. Rothe, *Nonparametric estimation of distributional policy effects*, J. Econom. 155 (2010), pp. 56–70.
[50] C. Rothe, *Partial distributional policy effects*, Econometrica 80 (2012), pp. 2269–2301.
[51] C. Rothe, *Decomposing the composition effect: the role of covariates in determining between-group differences in economic outcomes*, J. Business Econ. Stat. 33 (2015), pp. 323–337.
[52] J. Wooldridge, *Some alternative to the Box–Cox regression model*, Int. Econ. Rev. 33 (1992), pp. 935–955.
[53] C. Wu, *Asymptotic theory of nonlinear least squares estimation*, Ann. Stat. 9 (1981), pp. 501–513.
[54] I. Yeo and R. Johnson, *A new family of power transformations to improve normality or symmetry*, Biometrika 87 (2000), pp. 954–959.