ECINEQ
Society for the Study of Economic Inequality

Working Paper Series

Mincer earnings regression in the form of the double Pareto-lognormal model

Masato Okamoto

ECINEQ WP 2016 - 407

ECINEQ 2016 - 407
August 2016

www.ecineq.org

Mincer earnings regression in the form of the double Pareto-lognormal model

Masato Okamoto†
National Coalition of Independent Scholars (NCIS), U.S.A.

Abstract

In this study, the standard Mincer earnings regression equation in the form of the lognormal (LN) model is generalized into the form of the double-Pareto-lognormal (dPLN) model, substantially improving the goodness-of-fit to wage data. The empirical study contrasts the new and traditional models with respect to relationships between the wage and its determinant factors other than the primary equation for the conditional mean of log-wage, given potential work experience and education, such that, the wage distributions predicted by the dPLN-regression model faithfully reproduce the log-wage quantile regression results of the original data, whereas those by the LN-regression model fail such reproduction. Furthermore, the dPLN-regression model predicts that higher education has statistically significant positive effects on wage dispersion, particularly at the higher end, whereas the LN-regression model predicts insignificant negative effects even when heteroskedasticity in the error term is incorporated into the model. Thus, the new model is expected to be useful for not only accurately estimating contributions of wage determinant factors to wage dispersions and the shares of low-wage workers, but also improving the existing analysis methods using earnings equations such as the Oaxaca-Blinder decomposition and return of education by utilizing the dispersion regression equations.

Keywords: Distributional regression, heteroskedasticity, mixture distribution, quantile regression, wage dispersion.

JEL Classification: D31, D63, J31.

†**Contact details**: M. Okamoto, National Coalition of Independent Scholars (NCIS), San Antonio, TX, USA. E-mail: m_okamoto@ncis.org.

1 Introduction

Several parametric models that address the size distribution of income and earnings have been proposed and are thus far highly regarded, such as the Singh-Maddala distribution (Singh and Madalla 1975), Dagum distribution (Dagum 1977) and generalized beta distribution of the second kind, in short ‘GB2’ (McDonald 1984). Nevertheless, regression equations (abbreviated as ‘reg.’ and ‘eqs.’, respectively hereafter) in the form of the lognormal (LN) model (also called the ‘LN-regression model’), such as the Mincer earnings reg. eq., have been still utilized (often implicitly) to analyze the determinant factors of income and earnings. The better-fitted models have rarely been applied to reg. modeling[1], although the LN distribution does not necessarily fit well with income and earnings data. The reasons for the popularity of the LN linear reg. model likely include that the ordinary least square method is applicable to the model fitting after log-transforming the objective variable to obtain unbiased estimates of reg. coefficients (abbreviated as ‘coeffs.’ hereafter) in the primary eq. for the conditional mean of log-earnings of workers, given potential work experience and education (in years), even when the error term neither follows normal distribution nor satisfies homogeneity. However, when researchers attempt to estimate the effects of the determinant factors on other characteristics of the earnings distribution – such as the quantiles and dispersion – the LN-reg. model may possibly be unsuitable, even when heteroskedasticity in the error term is incorporated into the reg. model.

In this study, the Mincer earnings reg. eq. is generalized into a reg. eq. in the form of the double Pareto-lognormal (dPLN) model (or ‘dPLN-regression model’) to obtain the advantages of the dPLN distribution proposed by Reed (2003), such as its better goodness-of-fit to income/earnings distributions than the existing parametric models (Reed and Wu 2008; Okamoto 2012) and the parameters have clear roles on a basis of the LN model which allows natural generalization of the traditional LN-reg. model. An empirical study using Italian male wage data demonstrates that the dPLN-reg. model substantially improves the goodness-of-fit compared with that of the LN-reg. model. The wage distributions predicted by the dPLN-reg. model faithfully reproduce the results of log-wage quantile reg. applied to the original sample data, whereas the distributions predicted by the LN-reg. model fail the reproduction. Both types of reg. models are also contrasted with the predicted effects of wage determinant factors on wage dispersion, such that, whereas the LN-reg. predicts that

[1] Recently some researchers made attempts to apply more appropriate parametric models. Sohn et al. (2014) and Klein et al. (2015) analyzed the determinants of male earnings using the Dagum distribution as the base model although their regression models differ from the ordinary Mincer-type equations. They used the term ‘distributional regression’ for conditional wage distributions modeled using a parametric size distribution (more flexible than the LN) of which each parameter is regressed on covariates. In other research fields, the generalized gamma distribution is employed to model the relation between child age and clinical variables in a similar manner (Noufaily and Jones 2013). The major disadvantages of use of the existing beta- and gamma-type distribution models are that the roles of the parameters are unclear and the derived equations for the conditional mean are complicated.

longer education does not increase wage dispersion, the dPLN-reg. predicts the statistically significant positive effects on wage dispersion, particularly at the higher end. Such empirical findings highlight the importance of developing reg. models using a better-fit statistical distribution model as the base model for studies using earnings equations such as wage dispersion, the Oaxaca-Blinder decomposition and return of education.

The remainder of this article is organized as follows. In section 2, the LN-reg. model is generalized into the dPLN-reg. model, which is followed by the introduction of a procedure for fitting the models and measures for evaluating the goodness-of-fit. Then, Mincer-type LN- and dPLN-reg. models are fitted to Italian male wage data in the next section. Two LN-reg. and four dPLN-reg. models are devised based on whether and how they incorporate heteroskedasticity/heterogeneity in the error term. In addition to the usual goodness-of-fit comparisons, those reg. models are compared regarding the log-wage quantile reg. results of the predicted wage distributions and the predicted effects of the determinant factors on wage dispersion. The last section concludes with discussions and future issues. The formulas for the score functions and observed Fisher information matrices of the dPLN and dPLN-reg. models required for the maximum likelihood estimation as well as the fitting procedure in detail are deferred to appendices.

2 Methodology

2.1 Generalization of the Mincer earnings regression model

Reg. eqs. in the form of the LN model have been utilized (often implicitly) to analyze determinant factors of income and earnings, such as the Mincer earnings reg. eq., despite the fact that the LN distribution does not necessarily fit income and earnings data well. The LN distribution has the following probabilistic density function (pdf) and cumulative distribution function (cdf):

$$f_{LN}(y;\mu,\sigma) = \frac{1}{\sigma y}\phi\left(\frac{\log y-\mu}{\sigma}\right) \text{ and } F_{LN}(y;\mu,\sigma) = \Phi\left(\frac{\log y-\mu}{\sigma}\right) \text{ for } y>0, \qquad (1)$$

where μ denotes the location parameter and $\sigma(>0)$ denotes the dispersion parameter; moreover, $\phi(x) := \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right)$ and $\Phi(x) := \int_{-\infty}^{x}\phi(z)dz$ denote the pdf and cdf of the standard normal distribution, respectively. The first and higher order moments of the LN distribution are expressed as follows (*cf*. Kleiber and Kotz 2003):

$$E(Y^h) = e^{h\mu+\frac{1}{2}h^2\sigma^2}. \qquad (2)$$

The standard Mincer eq. regresses parameter μ on three regressors: years of potential work experience, the square of the years and years of education, as follows:

$$Y \sim LN\{\mu(\mathbf{x}), \sigma\} \text{ and}$$
$$\mu(\mathbf{x}) = \sum_i b_i x_i = b_0 + b_1 exp + b_2 exp^2 + b_3 educ, \tag{3}$$

where Y denotes a variable of earnings; $\mathbf{x} = \{x_0, x_1, x_2, x_3\} = \{1, exp, exp^2, educ\}$ denotes a regressor vector in which 'exp' and '$educ$' stand for years of potential work experience and education, respectively. Reg. eq. $\mu(\mathbf{x})$ shall be referred to the 'primary equation for the conditional mean' or simply 'primary equation' The ordinary least square (OLS) method can be applied to estimate partial reg. coeffs. b_0, b_1, b_2 and b_3. Model (3) is typically expressed as follows:

$$\log Y \sim b_0 + b_1 exp + b_2 exp^2 + b_3 educ + \varepsilon. \tag{4}$$

If Y strictly follows the LN-reg. model in (3), the error term ε follows a normal distribution with expectation zero and variance σ^2. The LN-reg. is widely used, even in the cases in which ε is neither normal nor homogenous, probably because the OLS estimates satisfy unbiasedness (under the assumption that ε is not correlated with regressors).[2] However, when researchers aim to estimate the effects of the determinant factors on the characteristics of earnings distributions other than the conditional means, the LN-reg. may possibly be unsuitable. To address this issue, this study shall examine wage reg. eqs. in the form of the dPLN model. The dPLN distribution has the following pdf and cdf:[3]

$$\begin{aligned} f_{dPLN}(y; \mu, \sigma, \alpha, \beta) \\ &= \frac{\alpha}{\alpha+\beta} \beta y^{\beta-1} e^{-\beta\mu+\frac{1}{2}\beta^2\sigma^2} \Phi^c\left(\frac{\log y - \mu + \beta\sigma^2}{\sigma}\right) \\ &+ \frac{\beta}{\alpha+\beta} \alpha y^{-\alpha-1} e^{\alpha\mu+\frac{1}{2}\alpha^2\sigma^2} \Phi\left(\frac{\log y - \mu - \alpha\sigma^2}{\sigma}\right) \text{ and} \end{aligned} \tag{5}$$

$$\begin{aligned} F_{dPLN}(y; \mu, \sigma, \alpha, \beta) \\ &= \frac{\alpha}{\alpha+\beta} y^{\beta} e^{-\beta\mu+\frac{1}{2}\beta^2\sigma^2} \Phi^c\left(\frac{\log y - \mu + \beta\sigma^2}{\sigma}\right) \\ &- \frac{\beta}{\alpha+\beta} y^{-\alpha} e^{\alpha\mu+\frac{1}{2}\alpha^2\sigma^2} \Phi\left(\frac{\log y - \mu - \alpha\sigma^2}{\sigma}\right) \\ &+ \Phi\left(\frac{\log y - \mu}{\sigma}\right) \quad \text{for } y > 0, \end{aligned} \tag{6}$$

[2] As the empirical example in this study uses microdata with weights for tabulation, the weighted least square method is applied instead. The corresponding prerequisite for the unbiasedness is that there is no weighted-correlation between ε and regressors.

[3] Reed (2003) derived the dPLN distribution from the assumption that individual log earnings follow Brownian motion with constant parameter μ and σ, and elapsed time from individual birth (entry into the labor market) to death (retirement from the labor market) follows an exponential distribution. However, Toda (2012) showed that individual earnings follow a dPLN distribution after a sufficiently long time has passed without the specific heterogeneity assumption if a mean-reverting independent transitory component exists with the Brownian motion.

where α and β are additional positive parameters; $\Phi^c(x) \coloneqq 1 - \Phi(x)$ is the complementary cdf of the standard normal distribution. The dPLN distribution follows power laws of orders determined by β and α in the left and right tails, as follows:

$$f_{dPLN}(y) \sim k_1 y^{\beta-1}\ (y \to 0) \text{ and } f_{dPLN}(y) \sim k_2 y^{-\alpha-1}\ (y \to \infty),$$

where k_1, k_2 denote positive constants. The dPLN distribution converges to the lognormal distribution with parameters μ and σ when $\alpha, \beta \to \infty$. The first and higher moments of the dPLN distribution are expressed as follows (Reed 2003):

$$E(Y^h) = \frac{\alpha\beta}{(\alpha - h)(\beta + h)} e^{h\mu + \frac{1}{2}h^2\sigma^2} \text{ for } \alpha > h. \tag{7}$$

The dPLN distribution has certain advantages as the base model for earnings reg. eqs.: Several researchers have shown that the dPLN is better fitted to income and earnings distributions than other parametric models such as the LN and GB2. Moreover, the parameters have clear roles on a basis of the LN model that allows the LN-reg. in (3) and (4) to be generalized naturally, so that the primary eq. for the conditional mean can be obtained in the same form as the LN-reg. Finally, the predicted earnings distribution has an analytic expression of the Gini index by regarding it as a mixture distribution of dPLNs (Okamoto 2012). Thus, the effects of the determinant factors on wage dispersion measured by the Gini index can be estimated from the predicted wage distribution.

The logarithm of a dPLN random variable Y follows a Normal-Laplace (NL) distribution (Reed and Wu 2008) as follows:

$$\log Y = \mu + \varepsilon = \mu + \sigma Z + \alpha^{-1}L_1 - \beta^{-1}L_2, \tag{8}$$

where $Z \sim N(0,1)$, i.e., a standard normal random variable, L_1 and L_2 follow exponential distributions with pdf e^{-x} and those three variables are assumed to be independent of one another. The expectation of the NL distribution in (8) is expressed as follows:

$$E(\log Y) = \mu + \alpha^{-1} - \beta^{-1}. \tag{9}$$

If $\varepsilon \coloneqq \sigma Z + \alpha^{-1}L_1 - \beta^{-1}L_2$ in (8) is regarded as the error term, then, the expectation of the error term is non-zero when $\alpha \neq \beta$. Thus, if only μ is regressed on $\mathbf{x}$ in the dPLN model, the reg. coeffs. do not agree with those of the LN-reg., i.e., the reg. of μ does not correspond to the primary eq. for the conditional mean of the LN-reg. When α and β are homogeneous, the difference emerges only at the intercept; otherwise, the difference also emerges with other coeffs. Let formula (8) be deformed as follows:

$$\log Y = \mu + \alpha^{-1} - \beta^{-1} + \varepsilon', \tag{10}$$

where $\varepsilon' \coloneqq \varepsilon - \alpha^{-1} + \beta^{-1}$, then, the expectation of ε' is zero. From eq. (10), it appears natural to regress the reciprocals of α and β on $\mathbf{x}$ in addition to μ as follows:

$$\begin{gathered} Y \sim dPLN\{\mu(\mathbf{x}), \sigma, 1/\,\alpha^{-1}(\mathbf{x}), 1/\,\beta^{-1}(\mathbf{x})\}, \\ \mu(\mathbf{x}) = \sum\nolimits_i c_i^{\mu} x_i\,, \alpha^{-1}(\mathbf{x}) = \sum\nolimits_i c_i^{\alpha} x_i \text{ and } \beta^{-1}(\mathbf{x}) = \sum\nolimits_i c_i^{\beta} x_i. \end{gathered} \tag{11}$$

The dPLN-reg. (11) creates asymptotic equivalences $b_i = c_i^{\mu} + c_i^{\alpha} + c_i^{\beta}$, which bring about the primary eq. of the dPLN-reg. Some discrepancies may arise in practice due to the misspecification, such as that the actual reg. functions are not strictly linear, but approximate equivalences are still expected to hold in the usual cases[4]. When σ is also heterogeneous, σ or its transformed value using an appropriate link function, such as a logarithm function, should also be regressed on $\mathbf{x}$. Thus, there can exist at most three reg. eqs. for the conditional dispersion of the error term. Nonetheless, the equivalence of the primary eqs. with the LN-reg. are expected to be (approximately) satisfied.

Despite the (approximate) equivalence of the primary eqs., there exist important differences between the LN- and dPLN-reg. models: The latter has two parameters, α and β, in addition to σ relating to the homogeneity/heterogeneity in the error term and that those two parameters also explicitly affect the primary eq. in contrast to σ. Thus, this parameter addition may affect the interpretation of the primary eq., such that reg. coeffs. do not represent the genuine returns of wage determinant factors; rather, the coeffs. include the contribution of population heterogeneity (see footnote 3). Note that, because of the equivalence property, the ambiguity in the interpretation also holds for the primary eq. of the LN-reg. when the dPLN-reg. model is appropriate as the actual wage distribution model.

2.2 Estimation procedure

The LN- and dPLN-reg. models briefly discussed in the previous subsection is fitted to wage data in such a manner as to maximize the following log-likelihood function:

$$\begin{aligned}\ell &= \sum_i w_i \log f\left[y_i; \mu\left(\mathbf{x}_i^{\mu}\right), \sigma, 1/\alpha^{-1}(\mathbf{x}_i^{\alpha}), 1/\beta^{-1}\left(\mathbf{x}_i^{\beta}\right)\right] \\ &= \sum_i w_i \log f\{y_i; \boldsymbol{\theta}(\mathbf{x}_i)\},\end{aligned} \tag{12}$$

where f denotes the pdf of the LN or dPLN distribution in (1) and (5); y_i denotes earnings of individual i; $\mathbf{x}_i^{\mu}$, $\mathbf{x}_i^{\alpha}$ and $\mathbf{x}_i^{\beta}$ denote values of regressors (only $\mathbf{x}_i^{\mu}$ for the LN-reg.) taken by individual i on which their respective parameters or reciprocals are regressed; and w_i denotes a weight for tabulation assigned to individual i. The weights are assumed to be normalized to make the total weights equal to the sample size n, i.e., $\sum_i w_i = n$. In the reg. models for the heterogeneous error term applied in the next section, the logarithm of σ is also regressed. In addition, the alternative link function (the logarithm of the reciprocal) is also considered for α and β in the next section. Notation $\boldsymbol{\theta}(\mathbf{x})$ is used as the general abbreviated notation of the reg. functions for all parameters, regardless of the link functions used.

To perform the maximum likelihood (ML) estimation, the Newton-Raphson (NR)

[4] Estimation bias can also be a major cause of the discrepancy when the sample size is small.

iterative algorithm is applied using the score function vector, denoted as ℓ' and the observed Fisher information matrix (FIM)[5] presented in Appendix 2. Beginning with an initial value vector $\boldsymbol{c}_0$ given appropriately, the NR algorithm iteratively updates the reg. coeffs. until convergence occurs as follows:

$$\boldsymbol{c}_k = \boldsymbol{c}_{k-1} - \rho \cdot FIM(\boldsymbol{c}_{k-1})^{-1}\ell'(\boldsymbol{c}_{k-1}) \text{ for } k = 1,2,\cdots, \quad (13)$$

where FIM^{-1} denotes the inverse matrix of the FIM, $\boldsymbol{c}_k$ denotes the tentative reg. coeff. vector at step k, and ρ represents for the adjustment multiplier, which is usually fixed to unity but reduced to less than unity in some cases to avoid convergence failure. In the case of reg. model (11), $\boldsymbol{c}_k$ consists of the tentative values of the following reg. coeffs. and parameter σ:

$$\boldsymbol{c} = \left\{c_0^{\mu}, c_1^{\mu}, c_2^{\mu}, c_3^{\mu};\ \sigma;\ c_0^{\alpha}, c_1^{\alpha}, c_2^{\alpha}, c_3^{\alpha};\ c_0^{\beta}, c_1^{\beta}, c_2^{\beta}, c_3^{\beta}\right\}.$$

Ways to create an initial value vector are crucial for the NR algorithm. In the empirical example of this study, the ML parameters/reg. coeffs. for a simpler model were used as part of an initial value vector for a more complex model and iterative partial maximization techniques were employed, which are described in Appendix 3. As for the reg. eq. of σ, the procedure in this study relied on the Nelder-Mead simplex algorithm (Nelder and Mead 1965) to obtain an initial value vector sufficiently close to the ML estimates due to the difficulty of finding an appropriate initial value vector for the NR algorithm.[6,7]

2.3 Evaluation measures for the goodness-of-fit

Two types of measures are employed to evaluate the goodness-of-fit of the reg. models. The first type of measures evaluate the suitability of the model for representing the relation between $\mathbf{x}$ and y. The AIC and BIC, which are the log-likelihood penalized based on the number of parameters/reg. coeffs. (denoted as '$\#\boldsymbol{\theta}$'), are used here.

$$AIC = -2\ell + 2 \cdot \#\boldsymbol{\theta},$$
$$BIC = -2\ell + \log n \cdot \#\boldsymbol{\theta}.$$

Another type of measures evaluate the proximity between the actual overall size distribution of y and that predicted by a reg. model. The relation between $\mathbf{x}$ and y does not matter explicitly for the second type of measures. The predicted size distribution of y is a mixture of $LN\{\boldsymbol{\theta}(\mathbf{x}_i)\}$ or $dPLN\{\boldsymbol{\theta}(\mathbf{x}_i)\}$ distributions with weights w_i s. Its pdf and cdf are

[5] To the author's best knowledge, the FIM formula of the dPLN has not been shown in the prior literature.
[6] The Nelder-Mead algorithm was performed using the 'optim' function implemented in the statistical analysis system R. Although no particular problem arises in the empirical example of this study, the Nelder-Mead algorithm becomes time-consuming with an increase of the number of regressors. This issue needs to be addressed in the future.
[7] A sample R-script for the ML estimation procedure and calculation of the evaluation measures, which is prepared for the empirical example in the next section, is available on the author's personal website (http://www.geocities.jp/stat_okamoto/).

expressed as follows:

$$f_M(y) = \frac{1}{\sum_i w_i} \sum_i w_i f\{y; \boldsymbol{\theta}(\mathbf{x}_i)\}, \tag{14}$$

$$F_M(y) = \frac{1}{\sum_i w_i} \sum_i w_i F\{y; \boldsymbol{\theta}(\mathbf{x}_i)\}. \tag{15}$$

The log-likelihood of this mixture distribution, ℓ_M (shown below), is employed as one of the second type of measures.

$$\ell_M = \sum_i w_i \log f_M(y_i). \tag{16}$$

Another of the second type of measures is the square root of the squared sum of errors between the empirical Lorenz curve of the overall earnings distribution and that predicted by the reg. model, denoted as 'L-RSSE'. Those additions are made because the predicted overall distribution, particularly with respect to its dispersion, is not necessarily accurate relative to those predicted by other reg. models and/or single statistical distribution models without reg. on $\mathbf{x}$ even when the first type of measures indicate its superiority to others. Let sequences $\{y_{[i]}\}$ and $\{w_{[i]}\}$ denote $\{y_i\}$ and $\{w_i\}$ arranged in ascending order of $\{y_i\}$, then, the L-RSSE is expressed as follows:

$$L\text{-}RSSE = \sqrt{\sum_i \{L_{[i]} - L_M(c_{[i]})\}^2}, \tag{17}$$

where $\{c_{[i]} = \sum_{k=1}^{i} w_{[k]} / \sum_{k=1}^{n} w_{[k]}\}$ denotes the cumulative population share and $\{L_{[i]} = \sum_{k=1}^{i} w_{[k]} y_{[k]} / \sum_{k=1}^{n} w_{[k]} y_{[k]}\}$ denotes the cumulative share of earnings, i.e., $(c_{[i]}, L_{[i]})$ corresponds to the (x, y) coordinate of the empirical Lorenz curve. The Lorenz curve of the predicted mixture distribution L_M is implicitly expressed as follows:

$$L_M(c) = F_M^{(1)}(z) = \frac{1}{\sum_i w_i m_i} \sum_i w_i m_i F^{(1)}\{z; \boldsymbol{\theta}(\mathbf{x}_i)\} \text{ and}$$
$$c = F_M(z),$$

where m_is denote the expectations of the LN or dPLN distributions comprising the predicted mixture distribution, i.e., the first moments in (2) or (7); $F^{(1)}(z; \boldsymbol{\theta}(\mathbf{x}_i))$ denotes the cdf of the first moment distribution of each component, expressed as follows for the LN distribution (*cf*. Kleiber and Kotz 2003):

$$F_{LN}^{(1)}(y; \mu, \sigma) = \Phi\left(\frac{\log y - \mu - \sigma^2}{\sigma}\right). \tag{18}$$

For the dPLN distribution, $F^{(1)}(z)$ is expressed as follows (Okamoto 2012 and 2013):

$$F_{dPLN}^{(1)}(y; \mu, \sigma, \alpha, \beta) = \frac{\alpha-1}{\alpha+\beta} y^{\beta+1} e^{-(\beta+1)\mu + \frac{1}{2}(\beta^2-1)\sigma^2} \Phi^c\left(\frac{\log y - \mu + \beta\sigma^2}{\sigma}\right) - \frac{\beta+1}{\alpha+\beta} y^{-\alpha+1} e^{(\alpha-1)\mu + \frac{1}{2}(\alpha^2-1)\sigma^2} \Phi\left(\frac{\log y - \mu - \alpha\sigma^2}{\sigma}\right) + \Phi\left(\frac{\log y - \mu - \sigma^2}{\sigma}\right). \tag{19}$$

Although the inverse cdf F_M^{-1} cannot be expressed explicitly, the univariate NR algorithm enables the calculation of $F_M^{-1}(c_{[i]})$.

The second type of measures, ℓ_M and L-RSSE, can be regarded as variants of measures employed by Okamoto (2012) to evaluate the goodness-of-fit of the overall income distributions expressed as a mixture distribution consisting of income distributions by age groups. Penalty for the number of parameters/reg. coeffs. is not imposed on either measure because, according to the simulation, the appropriate penalty on ℓ_M is likely to be very small even when really required (Okamoto 2012), and the L-RSSE is unlikely to require the penalty (Okamoto 2014).

Specific major inequality indices, i.e., the Gini index, second Theil index or equivalently mean log deviation (MLD), Theil index (T1) and squared coefficient of variation (SCV), are also calculated to examine the goodness-of-fit of the predicted overall distribution. The MLD and T1 of the LN distribution is equal to $\sigma^2/2$. The analytic formulas of those indices for the dPLN distribution were given by Okamoto (2012). The formulas of the SCV are derived from the first and second order moments in (2) and (7). Those indices of the LN and dPLN mixture distributions can be calculated using the well-known subgroup decomposition formula for the family of generalized entropy measures (Shorrocks 1980) as the build-up formula. The within-group and between-group inequality components in this formula represent inequalities within and between groups in which individuals have the same values for regressors $\mathbf{x}$. The derived two components of the inequality measures, including the Gini index, are also compared with those from the raw data.

3 Empirical example

3.1 Data

The LN- and dPLN-reg. models are fitted to public use microdata from the 2008, 2010, 2012 and 2014 Survey on Household Income and Wealth (SHIW), conducted by the Bank of Italy biennially, to analyze relations of payroll income (including fringe benefits) to the years of potential work experience and education of male employees under 55 years old. The reasons for using the SHIW data are that the microdata are publicly released without top-coding (i.e., high incomes are not replaced by artificial values) and are provided with replicate weights for estimating the sample variances consistently with the survey design. Only a few microdata satisfy those requirements.[8]

[8] The standard deviations (SDs) of data from ordinary (stratified) multistage sampling design are more than 1.5 times larger than those from the simple random sampling (SRS) in many cases (the SDs of the parameters/reg. coeffs. for the SRS are estimable using the FIM). To address this issue, 300–400 sets of replicate weights are provided to perform the Jackknife resampling procedure consistently with the SHIW survey design (Faiella 2008). As the replicate weights for the most recent 2014 SHIW have not been released, the statistical inference is omitted for 2014.

The sample size is approximately 2,500 for each year. Years of education is regarded as 3 for employees without educational qualification; 5, 8, 11, and 13 for those who have graduated from primary, lower secondary, vocational secondary, and upper secondary school, respectively; and 16, 18, 21 for those qualified with 3-year university, 5-year university, and postgraduate degrees, respectively. Years of potential work experience is calculated as age – (years of education + 6).

3.2 Variants of the earnings regression models applied

Two LN- and four dPLN-reg. models are fitted to the data. The LN-reg. model I, denoted as 'MLN I' by identifying it with the predicted LN mixture model, corresponds to the standard Mincer eq. (3). MLN II incorporates the heteroskedasticity in the error term by regressing the logarithm of σ using the same reg. function as $\mu(\mathbf{x})$ as follows:

$$\log\sigma(\mathbf{x}) = \varsigma(\mathbf{x}) = b_0^\sigma + b_1^\sigma exp + b_2^\sigma exp^2 + b_3^\sigma educ. \tag{20}$$

MdPLN I and II replace the LN model in MLN I and II with the dPLN model, respectively. In MdPLN II, σ is heterogeneous, whereas α and β are assumed to be homogeneous. MdPLN III fully incorporates the heterogeneity in the error term by regressing the reciprocals of α and β as well as $\log\sigma$ as follows:

$$\begin{aligned} \log\sigma(\mathbf{x}^\sigma) &= \varsigma(\mathbf{x}^\sigma) = c_0^\sigma + c_1^\sigma exp + c_3^\sigma educ, \\ \alpha^{-1}(\mathbf{x}) &= c_0^\alpha + c_1^\alpha exp + c_2^\alpha exp^2 + c_3^\alpha educ, \\ \beta^{-1}(\mathbf{x}) &= c_0^\beta + c_1^\beta exp + c_2^\beta exp^2 + c_3^\beta educ. \end{aligned} \tag{21}$$

MdPLN III′ replaces the link functions of α and β in MdPLN III with the logarithm of the reciprocals as follows:

$$\begin{aligned} -\log\alpha(\mathbf{x}) &= \tau(\mathbf{x}) = c_0^\alpha + c_1^\alpha exp + c_2^\alpha exp^2 + c_3^\alpha educ, \\ -\log\beta(\mathbf{x}) &= \upsilon(\mathbf{x}) = c_0^\beta + c_1^\beta exp + c_2^\beta exp^2 + c_3^\beta educ. \end{aligned} \tag{22}$$

In both models, the square of potential work experience years is excluded from the regressors in the reg. eq. of $\log\sigma = \varsigma(\mathbf{x})$ to avoid instability of the model fitting. MdPLN III can produces the primary eq. in the same form as $\mu(\mathbf{x})$ in MLN I and II, as explained in Section 2.

3.3 Comparisons of estimated regression models and goodness-of-fit of the models

Table 1 shows the results for the six reg. models fitted to the Italian male wage data. Significance tests of the reg. coeffs. were performed using the Jackknife replicate weights. In comparing the primary eq. of MdPLN III, $\mu(\mathbf{x}) + \alpha^{-1}(\mathbf{x}) - \beta^{-1}(\mathbf{x})$, with $\mu(\mathbf{x})$s of MLN I and II, there exist significant differences between the corresponding coeffs. at the

5% level for 2008 and 2010. Thus, the equivalence of the primary eqs. appears not to hold completely, but the differences remain small. If comparing $\mu(\mathbf{x})$s alone in the four MdPLN models with $\mu(\mathbf{x})$s in MLN I and II, the differences of the corresponding coeffs. are significant in most cases except for those of the squared terms of 'potential experience years'. However, the signs of all the coeffs. coincide with those in other models. By contrast, the MdPLN models clearly differ from MLN II as for the reg. eqs. of σ, such that the coeffs. for 'potential work experience years' are negative in MLN II, whereas they are positive in MdPLN II, III and III′, although the significance levels vary among the survey years; moreover, the coeffs. for 'education years' are insignificant in MLN II, whereas those are positive and in most cases significant at the 1% level in MdPLN II, III and III′. In the estimation results for the reg. eqs. of α and β in MdPLN III and III′, the coeffs. for 'education years' are positive in α's eq. and negative in β's eq. Those results are significant at the 5% level except for those in MdPLN III′ for 2012. As the squared term is also a regressor, a visible number of the coeffs. for 'potential work experience years' are insignificant at the 5% level, particularly in α's eq.; nevertheless, those are positive in α's and negative in β's eqs. for all survey years. In summary, the fitting results of the MdPLN models are by and large stable. From analytic formulas of inequality measures such as the MLD and T1 (Okamoto 2012), it can be posited that the coeffs. in α's and β's eqs. – with the opposite signs in MdPLN III and III′ – realize the complex relations between the regressors and wage dispersion, which are unable to be mimicked by the LN-reg. models, such that a value change of the same regressor in α's and β's eqs. contributes an increase of the (within-group) wage dispersion via one parameter, and simultaneously a decrease via another parameter.

In comparing the goodness-of-fit among the models, the MdPLN models substantially improve from the MLN models in terms of all four criteria, i.e., the AIC, BIC, ℓ_M and L-RSSE. Figure 1 illustrates that the pdfs of the MdPLN models closely resemble the density distribution of the original data relative to those of the MLN models.[9] Note that the pdf of MdPLN III′ is omitted from Figure 1 because the curve drawn by the pdf appears completely overlapped with that of MdPLN III. In comparing the four MdPLN models, model II, which incorporates the heterogeneity only in σ, outperforms model I in all four criteria. Models III and III′, which incorporate the heterogeneity in α and β in addition to that in σ, outperform model II in terms of the AIC, BIC and ℓ_M. As for the L-RSSE, model III attains smaller errors than model II for 2008, 2012 and 2014, whereas the converse holds for 2010. Thus, model III is also unlikely to be inferior to model II in terms of the L-RSSE. By contrast, model III′ attains a smaller error than model II only for 2008. Although model III′ tends to be even slightly better than model III in terms of the AIC, BIC and ℓ_M, the converse

[9] Using function 'density' implemented into the statistical analysis system R with the default options, the density distribution of the raw data was estimated by the kernel density estimation (KDE) method. The mode of the raw data in Table 2 and 3 derives from the estimated density distribution.

holds in terms of the L-RSSE except for 2010. Thus, model III′ is unlikely to be superior to models II and III in terms of the accuracy of the inequality estimation despite the superiority in terms of the three frequency-based evaluation measures.

Table 2 shows the mean, mode and major inequality indices estimated from the MLN and MdPLN models. MdPLN III′ predicts that the SCV is infinite for 2012 and 2014. The MLN models substantially underestimate the mode and overestimate other statistics, whereas the MdPLN models estimate them much more accurately; in particular, the errors of the Gini indices reduce to less than one-tenth. In comparing the MdPLN models, models III and III′ estimate the mean and MLD more accurately than models I and II, whereas the converse is true with respect to the mode. As for the accuracy of the Gini index and T1, similar tendencies are observed as with the L-RSSE; thus, a clear judgment of superiority or inferiority is hard to make regarding the estimation of the two indices.

In comparison with the MLN models, the MdPLN models also predict the ratios of the within-group inequality to the overall inequality closer to those in the original data. In particular, models III and III′, which fully incorporate the heterogeneity in the error term, tend to predict best among the MdPLN models.

The adjusted Gini index in Table 2 is defined as the Gini index after adjusting the mean in each group m_i to the overall mean m_M by uniformly multiplying all wages within the group by a constant value and by changing the population share of the group p_i to maintain the wage amount share $p_i m_i/m_M$; here, the groups are minutely classified according to the values of the regressors. Consider the following subgroup decomposition of the Gini index (Okamoto 2009):

$$G_M = \sum_i p_i \frac{m_i}{m_M} G_i + \frac{1}{m_M} \sum_{i<j} p_i p_j \int \{F_i(y) - F_j(y)\}^2 dy, \tag{23}$$

where $p_i = w_i/\sum_i w_i$ is the population share of group i; $m_i, G_i, F_i(y)$ represent the mean, Gini index and cdf in group i; and m_M represents the overall mean.[10] Then, the adjusted Gini index is decomposed as follows:

$$G_M^{adj} = \sum_i p_i \frac{m_i}{m_M} G_i + \frac{1}{m_M} \sum_{i<j} p_i p_j \frac{m_i}{m_M} \frac{m_j}{m_M} \int \left\{F_i\left(\frac{m_i}{m_M} y\right) - F_j\left(\frac{m_j}{m_M} y\right)\right\}^2 dy \tag{24}$$

The first term in (24), which correspond to the within-group component, is identical to that in (23). The second term in (24), which correspond to the between-group component, is non-negative; thus, the adjusted Gini index is equal to or larger than the within-group component. Provided that the cdfs in all groups become identical to one another by equalizing the means, then, the between-group component vanishes and the adjusted Gini index is equal to the within-group component. The MdPLN models come under this case

[10] The Gini index is also decomposable as $G_M = \sum_i p_i \frac{m_i}{m_M} G_i + \frac{1}{m_M} \sum_i p_i \int \{F_i(y) - F_M(y)\}^2 dy$. The within-group component is equivalent to that in the Gini stratification formula of Yitzhaki and Lerman (1991).

when wage distributions in all groups follow the dPLN distributions with common values for parameters σ, α, and β. As the between-group component in (24) measures the magnitude of differences of cdfs between groups in the form of a squared sum after equalizing the means, the component can be regarded as a measure for the magnitude of the heterogeneity in the error term predicted by the reg. models. Table 2 lists the ratios of the within-group components to the adjusted Gini indices (= 100% – the ratios of the between-group components) predicted by the reg. models. The ratios predicted by MdPLN III and III′ are smaller than those predicted by the other models; in other words, the models that fully incorporate the heterogeneity in the error term estimate that the magnitude of the heterogeneity is larger. Nonetheless, the ratios approximately 99% appear to underestimate the magnitude relative to the ratios approximately 98% that is estimated from the raw data. On this point, however, attention should be paid to the tendencies in the calculation when using a finite sample.[11] In practice, a simulation, conducted in accordance with the procedure for the calculation from the raw data, computes a 95% confidence interval ranging from 98.0 to 99.0% for the MdPLN III estimate. Thus, the difference is not statistically significant.

Finally, compared with the goodness-of-fit of major statistical size distributions (single distribution models without reg.) in Table 3, the four dPLN-reg. models are superior to the dPLN, GB2 and other single distribution models in terms of the AIC and BIC. As for the goodness-of-fit of the overall wage distribution predicted by the dPLN-reg. models, the log-likelihood values ℓ_Ms of MdPLN III and III′ are slightly larger than the maximum log-likelihood values of the single dPLN and GB2 except for 2010. Model III is also unlikely to be inferior to the single dPLN and GB2 in terms of the L-RSSE, whereas the converse holds for model III′. In comparing the pdfs between the single dPLN and MdPLN III in Figure 2, the density of the former is closer to that of the original data around the peak, whereas the peak location of the latter's is closer to that of the original density distribution. Note that the pdf of the single GB2 is omitted from Figure 2 because the curve drawn by the pdf looks completely overlapped with that of the single dPLN.

3.4 Quantile regression results when applied to wage distributions predicted by the regression models

[11] The within-group component and adjusted Gini indices of the raw data were approximated by the calculation for a subgrouping made by cross-classifying Italian male employees according to five grades of their educational qualification (i.e., no school or primary school; lower secondary; vocational secondary; upper secondary school; and university or higher degree) and four groups of the duration of the potential work experience, i.e., less than 10, 10-19, 20-29 and longer than or equal to 30 years. Three groups corresponding to less than 30 years of potential work experience and the lowest educational grade were collapsed into one group due to the small sample sizes. Thus, the entire sample was divided into 18 subsamples for the calculation.

The reg. eq. in the standard earnings reg. model (3) is fitted to wage distributions predicted by the reg. models using the quantile reg. method (Koenker and Bassett 1978) at quantiles 0.1, 0.5 and 0.9 in addition to the ordinary least square (LS) reg. method.[12] The LS reg. of the predicted wage distributions results in the primary eq., i.e., the coeffs. are identical to those of $\mu(\mathbf{x})$ in the cases of the MLN models and equal to those of $\mu(\mathbf{x}) + \alpha^{-1}(\mathbf{x}) + \beta^{-1}(\mathbf{x})$ in the cases of MdPLN I, II and III. As for MdPLN III′, the LS reg. is equivalent to the LS reg. of the expectations of log wage $\mu(\mathbf{x}_i) + e^{\tau(\mathbf{x_i})} - e^{\upsilon(\mathbf{x_i})}$s on $\mathbf{x}_i$s. The derived eq. corresponds to the primary eq. of the other models. The quantile reg. of the predicted wage distributions at quantile q is equivalent to the following minimization problem:

$$\underset{\boldsymbol{b}}{\operatorname{argmin}} \sum\nolimits_i w_i E\left\{\psi_q\left(\log Y_i - \sum\nolimits_j b_j x_j\right)\right\}, \qquad (25)$$

where $Y_i \sim LN\{\boldsymbol{\theta}(\mathbf{x}_i)\}$ or $dPLN\{\boldsymbol{\theta}(\mathbf{x}_i)\}$; $\psi_q(r)$ is defined as follows:

$$\psi_q(r) := \begin{cases} (q-1)r & \text{if } r \le 0 \\ qr & \text{if } r > 0 \end{cases}. \qquad (26)$$

The minimization problem (25) results in the following simultaneous eqs. for $\boldsymbol{b} = \{b_0, b_1, b_2, b_3\}$, to which the NR algorithm is applicable using the LS estimates as initial values:

$$0 = \sum\nolimits_i w_i x_{k,i}\left\{q - F_i\left(\sum\nolimits_j b_j x_j\,; \boldsymbol{\theta}(\mathbf{x}_i)\right)\right\}, \qquad k = 0,1,2,3 \qquad (27)$$

where $x_{k,i}$ represents the value of regressor k taken by individual i and $F_i(y)$ stands for the cdf of the normal or NL distribution with parameters $\boldsymbol{\theta}(\mathbf{x}_i)$. The cdf of the NL distribution is expressed as $F_{dPLN}\{e^y; \boldsymbol{\theta}(\mathbf{x}_i)\}$ using the cdf of the dPLN. Tables 4 presents the results for each year and the average of 2008, 2010 and 2012.[13]

In the results for the quantile regs. applied to the original data, the reg. coeff. at quantile 0.1 is larger than that at 0.5, and the coeff. at 0.9 is approximately equal to that at 0.5 in terms of ‘potential work experience years.’ The coeffs. at quantile 0.1 and 0.9 are larger than that at 0.5 in terms of ‘education years.’ The quantile regs. applied to MLN I and MdPLN I, which are reg. models formulated on the assumption of the homogeneous error term, produce equal coeffs. at all quantile points except those of the intercepts. Thus, both models fail to reproduce the results of the original data. The quantile regs. applied to MLN II and MdPLN II (i.e., the reg. model incorporating the heterogeneity in σ) also produce different results from those in the original data, such that, in the former case, the coeff. at quantile 0.9 is smaller than that at 0.5 as to ‘potential work experience years’ and, in the

[12] The quantile reg. was applied to the original data using function ‘rq’ implemented into package ‘quantreg’ of the statistical analysis system R. The significance tests are based on the standard deviations estimated by the Jackknife resampling procedure using the replicate weights.

[13] The quantile regression in (25)–(27) differs from ‘parametric quantile regressions’ including the method studied by Noufaily and Jones (2013). Their regression’s outcomes are conditional quantile functions directly derived from fitted regression models.

latter case, the coeff. at quantile 0.1 is smaller than that at 0.5 in terms of both 'potential work experience years' and 'education years.' Furthermore, as MdPLN II does not address heterogeneity in either α or β, the LS method also produces biased coeffs. By contrast, MdPLN III and III′ faithfully reproduce the quantile reg. results of the original data. The coeffs. including those of the intercepts at all quantile points and their differences between the quantile points do not differ significantly from those in the original data. Although the gaps in some coeffs. between the quantile and LS reg. differ from those in the original data significantly at the 5–10% level for 2010, all the differences are insignificant for 2008, 2012 and the average of 2008, 2010 and 2012.

The quantile reg. is applied on the assumption that the reg. coeffs. are heterogeneous such that the coeff. varies according to quantile point as follows:

$$\log y \sim b_0(q) + b_1(q) \cdot exp + b_2(q) \cdot exp^2 + b_3(q) \cdot educ + \varepsilon.$$

However, the results of MdPLN III demonstrate that the heterogeneity in the error term may make the reg. coeffs. look as if varying along with the quantile point, suggesting that this phenomenon actually arises in the application of quantile reg. to the real wage data. If that is the case, the interpretation of the results of the quantile reg. may differ depending on the cause of heterogeneity in the error term (see the last paragraph in section 2.1).

3.5 The effects of uniform change of wage determinants on wage dispersion estimated from the reg. models with heterogeneous error terms

The marginal change ratio of the mean or inequality index relative to the uniform infinitesimal change of a wage determinant factor shall be referred to as the 'effect' of the respective determinant factor x on the respective statistic I:

$$\text{Effect of x on } I \coloneqq \frac{\Delta \log I}{\Delta \mathrm{x}} = \frac{\Delta I / I}{\Delta \mathrm{x}}. \tag{28}$$

For example, provided that the number of years of education for each individual increases uniformly by $\Delta \mathrm{x}$ year, then the marginal change ratio of the Gini index relative to $\Delta \mathrm{x}$ is called the effect of education years on the Gini index. The effect of education years on any relative invariant inequality index is null in the case of the Mincer eq. model under the assumption of the homogeneous error term. Notably, the same is not the case with potential work experience years because its squared term is also a regressor in the model. Here, the four reg. models incorporating heterogeneity in the error term, out of the six reg. models, are compared with respect to the effects of the wage determinant factors because the aforementioned results indicate the existence of heterogeneity in the error term and this empirical study aims mainly at estimating the effects on the inequality indices.

Table 5 lists the effects of potential work experience and education years on Italian male wages estimated from MLN II, MdPLN II, III and III′. The table also includes each parameter's contribution to the effects, estimated by giving the infinitesimal uniform change of the respective factor only to the reg. eq. of the respective parameter. The effects on the mean wage are estimated as +1.5 - 2% for potential work experience years and +6 - 7% for education years. Although MdPLN III and III′, fully incorporating the heterogeneity in the error term, estimate the effects larger than MdPLN II, the estimates do not widely vary among MLN II and three MdPLN models. By contrast, the predicted effects on inequality indices exhibit sharp differences between MLN and MdPLNs. MLN II estimates that the effects of potential work experience years on inequality indices are negative. The effects on the MLD, Gini index and T1 are significant at the 5% level and that on the SCV is significant at the 10% level for the averages of 2008, 2010 and 2012. By contrast, the MdPLN models estimate that the effects on all inequality indices are insignificant at the 5% level. Although MdPLN II estimates that the effects on the T1 and SCV are significant at the 10% level, the estimates are positive. As for the effects of education years on inequality indices, MLN II estimates that those are negative but insignificant even at the 10% level, whereas the MdPLN models estimate that the effects are positive in most cases. For the averages of 2008, 2010 and 2012, MdPLN II estimates that the effects on all indices are significant at the 1% level. The significance level falls in the case of MdPLN III and III′; nonetheless, MdPLN III makes estimates the significant effects on the Gini index and T1 at 1% and those on the MLD and SCV at the 10% level; MdPLN III′ produces the significant effects on the Gini index and T1 at the 5% level and those on the MLD at the 10% level. The expanding effects of longer education years on inequality indices tend to be larger for indices sensitive to distributional changes in higher wage classes, such that, whereas the effects on the Gini index are equal to or less than +3%, those on the T1 are equal to or more than +6% (equal to or above the effects on the mean wage) and those on the SCV are more than +10% (substantially above the effects on the mean wage). Thus, the estimation from the MdPLN models imply that an increase in highly educated workers, such as university graduates, may contribute to increasing the wage dispersion.

MLN II estimates that both μ and σ contribute to the reducing effects of longer years of potential work experience on inequality indices (although the contributions of σ are evaluated as insignificant), whereas three MdPLN models estimate the reducing effects via μ are nullified by the expanding effects via σ. Furthermore, MdPLN III and III′ estimate that the expanding effects via α are cancelled out by the reducing effects via β. As for the effects of longer years of education on inequality indices, the MLN II estimates that the contributions of σ are negative but insignificant, whereas three MdPLN models estimate that the contributions of σ are positive and significant. MdPLN III and III′ estimate that the expanding effects via α are weakened by the reducing effects via β. The net contributions of the two parameters are insignificant or significant at loose statistical levels; nonetheless, the magnitudes of the net contributions are possibly non-ignorably large relative to those of the contributions via σ as to inequality indices sensitive to distributional changes in higher wage classes.

The results of MdPLN II are notable from the perspective of the reg. modeling for the following reasons: MdPLN II incorporates only the heterogeneity in σ. In this respect, the model resembles MLN II, whereas the estimated effects of the wage determinant factors on inequality indices show sharp differences from MLN II; instead, the estimates are similar to those of MdPLN III and III′, which fully incorporate the heterogeneity in the error term.

4 Discussion and concluding remarks

In this study, the traditional reg. eq. in the form of the LN model is generalized into that in the form of the dPLN model. The generalization substantially improves the goodness-of-fit of the Mincer reg. eq. to Italian male wage data. The primary eq. for the conditional mean of log-wage predicted by the dPLN-reg. is close to that predicted by the LN-reg. By contrast, sharp differences are observed between the two types of reg. models regarding the predicted relationships other than the conditional mean, such that the wage distributions predicted by the dPLN-reg. models incorporating the heterogeneity in the error term faithfully reproduce the log-wage quantile reg. results of the original data, whereas those predicted by the LN-reg. models fail the reproduction even when incorporating the heteroskedasticity in the error term. Furthermore, the two types of reg. models predict the relations of the wage dispersion to the wage determinant factors strikingly differently. Thus, the new model is expected to be useful, for example, for estimating contributions of wage determinants to wage dispersions and the shares of low-wage workers accurately. Furthermore, as the mean wage is determined by not only the mean of log-wage but also the dispersion (see (2), (7) and differences in the predicted overall means in Table 2), the new model has possibility to improve the existing analysis methods using earnings eqs. such as the Oaxaca-Blinder decomposition and return of education by utilizing the dispersion reg. eqs.

The author believes that the dPLN-reg. makes the Mincer equation closer to reality as never before; nonetheless, it should be listed as one of the future tasks to investigate reg. models based on the generalized models of the dPLN (Reed and Wu 2008; Okamoto 2014) or other appropriate statistical size distribution models (although the model fitting method is likely to become an issue due to the model complexity) because Figure 1 appears to show that there is room for further improvement of the reg. models.

To distinguish the genuine returns of wage determinant factors in the primary eq. from possible effects of the population heterogeneity, it is also a future task to incorporate panel data analysis methods into the dPLN-reg. model.

Appendix 1 Score function and observed Fisher information matrix for the dPLN

For presenting the formula of the score function and observed FIM of the dPLN, the pdfs of the left and right Pareto-lognormal distributions, f^L and f^R, are introduced:

$$f^L = f^L(y;\mu,\sigma,\beta) \coloneqq \beta y^{\beta-1} e^{-\beta\mu+\frac{1}{2}\beta^2\sigma^2} \Phi^c\left(\frac{\log y - \mu + \beta\sigma^2}{\sigma}\right) \text{ and}$$

$$f^R = f^R(y;\mu,\sigma,\alpha) \coloneqq \alpha y^{-\alpha-1} e^{\alpha\mu+\frac{1}{2}\alpha^2\sigma^2} \Phi\left(\frac{\log y - \mu - \alpha\sigma^2}{\sigma}\right).$$

Then, the pdf of the dPLN can be expressed as follows, using f^L and f^R:

$$f = f_{dPLN}(y;\mu,\sigma,\alpha,\beta) = \frac{\alpha}{\alpha+\beta} f^L(y;\mu,\sigma,\beta) + \frac{\beta}{\alpha+\beta} f^R(y;\mu,\sigma,\alpha).$$

In connection with f^L and f^R, six functions are additionally defined as follows:

$$g^L = g^L(y;\mu,\sigma,\beta) \coloneqq \beta y^{\beta-1} e^{-\beta\mu+\frac{1}{2}\beta^2\sigma^2} \phi\left(\frac{\log y - \mu + \beta\sigma^2}{\sigma}\right),$$

$$g^R = g^R(y;\mu,\sigma,\alpha) \coloneqq \alpha y^{-\alpha-1} e^{\alpha\mu+\frac{1}{2}\alpha^2\sigma^2} \phi\left(\frac{\log y - \mu - \alpha\sigma^2}{\sigma}\right),$$

$$m_{\alpha+} = m_{\alpha+}(y;\mu,\sigma,\alpha) \coloneqq \log y - \mu + \alpha\sigma^2,$$
$$m_{\alpha-} = m_{\alpha-}(y;\mu,\sigma,\alpha) \coloneqq \log y - \mu - \alpha\sigma^2,$$
$$m_{\beta+} = m_{\beta+}(y;\mu,\sigma,\beta) \coloneqq \log y - \mu + \beta\sigma^2 \text{ and}$$
$$m_{\beta-} = m_{\beta-}(y;\mu,\sigma,\beta) \coloneqq \log y - \mu - \beta\sigma^2.$$

In the case in which the logarithm of σ, $\varsigma = \log\sigma$, and the reciprocals of α and β are used as the parameters instead of σ, α and β for the dPLN, the score function, i.e. the vector of the first-order partial derivatives of the log-likelihood with respect to the parameters, is expressed as follows, using the first-order partial derivatives of f with respect to the parameters, denoted as f_μ, f_σ, $f_{\alpha^{-1}}$ and $f_{\beta^{-1}}$, respectively:

$$\ell'_{dPLN} = \sum\nolimits_i w_i D_{dPLN}(y_i;\mu,\varsigma,\alpha^{-1},\beta^{-1}) = \sum\nolimits_i w_i D_{dPLN}(y_i;\boldsymbol{\theta}),$$

$$D_{dPLN}(y;\boldsymbol{\theta}) = \left[d_\mu, d_\varsigma, d_{\alpha^{-1}}, d_{\beta^{-1}}\right]^T = \tfrac{1}{f} f_{\boldsymbol{\theta}}, \text{ where } f_{\boldsymbol{\theta}} = \left[f_\mu, f_\varsigma, f_{\alpha^{-1}}, f_{\beta^{-1}}\right]^T.$$

The observed FIM is expressed as follows, using the first- and second-order partial derivatives of f with respect to the parameters, denoted such as f_μ, $f_{\alpha^{-1}}$, $f_{\mu\mu}$ and $f_{\varsigma\beta^{-1}}$:

$$FIM_{dPLN} = \sum\nolimits_i w_i I_{dPLN}(y_i;\mu,\varsigma,\alpha^{-1},\beta^{-1}) = \sum\nolimits_i w_i I_{dPLN}(y_i;\boldsymbol{\theta}),$$

$$I_{dPLN}(y;\boldsymbol{\theta}) = -\begin{bmatrix} l_{\mu\mu} & l_{\mu\varsigma} & l_{\mu\alpha^{-1}} & l_{\mu\beta^{-1}} \\ l_{\mu\varsigma} & l_{\varsigma\varsigma} & l_{\varsigma\alpha^{-1}} & l_{\varsigma\beta^{-1}} \\ l_{\mu\alpha^{-1}} & l_{\varsigma\alpha^{-1}} & l_{\alpha^{-1}\alpha^{-1}} & l_{\alpha^{-1}\beta^{-1}} \\ l_{\mu\beta^{-1}} & l_{\varsigma\beta^{-1}} & l_{\alpha^{-1}\beta^{-1}} & l_{\beta^{-1}\beta^{-1}} \end{bmatrix} = -\frac{1}{f} f_{\boldsymbol{\theta\theta}} + \frac{1}{f^2} f_{\boldsymbol{\theta}} \cdot f_{\boldsymbol{\theta}}{}^T,$$

$$\text{where } f_{\boldsymbol{\theta\theta}} = \begin{bmatrix} f_{\mu\mu} & f_{\mu\varsigma} & f_{\mu\alpha^{-1}} & f_{\mu\beta^{-1}} \\ f_{\mu\varsigma} & f_{\varsigma\varsigma} & f_{\varsigma\alpha^{-1}} & f_{\varsigma\beta^{-1}} \\ f_{\mu\alpha^{-1}} & f_{\varsigma\alpha^{-1}} & f_{\alpha^{-1}\alpha^{-1}} & f_{\alpha^{-1}\beta^{-1}} \\ f_{\mu\beta^{-1}} & f_{\varsigma\beta^{-1}} & f_{\alpha^{-1}\beta^{-1}} & f_{\beta^{-1}\beta^{-1}} \end{bmatrix}.$$

In the above formulas, v^T represents the transpose vector of v. The first- and second-order partial derivatives of f are expressed as follows, using the functions introduced above:

$$f_\mu = \frac{\alpha\beta}{\alpha+\beta}(-f^L + f^R) + \frac{1}{\alpha+\beta}\frac{1}{\sigma}(\alpha g^L - \beta g^R),$$

$$f_\varsigma = \frac{\alpha\beta}{\alpha+\beta}\sigma^2(\beta f^L + \alpha f^R) + \frac{1}{\alpha+\beta}\frac{1}{\sigma}\{\alpha m_{\beta-} g^L - \beta m_{\alpha+} g^R\},$$

$$f_{\alpha^{-1}} = -\frac{\alpha^2\beta}{(\alpha+\beta)^2}\left[f^L + \left\{\frac{\beta}{\alpha} - (\alpha+\beta)m_{\alpha-}\right\}f^R\right] + \frac{\alpha^2\beta}{\alpha+\beta}\sigma g^R,$$

$$f_{\beta^{-1}} = -\frac{\alpha\beta^2}{(\alpha+\beta)^2}\left[\left\{\frac{\alpha}{\beta} + (\alpha+\beta)m_{\beta+}\right\}f^L + f^R\right] + \frac{\alpha\beta^2}{\alpha+\beta}\sigma g^L,$$

$$f_{\mu\mu} = \frac{\alpha\beta}{\alpha+\beta}(\beta f^L + \alpha f^R) + \frac{1}{\alpha+\beta}\frac{1}{\sigma^3}\{\alpha(-2\beta\sigma^2 + m_{\beta+})g^L - \beta(2\alpha\sigma^2 + m_{\alpha-})g^R\},$$

$$f_{\mu\varsigma} = \frac{\alpha\beta}{\alpha+\beta}\sigma^2(-\beta^2 f^L + \alpha^2 f^R) + \frac{1}{\alpha+\beta}\frac{1}{\sigma}\left[\alpha\left\{\frac{m_{\beta+}m_{\beta-}}{\sigma^2} - \beta m_{\beta-} + \beta^2\sigma^2 - 1\right\}g^L - \beta\left\{\frac{m_{\alpha+}m_{\alpha-}}{\sigma^2} + \right.\right.$$

$$\left.\left.\alpha m_{\alpha+} + \alpha^2\sigma^2 - 1\right\}g^R\right],$$

$$f_{\mu\alpha^{-1}} = \frac{\alpha^2\beta}{(\alpha+\beta)^2}[\beta f^L - \{\alpha + 2\beta - \alpha(\alpha+\beta)m_{\alpha-}\}f^R] - \frac{\alpha^2\beta}{(\alpha+\beta)^2}\frac{1}{\sigma}\left[g^L - \left\{\frac{\beta}{\alpha} + \alpha(\alpha+\beta)\sigma^2\right\}g^R\right],$$

$$f_{\mu\beta^{-1}} = \frac{\alpha\beta^2}{(\alpha+\beta)^2}[\{2\alpha + \beta + \beta(\alpha+\beta)m_{\beta+}\}f^L - \alpha f^R] - \frac{\alpha\beta^2}{(\alpha+\beta)^2}\frac{1}{\sigma}\left[\left\{\frac{\alpha}{\beta} + \beta(\alpha+\beta)\sigma^2\right\}g^L - g^R\right],$$

$$f_{\varsigma\varsigma} = \frac{\alpha\beta}{\alpha+\beta}\sigma^2\{\beta(2+\beta^2\sigma^2)f^L + \alpha(2+\alpha^2\sigma^2)f^R\} + \frac{1}{\alpha+\beta}\frac{1}{\sigma^3}[\alpha\{2\beta^2\sigma^4 m_{\beta-} - \sigma^2 m_{\beta+} +$$

$$m_{\beta-}^2 m_{\beta+}\}g^L + \beta\{-2\alpha^2\sigma^4 m_{\alpha+} + \sigma^2 m_{\alpha-} - m_{\alpha+}^2 m_{\alpha-}\}g^R],$$

$$f_{\varsigma\alpha^{-1}} = -\frac{\alpha^2\beta}{(\alpha+\beta)^2}\sigma^2[\beta^2 f^L + \alpha\{2\alpha + 3\beta - \alpha(\alpha+\beta)m_{\alpha-}\}f^R] - \frac{\alpha^2\beta}{(\alpha+\beta)^2}\frac{1}{\sigma}\left[m_{\beta-}g^L - \left\{(\alpha + \right.\right.$$

$$\left.\left.\beta)(\alpha^2\sigma^2 + 1)\sigma^2 + \frac{\beta}{\alpha}m_{\alpha+}\right\}g^R\right],$$

$$f_{\varsigma\beta^{-1}} = -\frac{\alpha\beta^2}{(\alpha+\beta)^2}\sigma^2[\beta\{3\alpha + 2\beta + \beta(\alpha+\beta)m_{\beta+}\}f^L + \alpha^2 f^R] + \frac{\alpha\beta^2}{(\alpha+\beta)^2}\frac{1}{\sigma}\left[\left\{(\alpha+\beta)(\beta^2\sigma^2 + 1)\sigma^2 - \right.\right.$$

$$\left.\left.\frac{\alpha}{\beta}m_{\beta-}\right\}g^L + m_{\alpha+}g^R\right],$$

$$f_{\alpha^{-1}\alpha^{-1}} = \frac{\alpha^4\beta}{(\alpha+\beta)^3}\left[2\frac{\beta}{\alpha}f^L + \left\{2\frac{\beta^2}{\alpha^2} + (\alpha+\beta)^2\sigma^2 - \frac{\alpha+\beta}{\alpha}(2\alpha + 4\beta)m_{\alpha-} + (\alpha+\beta)^2 m_{\alpha-}^2\right\}f^R\right] -$$

$\frac{\alpha^4\beta}{(\alpha+\beta)^2}\sigma\left\{\frac{2\alpha+4\beta}{\alpha}-(\alpha+\beta)m_{\alpha-}\right\}g^R,$

$f_{\alpha^{-1}\beta^{-1}}=\frac{\alpha^2\beta^2}{(\alpha+\beta)^3}\left[\{2\alpha+\beta(\alpha+\beta)m_{\beta+}\}f^L+\{2\beta-\alpha(\alpha+\beta)m_{\alpha-}\}f^R\right]-\frac{\alpha^2\beta^2}{(\alpha+\beta)^2}\sigma(\beta g^L+\alpha g^R)$ and

$f_{\beta^{-1}\beta^{-1}}=\frac{\alpha\beta^4}{(\alpha+\beta)^3}\left[\left\{2\frac{\alpha^2}{\beta^2}+(\alpha+\beta)^2\sigma^2+\frac{\alpha+\beta}{\beta}(4\alpha+2\beta)m_{\beta+}+(\alpha+\beta)^2m_{\beta+}^2\right\}f^L+2\frac{\alpha}{\beta}f^R\right]-$

$\frac{\alpha\beta^4}{(\alpha+\beta)^2}\sigma\left\{\frac{4\alpha+2\beta}{\beta}+(\alpha+\beta)m_{\beta+}\right\}g^L.$

In the case in which $\tau=-\log\alpha$ and $\upsilon=-\log\beta$ are used instead of α and β, the respective partial derivatives of f are replaced as follows:

$f_\tau=-\frac{\alpha\beta}{(\alpha+\beta)^2}\left[f^L+\left\{\frac{\beta}{\alpha}-(\alpha+\beta)m_{\alpha-}\right\}f^R\right]+\frac{\alpha\beta}{\alpha+\beta}\sigma g^R,$

$f_\upsilon=-\frac{\alpha\beta}{(\alpha+\beta)^2}\left[\left\{\frac{\alpha}{\beta}+(\alpha+\beta)m_{\beta+}\right\}f^L+f^R\right]+\frac{\alpha\beta}{\alpha+\beta}\sigma g^L,$

$f_{\mu\tau}=\frac{\alpha\beta}{(\alpha+\beta)^2}[\beta f^L-\{\alpha+2\beta-\alpha(\alpha+\beta)m_{\alpha-}\}f^R]-\frac{\alpha\beta}{(\alpha+\beta)^2}\frac{1}{\sigma}\left[g^L-\left\{\frac{\beta}{\alpha}+\alpha(\alpha+\beta)\sigma^2\right\}g^R\right],$

$f_{\mu\upsilon}=\frac{\alpha\beta}{(\alpha+\beta)^2}\left[\{2\alpha+\beta+\beta(\alpha+\beta)m_{\beta+}\}f^L-\alpha f^R\right]-\frac{\alpha\beta}{(\alpha+\beta)^2}\frac{1}{\sigma}\left[\left\{\frac{\alpha}{\beta}+\beta(\alpha+\beta)\sigma^2\right\}g^L-g^R\right],$

$f_{\varsigma\tau}=-\frac{\alpha\beta}{(\alpha+\beta)^2}\sigma^2[\beta^2f^L+\alpha\{2\alpha+3\beta-\alpha(\alpha+\beta)m_{\alpha-}\}f^R]-\frac{\alpha\beta}{(\alpha+\beta)^2}\frac{1}{\sigma}\left[m_{\beta-}g^L-\left\{(\alpha+\beta)(\alpha^2\sigma^2+\right.\right.$

$\left.\left.1)\sigma^2+\frac{\beta}{\alpha}m_{\alpha+}\right\}g^R\right],$

$f_{\varsigma\upsilon}=-\frac{\alpha\beta}{(\alpha+\beta)^2}\sigma^2\left[\beta\{3\alpha+2\beta+\beta(\alpha+\beta)m_{\beta+}\}f^L+\alpha^2f^R\right]+\frac{\alpha\beta}{(\alpha+\beta)^2}\frac{1}{\sigma}\left[\left\{(\alpha+\beta)(\beta^2\sigma^2+1)\sigma^2-\right.\right.$

$\left.\left.\frac{\alpha}{\beta}m_{\beta-}\right\}g^L+m_{\alpha+}g^R\right],$

$f_{\tau\tau}=\frac{\alpha\beta}{(\alpha+\beta)^3}\left[-(\alpha-\beta)f^L+\left\{-\frac{\beta}{\alpha}(\alpha-\beta)+\alpha(\alpha+\beta)^2\sigma^2-(\alpha+\beta)(\alpha+3\beta)m_{\alpha-}+\alpha(\alpha+\right.\right.$

$\left.\left.\beta)^2m_{\alpha-}^2\right\}f^R\right]-\frac{\alpha^2\beta}{(\alpha+\beta)^2}\sigma\left\{\frac{\alpha+3\beta}{\alpha}-(\alpha+\beta)m_{\alpha-}\right\}g^R$

$f_{\tau\upsilon}=\frac{\alpha\beta}{(\alpha+\beta)^3}\left[\{2\alpha+\beta(\alpha+\beta)m_{\beta+}\}f^L+\{2\beta-\alpha(\alpha+\beta)m_{\alpha-}\}f^R\right]-\frac{\alpha\beta}{(\alpha+\beta)^2}\sigma(\beta g^L+\alpha g^R)$ and

$f_{\upsilon\upsilon}=\frac{\alpha\beta}{(\alpha+\beta)^3}\left[\left\{\frac{\alpha}{\beta}(\alpha-\beta)+\beta(\alpha+\beta)^2\sigma^2+(\alpha+\beta)(3\alpha+\beta)m_{\beta+}+\beta(\alpha+\beta)^2m_{\beta+}^2\right\}f^L+(\alpha-\right.$

$\left.\beta)f^R\right]-\frac{\alpha\beta^2}{(\alpha+\beta)^2}\sigma\left\{\frac{3\alpha+\beta}{\beta}+(\alpha+\beta)m_{\beta+}\right\}g^L.$

Appendix 2 Score functions and observed Fisher information matrices for the dPLN-regression models

In model MdPLN III in which μ, $\varsigma = \log\sigma$, α^{-1} and β^{-1} are linearly regressed on $\mathbf{x}^{\mu}, \mathbf{x}^{\sigma}, \mathbf{x}^{\alpha}$ and $\mathbf{x}^{\beta}$, respectively, the score function is expressed as follows, using the elements in D_{dPLN} in Appendix 1:

$$\ell'_{MdPLN} = \sum_i w_i D_{MdPLN}\left(y_i; \mathbf{x}_i^{\mu}, \mathbf{x}_i^{\sigma}, \mathbf{x}_i^{\alpha}, \mathbf{x}_i^{\beta}\right) = \sum_i w_i D_{MdPLN}(y_i; \mathbf{x}_i),$$

$$D_{MdPLN}\left(y_i; \mathbf{x}_i^{\mu}, \mathbf{x}_i^{\sigma}, \mathbf{x}_i^{\alpha}, \mathbf{x}_i^{\beta}\right) = \left[d_{\mu}\{\boldsymbol{\theta}(\mathbf{x}_i)\}\mathbf{x}_i^{\mu}, d_{\varsigma}\{\boldsymbol{\theta}(\mathbf{x}_i)\}\mathbf{x}_i^{\sigma}, d_{\alpha^{-1}}\{\boldsymbol{\theta}(\mathbf{x}_i)\}\mathbf{x}_i^{\alpha}, d_{\beta^{-1}}\{\boldsymbol{\theta}(\mathbf{x}_i)\}\mathbf{x}_i^{\beta}\right]^T.$$

The observed FIM is expressed as follows, using the elements in I_{dPLN} in Appendix 1:

$$FIM_{MdPLN} = \sum_i w_i I_{MdPLN}\left(y_i; \mathbf{x}_i^{\mu}, \mathbf{x}_i^{\sigma}, \mathbf{x}_i^{\alpha}, \mathbf{x}_i^{\beta}\right),$$

$$I_{MdPLN}\left(y_i; \mathbf{x}_i^{\mu}, \mathbf{x}_i^{\sigma}, \mathbf{x}_i^{\alpha}, \mathbf{x}_i^{\beta}\right) =$$

$$-\begin{bmatrix} l_{\mu\mu}\{\boldsymbol{\theta}(\mathbf{x}_i)\}\mathbf{x}_i^{\mu}\mathbf{x}_i^{\mu T} & l_{\mu\varsigma}\{\boldsymbol{\theta}(\mathbf{x}_i)\}\mathbf{x}_i^{\mu}\mathbf{x}_i^{\sigma T} & l_{\mu\alpha^{-1}}\{\boldsymbol{\theta}(\mathbf{x}_i)\}\mathbf{x}_i^{\mu}\mathbf{x}_i^{\alpha T} & l_{\mu\beta^{-1}}\{\boldsymbol{\theta}(\mathbf{x}_i)\}\mathbf{x}_i^{\mu}\mathbf{x}_i^{\beta T} \\ l_{\mu\varsigma}\{\boldsymbol{\theta}(\mathbf{x}_i)\}\mathbf{x}_i^{\mu}\mathbf{x}_i^{\sigma T} & l_{\varsigma\varsigma}\{\boldsymbol{\theta}(\mathbf{x}_i)\}\mathbf{x}_i^{\sigma}\mathbf{x}_i^{\sigma T} & l_{\varsigma\alpha^{-1}}\{\boldsymbol{\theta}(\mathbf{x}_i)\}\mathbf{x}_i^{\sigma}\mathbf{x}_i^{\alpha T} & l_{\varsigma\beta^{-1}}\{\boldsymbol{\theta}(\mathbf{x}_i)\}\mathbf{x}_i^{\sigma}\mathbf{x}_i^{\beta T} \\ l_{\mu\alpha^{-1}}\{\boldsymbol{\theta}(\mathbf{x}_i)\}\mathbf{x}_i^{\mu}\mathbf{x}_i^{\alpha T} & l_{\varsigma\alpha^{-1}}\{\boldsymbol{\theta}(\mathbf{x}_i)\}\mathbf{x}_i^{\sigma}\mathbf{x}_i^{\alpha T} & l_{\alpha^{-1}\alpha^{-1}}\{\boldsymbol{\theta}(\mathbf{x}_i)\}\mathbf{x}_i^{\alpha}\mathbf{x}_i^{\alpha T} & l_{\alpha^{-1}\beta^{-1}}\{\boldsymbol{\theta}(\mathbf{x}_i)\}\mathbf{x}_i^{\alpha}\mathbf{x}_i^{\beta T} \\ l_{\mu\beta^{-1}}\{\boldsymbol{\theta}(\mathbf{x}_i)\}\mathbf{x}_i^{\mu}\mathbf{x}_i^{\beta T} & l_{\varsigma\beta^{-1}}\{\boldsymbol{\theta}(\mathbf{x}_i)\}\mathbf{x}_i^{\sigma}\mathbf{x}_i^{\beta T} & l_{\alpha^{-1}\beta^{-1}}\{\boldsymbol{\theta}(\mathbf{x}_i)\}\mathbf{x}_i^{\alpha}\mathbf{x}_i^{\beta T} & l_{\beta^{-1}\beta^{-1}}\{\boldsymbol{\theta}(\mathbf{x}_i)\}\mathbf{x}_i^{\beta}\mathbf{x}_i^{\beta T} \end{bmatrix}.$$

In the above formulas, $\mathbf{x}_i^{\mu}$, $\mathbf{x}_i^{\sigma}$, $\mathbf{x}_i^{\alpha}$ and $\mathbf{x}_i^{\beta}$ represent the vectors of the respective regressors' values taken by individual i. The intercepts are regarded to be regressors which always take a value of unity. $d_{\mu}\{\theta(\mathbf{x}_i)\}, d_{\varsigma}\{\theta(\mathbf{x}_i)\}, \cdots$ and $l_{\mu\mu}\{\theta(\mathbf{x}_i)\}, l_{\mu\varsigma}\{\theta(\mathbf{x}_i)\}, \cdots$ represent values of the elements in $D_{dPLN}\{y; \boldsymbol{\theta}(\mathbf{x})\}$ and $I_{dPLN}\{y; \boldsymbol{\theta}(\mathbf{x})\}$ when the regressors take the values $\left\{\mathbf{x}_i^{\mu}, \mathbf{x}_i^{\sigma}, \mathbf{x}_i^{\alpha}, \mathbf{x}_i^{\beta}\right\}$.

In MdPLN III′ in which μ, $\varsigma = \log\sigma$, $\tau = -\log\alpha$ and $\upsilon = -\log\beta$ are linearly regressed on $\mathbf{x}^{\mu}, \mathbf{x}^{\sigma}, \mathbf{x}^{\alpha}$ and $\mathbf{x}^{\beta}$, respectively, subscripts α^{-1} and β^{-1} of $d_{\cdot}$s and $l_{\cdot\cdot}$s in the above formulas should be replaced with τ and υ, respectively.

Appendix 3 Procedure for fitting the dPLN-regression models

The procedure is described below, assuming μ, $\varsigma = \log\sigma$, α^{-1} and β^{-1} are linearly regressed on the respective regressors. The procedure is essentially the same in the case in which μ, $\varsigma = \log\sigma$, $\tau = -\log\alpha$ and $\upsilon = -\log\beta$ are linearly regressed.

Step 1 Fitting of the (single) dPLN model

In principle, the initial parameter values $\{\mu_{(0)}; \sigma_{(0)}; \alpha_{(0)}; \beta_{(0)}\}$ are calculated by applying the method of moment estimation (MME) devised by Reed and Jorgensen (2004), and the ML parameters $\{\mu_{(1)}; \sigma_{(1)}; \alpha_{(1)}; \beta_{(1)}\}$ are then found by the Nelder-Mead simplex algorithm. However, it should be noted that the MME sometimes computes a negative $\sigma^2_{(0)}$. As that is the case in the example in this study, an exceptional rule is introduced, such that $\mu_{(0)}, \alpha_{(0)}$ and $\beta_{(0)}$ are computed by the MME and $\sigma_{(0)}$ is set to 0.1. The NR procedure is not applied in this first step because there are cases of convergence failure.

Step 2 Fitting of the standard wage regression model using the least square method

The reg. coeffs. b_0, b_1, b_2 and b_3 in (3) shall be estimated using the least square method.

Step 3 Fitting of MdPLN I - regression of only parameter μ

Step 3.1 Under constraints that ς, α^{-1} and β^{-1} are fixed to $\varsigma_{(1)}, \alpha^{-1}_{(1)}$, and $\beta^{-1}_{(1)}$, respectively, beginning with the initial values $\{b_0 - \alpha^{-1}_{(1)} + \beta^{-1}_{(1)}, b_1, b_2, b_3\}$, the tentative reg. coeffs. of μ, $\left\{c^{\mu}_{0,(3.1)}, c^{\mu}_{1,(3.1)}, c^{\mu}_{2,(3.1)}, c^{\mu}_{3,(3.1)}\right\}$, are determined to maximize the likelihood (i.e., partial maximization) using the Newton-Raphson (NR) algorithm.

Step 3.2 Under constraints that the reg. coeffs. of μ are fixed to $\left\{c^{\mu}_{0,(3.1)}, c^{\mu}_{1,(3.1)}, c^{\mu}_{2,(3.1)}, c^{\mu}_{3,(3.1)}\right\}$, beginning with the initial values $\{\varsigma_{(1)}; \alpha^{-1}_{(1)}; \beta^{-1}_{(1)}\}$, the tentative parameter values $\{\varsigma_{(3.2)}; \alpha^{-1}_{(3.2)}; \beta^{-1}_{(3.2)}\}$ are determined to maximize the likelihood using the NR algorithm with <u>the adjustment multiplier $\rho = 0.8$</u> (see (13)). The adjustment factor is set to less than unity only in this step.

Step 3.3 beginning from the initial values $\left\{c^{\mu}_{0,(3.1)}, c^{\mu}_{1,(3.1)}, c^{\mu}_{2,(3.1)}, c^{\mu}_{3,(3.1)}; \varsigma_{(3.2)}; \alpha^{-1}_{(3.2)}; \beta^{-1}_{(3.2)}\right\}$, the ML parameters/reg. coeffs. of MdPLN I, $\left\{c^{\mu}_{0,(3.3)}, c^{\mu}_{1,(3.3)}, c^{\mu}_{2,(3.3)}, c^{\mu}_{3,(3.3)}; \varsigma_{(3.3)}; \alpha^{-1}_{(3.3)}; \beta^{-1}_{(3.3)}\right\}$, are found using the NR algorithm.

Step 4 Fitting of the model with regression equations of μ and $\log \sigma$

Step 4.1 Under constraints that the reg. coeffs. of μ are fixed to $\left\{c^{\mu}_{0,(3.3)}, c^{\mu}_{1,(3.3)}, c^{\mu}_{2,(3.3)}, c^{\mu}_{3,(3.3)}\right\}$ and parameters α^{-1} and β^{-1} are fixed to $\alpha^{-1}_{(3.3)}$ and $\beta^{-1}_{(3.3)}$, respectively, beginning with the initial values $\{\varsigma_{(3.3)}, 0{,}0\}$, the tentative reg. coeffs. of $\varsigma = \log\sigma$, $\{c^{\sigma}_{0,(4.1)}, c^{\sigma}_{1,(4.1)}, c^{\sigma}_{2,(4.1)}\}$, are determined to maximize the likelihood using the Nelder-Mead simplex algorithm. (Initial values are set to $\{\varsigma_{(3.3)}, 0{,}0{,}0\}$ and the reg. coeffs. are replaced relevantly, hereafter in the case of MdPLN II.)

Step 4.2 Under constraints that the reg. coeffs. of $\log\sigma$ are fixed to $\{c^{\sigma}_{0,(4.1)}, c^{\sigma}_{1,(4.1)}, c^{\sigma}_{2,(4.1)}\}$, beginning from the initial values of the reg. coeffs. $\{c^{\mu}_{0,(3.3)}, c^{\mu}_{1,(3.3)}, c^{\mu}_{2,(3.3)}, c^{\mu}_{3,(3.3)}\}$ and parameter values $\{\alpha^{-1}_{(3.3)}; \beta^{-1}_{(3.3)}\}$, the tentative reg. coeffs. of μ, $\{c^{\mu}_{0,(4.2)}, c^{\mu}_{1,(4.2)}, c^{\mu}_{2,(4.2)}, c^{\mu}_{3,(4.2)}\}$, and parameters $\{\alpha^{-1}_{(4.2)}; \beta^{-1}_{(4.2)}\}$ are determined to maximize the likelihood using the Nelder-Mead simplex algorithm.

Iterations of Steps 4.1 and 4.2 Step 4.1 is performed again after replacing the constraints with the respective reg. coeffs./parameters being fixed to $\{c^{\mu}_{0,(4.2)}, c^{\mu}_{1,(4.2)}, c^{\mu}_{2,(4.2)}, c^{\mu}_{3,(4.2)}\}$, $\alpha^{-1}_{(4.2)}$ and $\beta^{-1}_{(4.2)}$ and replacing the initial values with $\{c^{\sigma}_{0,(4.1)}, c^{\sigma}_{1,(4.1)}, c^{\sigma}_{2,(4.1)}\}$; then, Step 4.2 is again performed after replacing the initial values with $\{c^{\mu}_{0,(4.2)}, c^{\mu}_{1,(4.2)}, c^{\mu}_{2,(4.2)}, c^{\mu}_{3,(4.2)}\}$ and $\{\alpha^{-1}_{(4.2)}; \beta^{-1}_{(4.2)}\}$.

Step 4.3 Beginning from the initial values $\left\{\begin{matrix} c^{\mu}_{0,(4.2)}, c^{\mu}_{1,(4.2)}, c^{\mu}_{2,(4.2)}, c^{\mu}_{3,(4.2)}; \\ c^{\sigma}_{0,(4.1)}, c^{\sigma}_{1,(4.1)}, c^{\sigma}_{2,(4.1)}; \alpha^{-1}_{(4.2)}; \beta^{-1}_{(4.2)} \end{matrix}\right\}$, the ML reg. coeffs./parameters $\{c^{\mu}_{0,(4.3)}, c^{\mu}_{1,(4.3)}, c^{\mu}_{2,(4.3)}, c^{\mu}_{3,(4.3)}; c^{\sigma}_{0,(4.3)}, c^{\sigma}_{1,(4.3)}, c^{\sigma}_{2,(4.3)}; \alpha^{-1}_{(4.3)}; \beta^{-1}_{(4.3)}\}$ are found using the NR algorithm.

Step 5 Fitting of the model with regression equations of $\boldsymbol{\mu}, \log\boldsymbol{\sigma}$ and $\boldsymbol{\beta^{-1}}$

Step 5.1 Under constraints that the reg. coeffs. of μ and $\log\sigma$ are fixed to $\{c^{\mu}_{0,(4.3)}, c^{\mu}_{1,(4.3)}, c^{\mu}_{2,(4.3)}, c^{\mu}_{3,(4.3)}\}$ and $\{c^{\sigma}_{0,(4.3)}, c^{\sigma}_{1,(4.3)}, c^{\sigma}_{2,(4.3)}\}$, respectively and parameter α^{-1} is fixed to $\alpha^{-1}_{(4.3)}$, beginning with the initial values of the reg. coeffs. $\{\beta^{-1}_{(4.3)}, 0{,}0{,}0\}$, the tentative reg. coeffs. of β^{-1}, $\{c^{\beta}_{0,(5.1)}, c^{\beta}_{1,(5.1)}, c^{\beta}_{2,(5.1)}, c^{\beta}_{3,(5.1)}\}$, are determined to maximize the likelihood using the NR algorithm.

Step 5.2 Beginning from the initial values $\{c^{\mu}_{0,(4.3)}, c^{\mu}_{1,(4.3)}, c^{\mu}_{2,(4.3)}, c^{\mu}_{3,(4.3)}; c^{\sigma}_{0,(4.3)}, c^{\sigma}_{1,(4.3)}, c^{\sigma}_{2,(4.3)}; \alpha^{-1}_{(4.3)}; c^{\beta}_{0,(5.1)}, c^{\beta}_{1,(5.1)}, c^{\beta}_{2,(5.1)}, c^{\beta}_{3,(5.1)}\}$, the ML reg.

coeffs./parameter

$\left\{c^{\mu}_{0,(5.2)}, c^{\mu}_{1,(5.2)}, c^{\mu}_{2,(5.2)}, c^{\mu}_{3,(5.2)}; c^{\sigma}_{0,(5.2)}, c^{\sigma}_{1,(5.2)}, c^{\sigma}_{2,(5.2)}; \alpha^{-1}_{(5.2)}; c^{\beta}_{0,(5.2)}, c^{\beta}_{1,(5.2)}, c^{\beta}_{2,(5.2)}, c^{\beta}_{3,(5.2)}\right\}$ are found using the NR algorithm.

Step 6 Fitting of MdPLN III – regression of all four parameters

Step 6.1 Under constraints that the reg. coeffs. of $\mu, \log\sigma$ and β^{-1} are fixed to $\left\{c^{\mu}_{0,(5.2)}, c^{\mu}_{1,(5.2)}, c^{\mu}_{2,(5.2)}, c^{\mu}_{3,(5.2)}\right\}$, $\left\{c^{\sigma}_{0,(5.2)}, c^{\sigma}_{1,(5.2)}, c^{\sigma}_{2,(5.2)}\right\}$ and $\left\{c^{\beta}_{0,(5.2)}, c^{\beta}_{1,(5.2)}, c^{\beta}_{2,(5.2)}, c^{\beta}_{3,(5.2)}\right\}$, respectively, beginning with the initial values $\left\{\alpha^{-1}_{(5.2)}, 0,0,0\right\}$, the tentative reg. coeffs. of α^{-1}, $\left\{c^{\alpha}_{0,(6.1)}, c^{\alpha}_{1,(6.1)}, c^{\alpha}_{2,(6.1)}, c^{\alpha}_{3,(6.1)}\right\}$, are determined to maximize the likelihood using the NR algorithm.

Step 6.2 Under constraints that the reg. coeffs. of α^{-1} are fixed to $\left\{c^{\alpha}_{0,(6.1)}, c^{\alpha}_{1,(6.1)}, c^{\alpha}_{2,(6.1)}, c^{\alpha}_{3,(6.1)}\right\}$, beginning with the initial values $\left\{c^{\mu}_{0,(5.2)}, c^{\mu}_{1,(5.2)}, c^{\mu}_{2,(5.2)}, c^{\mu}_{3,(5.2)}\right\}$, $\left\{c^{\sigma}_{0,(5.2)}, c^{\sigma}_{1,(5.2)}, c^{\sigma}_{2,(5.2)}\right\}$ and $\left\{c^{\beta}_{0,(5.2)}, c^{\beta}_{1,(5.2)}, c^{\beta}_{2,(5.2)}, c^{\beta}_{3,(5.2)}\right\}$, respectively, the tentative reg. coeffs. of $\mu, \log\sigma$ and β^{-1}, $\left\{c^{\mu}_{0,(6.2)}, c^{\mu}_{1,(6.2)}, c^{\mu}_{2,(6.2)}, c^{\mu}_{3,(6.2)}\right\}$, $\left\{c^{\sigma}_{0,(6.2)}, c^{\sigma}_{1,(6.2)}, c^{\sigma}_{2,(6.2)}\right\}$ and $\left\{c^{\beta}_{0,(6.2)}, c^{\beta}_{1,(6.2)}, c^{\beta}_{2,(6.2)}, c^{\beta}_{3,(6.2)}\right\}$, are determined to maximize the likelihood using the NR algorithm.

Iterations of Steps 6.1 and 6.2 Steps 6.1 and 6.2 are iterated by replacing the constraints and initial values relevantly until the time when the tentative reg. coeffs. change only slightly. Both steps were iterated 12 times for the empirical example in this study.

Step 6.3 (Final step) Beginning with the initial values $\left\{\begin{matrix} c^{\mu}_{0,(6.2)}, c^{\mu}_{1,(6.2)}, c^{\mu}_{2,(6.2)}, c^{\mu}_{3,(6.2)}; c^{\sigma}_{0,(6.2)}, c^{\sigma}_{1,(6.2)}, c^{\sigma}_{2,(6.2)}; \\ c^{\alpha}_{0,(6.1)}, c^{\alpha}_{1,(6.1)}, c^{\alpha}_{2,(6.1)}, c^{\alpha}_{3,(6.1)}; c^{\beta}_{0,(6.2)}, c^{\beta}_{1,(6.2)}, c^{\beta}_{2,(6.2)}, c^{\beta}_{3,(6.2)} \end{matrix}\right\}$, the ML reg. coeffs. of MdPLN III are found using the NR algorithm.

References

Clementi, F., Gallegati, M., Kaniadakis, G.: κ-generalized statistics in personal income distribution. Eur. Phys. B **57**(2), 187–193 (2007)

Dagum, C.: A new model of personal income distribution: specification and estimation. Economie

Appliquée **30**(3), 413–437 (1977)

Faiella, I.: Accounting for sampling design in the SHIW. Temi di discussione 662, Bank of Italy (2008)

Kleiber, C., Kotz, S.: Statistical Size Distributions in Economics and Actuarial Sciences. Wiley, Hoboken (2003)

Klein, N., Kneib, T., Lang, S., Sohn, A.: Bayesian structured additive distributional regression with an application to regional income inequality in Germany. Ann. Appl. Stat. **9**(2), 1024–1052 (2015)

Koenker, R.W., Bassett, G.W.: Regression quantiles. Econometrica **46**(1), 33–50 (1978)

McDonald, J.B.: Some generalized functions for the size distribution of income. Econometrica **52**(3), 647–663 (1984)

Mincer, J.: Schooling, Experience and Earnings. Columbia University Press, New York (1974)

Nelder, J.A., Mead, R.: A simplex method for function minimization. Comput. J. **7**(4), 308–313 (1965)

Noufaily, A., Jones, M.C.: Parametric quantile regression based on the generalized gamma distribution. J. Roy. Stat. Soc. C-App., **62**(5), 723–740 (2013)

Okamoto, M.: Decomposition of Gini and multivariate Gini indices. J. Econ. Inequal. **7**(2), 153–177 (2009).

Okamoto, M.: Evaluation of the goodness of fit of new statistical size distributions with consideration of accurate income inequality estimation. Econ. Bull. **32**(4), 2969–2982. Http://www.accessecon.com/Pubs/EB/2012/Volume32/EB-12-V32-I4-P285.pdf (2012)

Okamoto, M.: Erratum to “Evaluation of the goodness of fit of new statistical size distributions with consideration of accurate income inequality estimation”. Econ. Bull. **33**(3), 2443–2444. Http://www.accessecon.com/Pubs/EB/2013/Volume33/EB-13-V33-I3-P229.pdf (2013)

Okamoto, M.: A flexible descriptive model for the size distribution of incomes. Econ. Bull. **34**(3), 1600–1610. Http://www.accessecon.com/Pubs/EB/2014/Volume34/EB-14-V34-I3-P144.pdf (2014)

Reed, W.J.: The Pareto law of incomes – an explanation and an extension. Physica A **319**, 469–486 (2003)

Reed, W.J., Jorgensen, M.: The double Pareto-lognormal distribution – a new parametric model for size distribution. Commun. Stat.-Theor. M. **33**(8), 1733–1753 (2004)

Reed, W.J., Wu, F.: New four- and five-parameter models for income distributions. In: Chotikapanich (ed.): Modeling Income Distributions and Lorenz Curve, pp. 211–223. Springer, New York (2008)

Shorrocks, A.F.: The class of additively decomposable inequality measures. Econometrica **48**(3), 613–626 (1980)

Singh, S.K., Maddala, G.S.: A function for the size distribution of incomes. Econometrica **44**(5), 963–970 (1975)

Sohn, A,, Klein, N., Kneib, T.: A new semiparametric approach to analysing conditional income distributions. SOEP papers on multidisciplinary panel data research 676, DIW Berlin (2014)

Toda, A.A.: The double power law in income distribution: explanations and evidence. J. Econ. Behav. Organ. **84**(1), 364 – 381 (2012)

Yitzhaki, S., Lerman, R.I.: Income stratification and income inequality. Rev. Income Wealth **37**(3), 313–329 (1991)

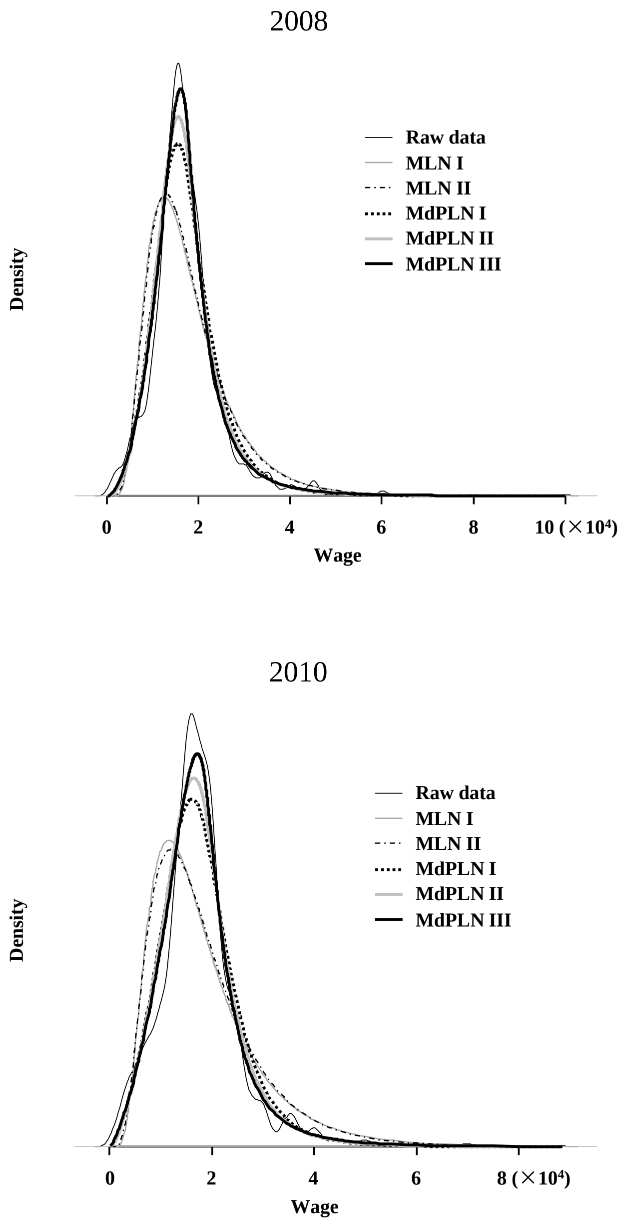

Fig. 1 The pdfs of wage distributions predicted by LN- and dPLN-reg. models fitted to Italian male wages. Note: KDE overestimates the density of the raw data around a null wage

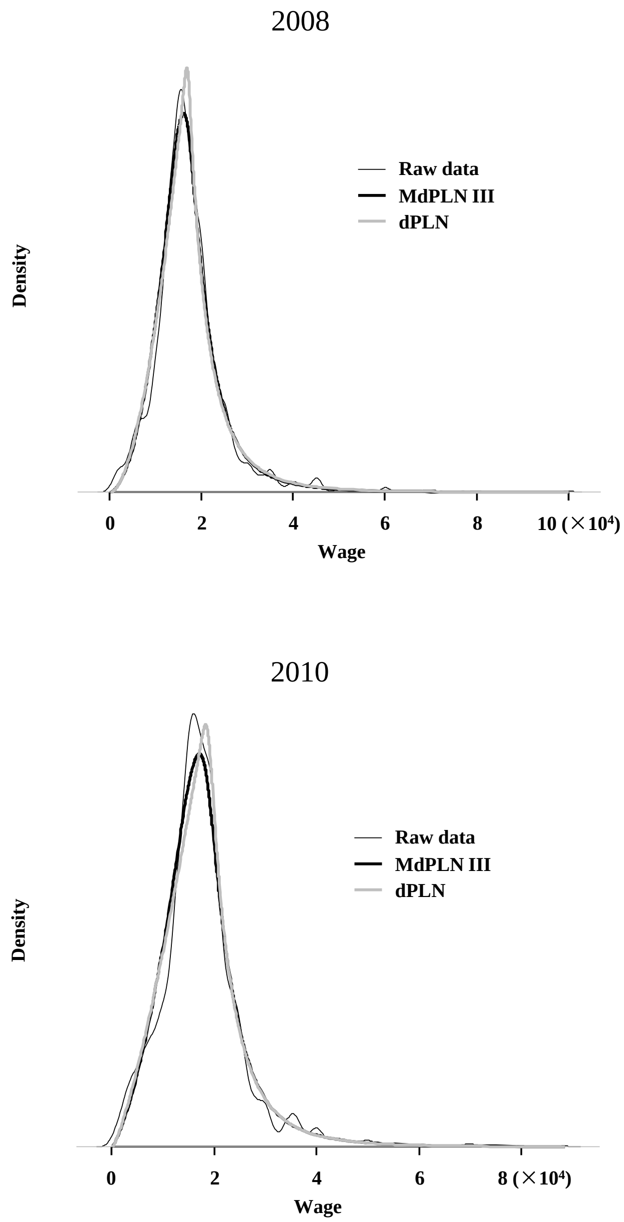

Fig. 2 The pdfs of wage distributions predicted by the dPLN and dPLN-reg. models fitted to Italian male wages. Note: As under Fig. 1

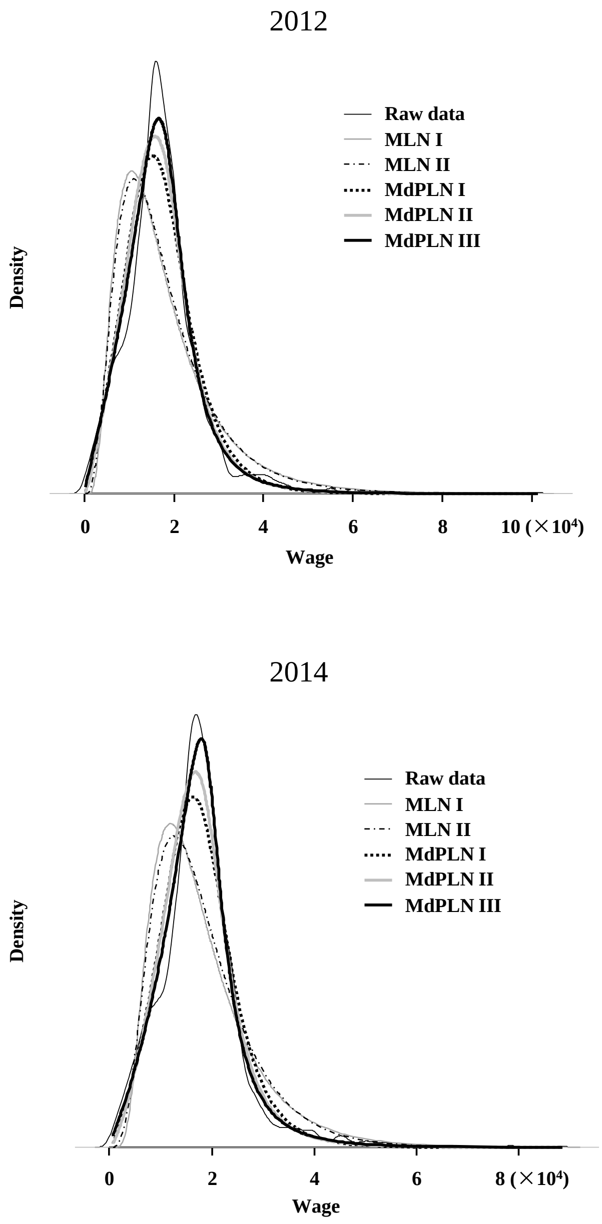

Fig. 1 (cont'd)

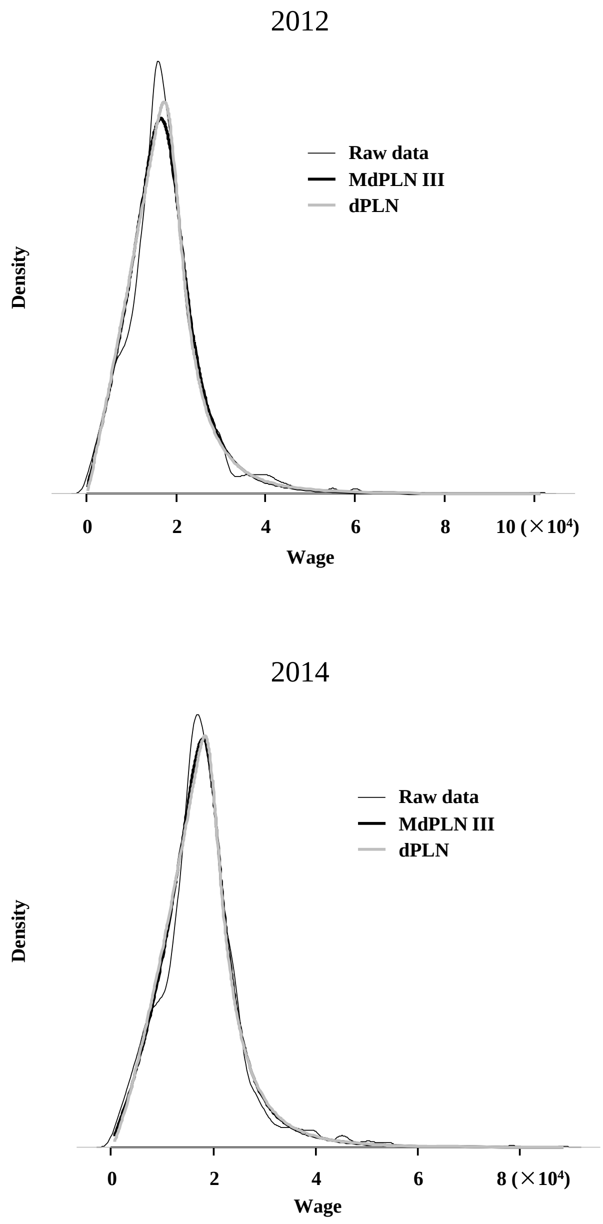

Fig. 2 (cont'd)

Table 1 Estimated equations of LN- and dPLN-regression models fitted to Italian male wages

2008

Para.	Regressor	MLN I	MLN II	MdPLN I	MdPLN II	MdPLN III	MdPLN III′
μ	intercept	8.35***	8.37***	8.65***	8.72***	9.05***	9.08***
	exp	4.53***	4.39***	3.70***	3.59***	2.66***	2.62***
	exp^2	-6.16***	-5.88***	-5.00***	-4.87***	-4.12**	-4.05**
	educ	6.24***	6.20***	5.68***	5.31***	3.46***	3.30***
log σ	intercept	-0.85***	-0.77***	-2.64^	-6.71***	-6.95***	-6.78***
	exp		-1.36		15.59	5.51*	4.92^
	exp^2		1.63		-23.27		
	educ		1.07		18.85***	24.92***	23.88***
1/α or	intercept			0.20***	0.18***	-0.05	-3.11***
-log α	*exp*					0.69^	3.79^
	exp^2					-0.71	-3.50
	educ					1.21**	7.66***
1/β or	intercept			0.32***	0.32***	0.59***	-0.35
-log β	*exp*					-0.91^	-2.97^
	exp^2					0.83	3.10
	educ					-1.17*	-3.40*
μ+1/α	intercept			8.53***	8.58***	8.42***	
−1/β	*exp*					4.26***	
	exp^2					-5.66***	
	educ					5.85***	
ℓ		-28437	-28412	-27929	-27890	-27855	-27851
AIC		56884	56840	55872	55799	55739	55733
BIC		56914	56887	55913	55858	55828	55822
ℓ_M		-28823	-28734	-28461	-28430	-28395	-28392
L-RSSE		1.304	1.235	0.368	0.242	0.122	0.176

Notes: Reg. coeffs. for years of potential work experience (*exp*) and education (*educ*) are multiplied by 10^2 and those for the square of potential work experience years (*exp*^{2}) are multiplied by 10^4. The coeffs. are statistically significant at *** 0.1, ** 1, * 5, and ^ 10% level.

Table 1 (cont'd)

2010

Para.	Regressor	MLN I	MLN II	MdPLN I	MdPLN II	MdPLN III	MdPLN III′
μ	intercept	8.31***	8.33***	8.75***	8.82***	9.16***	9.15***
	exp	4.33***	4.18***	3.40***	3.43***	2.74***	2.66***
	exp^2	-5.62***	-5.30***	-4.36***	-4.53***	-4.25***	-4.03***
	educ	6.52***	6.46***	5.69***	5.30***	3.14***	3.26***
log σ	intercept	-0.68***	-0.36^	-2.41***	-5.44***	-5.81***	-5.51***
	exp		-0.46		5.96	2.76	2.40
	exp^2		-0.32		-6.73		
	educ		-1.86		18.79***	22.65***	21.37***
1/α or	intercept			0.18***	0.16***	-0.04	-2.85***
-log α	*exp*					0.18	0.99
	exp^2					0.40	2.08
	educ					1.34**	6.58***
1/β or	intercept			0.40***	0.39***	0.79***	0.19
-log β	*exp*					-1.09^	-3.36*
	exp^2					1.01	3.73
	educ					-2.06**	-5.97***
μ+1/α	intercept			8.53***	8.58***	8.33***	
−1/β	*exp*					4.00***	
	exp^2					-4.86***	
	educ					6.53***	
ℓ		-26191	-26179	-25599	-25565	-25527	-25527
AIC		52392	52375	51213	51149	51083	51083
BIC		52422	52421	51254	51208	51171	51171
ℓ_M		-26477	-26381	-26022	-25996	-25965	-25961
L-RSSE		2.060	1.873	0.296	0.222	0.343	0.305

Table 1 (cont'd)

2012

Para.	Regressor	MLN		MdPLN			
		I	II	I	II	III	III′
μ	intercept	8.00***	8.00***	8.55***	8.67***	9.19***	9.26***
	exp	5.88***	5.79***	4.24***	4.00***	1.82^	1.42
	exp^2	-8.22***	-8.04***	-5.26***	-4.98***	-1.45	-0.34
	educ	6.94***	7.03***	6.31***	5.71***	3.45***	3.03**
log σ	intercept	-0.63***	-0.16	-2.13***	-4.62***	-4.45***	-3.69***
	exp		-3.97**		1.13	2.38	2.33
	exp^2		7.99^		5.08		
	educ		-0.57		15.94***	14.87***	9.34
α^{-1} or	intercept			0.16***	0.15***	-0.09	-4.52***
-log α	*exp*					0.74	9.70
	exp^2					-0.91	-17.14
	educ					1.15*	11.94*
β^{-1} or	intercept			0.45***	0.44***	1.07***	0.44
-log β	*exp*					-3.15***	-6.52***
	exp^2					5.54*	11.59*
	educ					-2.27*	-4.73^
$\mu+\alpha^{-1}$	intercept			8.26***	8.37***	8.03***	
$-\beta^{-1}$	*exp*					5.71***	
	exp^2					-7.90***	
	educ					6.87***	
ℓ		-24830	-24805	-24319	-24294	-24255	-24251
AIC		49669	49626	48653	48608	48539	48533
BIC		49698	49672	48693	48665	48626	48619
ℓ_M		-25116	-24992	-24757	-24732	-24702	-24703
L-RSSE		2.041	1.827	0.291	0.190	0.135	0.219

Table 1 (cont'd)

2014

Para.	Regressor	MLN I	MLN II	MdPLN I	MdPLN II	MdPLN III	MdPLN III′
μ	intercept	8.19	8.19	8.76	8.87	9.48	9.54
	exp	5.04	4.83	3.31	3.13	1.47	1.08
	exp^2	-6.77	-6.23	-3.90	-3.71	-1.83	-1.28
	educ	6.55	6.64	5.52	4.98	1.72	1.82
log σ	intercept	-0.75	-0.09	-2.40	-5.85	-5.28	-3.36
	exp		-3.72		0.76	2.37	-1.26
	exp^2		6.76		9.98		
	educ		-2.22		21.93	19.12	12.03
$1/\alpha$ or	intercept			5.90	0.15	-0.12	-5.93
$-\log\alpha$	*exp*					0.38	13.82
	exp^2					0.13	-15.47
	educ					1.61	13.51
$1/\beta$ or	intercept			2.43	0.41	1.14	0.74
$-\log\beta$	*exp*					-2.79	-6.32
	exp^2					4.14	9.98
	educ					-3.26	-7.88
$\mu+1/\alpha$	intercept			12.24	8.61	8.21	
$-1/\beta$	*exp*					4.65	
	exp^2					-5.83	
	educ					6.59	
ℓ		-21328	-21294	-20993	-20965	-20905	-20900
AIC		42665	42603	41999	41949	41840	41829
BIC		42693	42648	42039	42006	41925	41914
ℓ_M		-21587	-21488	-21318	-21292	-21264	-21268
L-RSSE		1.536	1.273	0.239	0.159	0.080	0.162

Table 2 Central tendency and dispersion statistics estimated from LN- and dPLN-regression models fitted to Italian male wages

2008

Statistics		Raw data	MLN		MdPLN			
			I	II	I	II	III	III′
Mean (€)	(a)	17,652	17,892	17,876	17,374	17,308	17,443	17,535
Mode (€)	(b)	15,410	12,375	12,667	15,395	15,498	16,015	16,070
Gini	(c)	0.2332	0.2729	0.2708	0.2291	0.2290	0.2345	0.2381
Adj. Gini	(d)	0.1974*	0.2386	0.2378	0.1935	0.1977	0.2006	0.2040
Within-Gini	(e)	0.1930*	0.2386	0.2372	0.1935	0.1962	0.1984	0.2012
(e)/(c) in %		82.7	87.4	87.6	84.4	85.7	84.6	84.5
(e)/(d) in %		97.8	100.0	99.8	100.0	99.2	98.9	98.7
MLD	(f)	0.1082	0.1217	0.1207	0.0923	0.0926	0.0999	0.1039
Within-MLD	(g)	0.0766*	0.0922	0.0921	0.0697	0.0686	0.0703	0.0716
(g)/(f) in %		70.9	75.7	76.3	75.5	74.1	70.3	68.9
T1	(h)	0.1047	0.1219	0.1210	0.0890	0.0932	0.1039	0.1123
Within-T1	(i)	0.0714*	0.0922	0.0920	0.0661	0.0682	0.0731	0.0782
(i)/(h) in %		68.2	75.6	76.0	74.2	73.2	70.4	69.6
SCV	(j)	0.1276	0.1384	0.1382	0.0985	0.1128	0.1434	0.2470
Within-SCV	(k)	0.0849*	0.1075	0.1078	0.0746	0.0859	0.1102	0.2093
(k)/(j) in %		66.5	77.7	78.0	75.8	76.1	76.8	84.7

2010

Statistics		Raw data	MLN		MdPLN			
			I	II	I	II	III	III'
Mean (€)	(a)	17,934	18,569	18,484	17,661	17,652	17,876	17,856
Mode (€)	(b)	16,014	11,472	11,855	16,022	16,367	17,025	17,063
Gini	(c)	0.2409	0.3097	0.3033	0.2429	0.2439	0.2512	0.2500
Adj. Gini	(d)	0.2087*	0.2788	0.2745	0.2099	0.2146	0.2165	0.2153
Within-Gini	(e)	0.2048*	0.2788	0.2741	0.2099	0.2132	0.2141	0.2129
(e)/(c) in %		85.0	90.0	90.4	86.4	87.4	85.2	85.2
(e)/(d) in %		98.1	100.0	99.9	100.0	99.4	98.9	98.9
MLD	(f)	0.1239	0.1587	0.1539	0.1085	0.1095	0.1189	0.1187
Within-MLD	(g)	0.0952*	0.1274	0.1273	0.0862	0.0857	0.0861	0.0859
(g)/(f) in %		76.9	80.3	82.7	79.4	78.3	72.4	72.4
T1	(h)	0.1066	0.1589	0.1504	0.0988	0.1037	0.1172	0.1158
Within-T1	(i)	0.0777*	0.1274	0.1236	0.0762	0.0791	0.0828	0.0817
(i)/(h) in %		72.9	80.1	82.2	77.1	76.3	70.6	70.6
SCV	(j)	0.1160	0.1874	0.1711	0.1043	0.1184	0.1542	0.1514
Within-SCV	(k)	0.0808*	0.1546	0.1433	0.0808	0.0922	0.1166	0.1144
(k)/(j) in %		69.7	82.5	83.7	77.5	77.9	75.6	75.5

Note: Figures denoted by an asterisk (*) were calculated from the raw data cross-classified by four work experience-year groups and five grades of educational qualification

Table 2 (cont'd)

2012

Statistics		Raw data	MLN		MdPLN			
			I	II	I	II	III	III′
Mean (€)	(a)	17,438	18,088	18,002	17,167	17,091	17,236	17,373
Mode (€)	(b)	16,070	10,550	11,025	15,182	15,730	16,540	16,330
Gini	(c)	0.2598	0.3283	0.3211	0.2607	0.2583	0.2607	0.2664
Adj. Gini	(d)	0.2210*	0.2931	0.2876	0.2223	0.2255	0.2224	0.2282
Within-Gini	(e)	0.2172*	0.2931	0.2870	0.2223	0.2247	0.2210	0.2258
(e)/(c) in %		83.6	89.3	89.4	85.3	87.0	84.8	84.7
(e)/(d) in %		98.3	100.0	99.8	100.0	99.7	99.3	98.9
MLD	(f)	0.1432	0.1797	0.1748	0.1275	0.1255	0.1340	0.1398
Within-MLD	(g)	0.1089*	0.1414	0.1415	0.0993	0.0990	0.0996	0.1011
(g)/(f) in %		76.0	78.7	80.9	77.9	78.8	74.4	72.3
T1	(h)	0.1230	0.1795	0.1703	0.1128	0.1134	0.1196	0.1334
Within-T1	(i)	0.0876*	0.1414	0.1367	0.0844	0.0865	0.0846	0.0928
(i)/(h) in %		71.2	78.8	80.3	74.8	76.2	70.8	69.6
SCV	(j)	0.1367	0.2158	0.1983	0.1170	0.1225	0.1346	Inf.
Within-SCV	(k)	0.0920*	0.1763	0.1633	0.0876	0.0943	0.0976	Inf.
(k)/(j) in %		67.3	81.7	82.3	74.9	77.0	72.5	n.a.

2014

Statistics		Raw data	MLN		MdPLN			
			I	II	I	II	III	III'
Mean (€)	(a)	17,809	18,230	18,123	17,598	17,511	17,700	17,799
Mode (€)	(b)	16,949	11,844	12,580	16,330	16,740	17,988	17,750
Gini	(c)	0.2404	0.2943	0.2847	0.2397	0.2381	0.2409	0.2449
Adj. Gini	(d)	0.2080*	0.2607	0.2529	0.2096	0.2127	0.2093	0.2133
Within-Gini	(e)	0.2037*	0.2607	0.2521	0.2096	0.2116	0.2071	0.2102
(e)/(c) in %		84.7	88.6	88.5	87.4	88.9	86.0	85.8
(e)/(d) in %		97.9	100.0	99.7	100.0	99.5	98.9	98.6
MLD	(f)	0.1194	0.1428	0.1368	0.1075	0.1065	0.1160	0.1195
Within-MLD	(g)	0.0902*	0.1107	0.1103	0.0873	0.0868	0.0876	0.0878
(g)/(f) in %		75.6	77.5	80.6	81.3	81.5	75.5	73.5
T1	(h)	0.1047	0.1426	0.1315	0.0959	0.0977	0.1053	0.1157
Within-T1	(i)	0.0745*	0.1107	0.1049	0.0757	0.0777	0.0762	0.0822
(i)/(h) in %		71.1	77.6	79.7	78.9	79.4	72.3	71.0
SCV	(j)	0.1118	0.1648	0.1448	0.0989	0.1063	0.1222	Inf.
Within-SCV	(k)	0.0747*	0.1320	0.1173	0.0781	0.0853	0.0914	Inf.
(k)/(j) in %		66.8	80.1	81.0	79.0	80.2	74.8	n.a.

Table 3 Estimation results of parametric size distribution models fitted to Italian male wages

2008

	Raw	4 para.		3 para.				2 para.
	data	dPLN	GB2	SM	Dagum	κG	GG	LN
Mean (€)	17,652	17,564	17,539	17,519	17,233	17,552	17,650	17,890
Mode (€)	15,410	16,708	16,493	14,635	15,395	15,248	13,710	12,375
Gini	0.2332	0.2377	0.2365	0.2312	0.2259	0.2314	0.2525	0.2726
MLD	0.1082	0.1030	0.1017	0.0906	0.0898	0.0924	0.1080	0.1215
Theil	0.1047	0.1069	0.1050	0.0901	0.0860	0.0933	0.1019	0.1215
SCV	0.1276	0.1424	0.1378	0.1019	0.0937	0.1125	0.1068	0.1375
ℓ		-28398	-28397	-28513	-28473	-28464	-28704	-28821
AIC		56804	56802	57033	56952	56934	57413	57645
BIC		56828	56826	57051	56970	56951	57431	57657
L-RSSE		0.161	0.154	0.329	0.435	0.238	0.845	1.297

2010

	Raw	4 para.		3 para.				2 para.
	data	dPLN	GB2	SM	Dagum	κG	GG	LN
Mean (€)	17,934	17,798	17,802	17,848	17,557	17,865	17,918	18,563
Mode (€)	16,014	18,314	18,382	15,040	16,353	15,626	14,295	11,472
Gini	0.2409	0.2465	0.2465	0.2461	0.2397	0.2444	0.2630	0.3094
MLD	0.1239	0.1163	0.1163	0.1063	0.1068	0.1059	0.1231	0.1584
Theil	0.1066	0.1101	0.1101	0.0993	0.0962	0.1009	0.1108	0.1584
SCV	0.1160	0.1316	0.1316	0.1051	0.1002	0.1117	0.1126	0.1864
ℓ		-25960	-25960	-26093	-26026	-26053	-26192	-26466
AIC		51928	51929	52191	52057	52112	52391	52936
BIC		51952	51952	52209	52075	52129	52408	52948
L-RSSE		0.200	n.a.	0.389	0.305	0.301	0.830	2.050

Notes: ‘κG’ stands for the κ-generalized distribution (Clementi *et al*. 2007). ‘GG’ stands for the generalized Gamma distribution. For the year 2010, the L-RSSE of the GB2 is unable to be properly computed due to the limitations of the computational precision of the statistical analysis system used for this study

Table 3 (cont'd)

2012

	Raw	4 para.		3 para.				2 para.
	data	dPLN	GB2	Singh-Maddala	Dagum	κG	GG	LN
Mean (€)	17,438	17,331	17,321	17,350	17,093	17,359	17,427	18,090
Mode (€)	16,070	17,308	17,495	14,175	15,730	14,775	13,100	10,490
Gini	0.2598	0.2658	0.2659	0.2650	0.2582	0.2633	0.2813	0.3286
MLD	0.1432	0.1367	0.1369	0.1252	0.1267	0.1248	0.1424	0.1799
T1	0.1230	0.1273	0.1274	0.1149	0.1118	0.1163	0.1270	0.1799
SCV	0.1367	0.1542	0.1544	0.1216	0.1169	0.1280	0.1301	0.2165
ℓ		-24707	-24707	-24806	-24745	-24774	-24890	-25117
AIC		49421	49422	49618	49496	49555	49786	50238
BIC		49444	49445	49636	49513	49572	49804	50250
L-RSSE		0.201	0.203	0.370	0.268	0.269	0.786	2.050

2014

	Raw	4 para.		3 para.				2 para.
	data	dPLN	GB2	SM	Dagum	κG	GG	LN
Mean (€)	17,809	17,704	17,703	17,754	17,491	17,765	17,803	18,238
Mode (€)	16,949	18,205	18,353	15,050	16,575	15,667	14,284	11,844
Gini	0.2404	0.2423	0.2423	0.2443	0.2371	0.2419	0.2590	0.2947
MLD	0.1194	0.1136	0.1136	0.1051	0.1055	0.1042	0.1188	0.1430
Theil	0.1047	0.1048	0.1048	0.0975	0.0940	0.0981	0.1074	0.1430
SCV	0.1118	0.1195	0.1194	0.1023	0.0966	0.1065	0.1091	0.1656
ℓ		-21270	-21270	-21369	-21315	-21339	-21436	-21591
AIC		42548	42548	42744	42637	42684	42879	43187
BIC		42570	42571	42761	42654	42700	42896	43198
L-RSSE		0.113	0.114	0.330	0.242	0.254	0.663	1.547

Table 4 Least square and quantile regression results for conditional wage distributions predicted by LN- and dPLN-regression models fitted to Italian male wages

2008, 2010 and 2012 on average

Model	Regressor	LS	Quantile regression			Diff. b/w quantile points		Diff. from LS		
			10%	50%	90%	(c)-(b)	(d)-(c)	(b)-(a)	(c)-(a)	(d)-(a)
		(a)	(b)	(c)	(d)					
Raw data	intercept	8.22***	7.34***	8.60***	8.70***	***		***	***	***
	exp	4.92***	7.20***	3.85***	3.93***	*			**	*
	exp^2	-6.67***	-10.33**	-5.30***	-4.58***				^	^
	educ	6.56***	7.22***	5.27***	7.21***	**	***		***	
MLN I	intercept	8.22***	7.59***	8.22***	8.85***	***	***	***		***
	exp	4.92***	4.92***	4.92***	4.92***					
	exp^2	-6.67***	-6.67***	-6.67***	-6.67***					
	educ	6.56***	6.56***	6.56***	6.56***					
MLN II	intercept	8.23***	7.41***	8.23***	9.06***	**	**	**		**
	exp	4.79***	6.13***	4.79***	3.45***	^	^	^		^
	exp^2	-6.41***	-8.63***	-6.41***	-4.18***					
	educ	6.56***	6.88***	6.56***	6.24***					
MdPLN I	intercept	8.44***	7.89***	8.52***	8.89***	***	***			
	exp	3.78***	3.78***	3.78***	3.78***					
	exp^2	-4.87***	-4.87***	-4.87***	-4.87***					
	educ	5.90***	5.90***	5.90***	5.90***					
MdPLN II	intercept	8.51***	8.13***	8.64***	8.72***	***	*	***	***	***
	exp	3.67***	3.34***	3.58***	4.10***	**	*	**	^	*
	exp^2	-4.79***	-4.45***	-4.71***	-5.19***	^				
	educ	5.44***	4.37***	5.12***	6.91***	***	***	***	***	***
MdPLN III	intercept	8.26***	7.49***	8.57***	8.69***	***	**	***	***	***
	exp	4.66***	6.22***	3.91***	3.93***	***		**	***	*
	exp^2	-6.14***	-8.72***	-5.22***	-4.62***	*		*	*	*
	educ	6.42***	6.65***	5.27***	7.21***	**	***		***	**
MdPLN III′	intercept	8.24***	7.47***	8.57***	8.68***	***	**	***	***	***
	exp	4.80***	6.47***	3.94***	3.97***	***		***	***	**
	exp^2	-6.43***	-9.28***	-5.27***	-4.72***	**		*	**	*
	educ	6.50***	6.68***	5.29***	7.26***	**	***		***	*

Notes: As under Table 1

Table 4 (cont'd)

2008

Model	Regressor	LS	Quantile regression			Diff. b/w quantile points		Diff. from LS		
			10%	50%	90%	(c)-	(d)-	(b)-	(c)-	(d)-
		(a)	(b)	(c)	(d)	(b)	(c)	(a)	(a)	(a)
Raw data	intercept	8.35***	7.88***	8.66***	8.66***	**		^	***	*
	exp	4.53***	6.02**	3.49***	4.61***				*	
	exp^2	-6.16***	-8.79*	-4.79***	-6.02*					
	educ	6.24***	5.19***	5.25***	6.98***		*		*	
MLN I	intercept	8.35***	7.80***	8.35***	8.90***	***	***	***		***
	exp	4.53***	4.53***	4.53***	4.53***					
	exp^2	-6.16***	-6.16***	-6.16***	-6.16***					
	educ	6.24***	6.24***	6.24***	6.24***					
MLN II	intercept	8.37***	7.77***	8.37***	8.97***	***	***	***		***
	exp	4.39***	5.23***	4.39***	3.56***					
	exp^2	-5.88***	-6.97**	-5.88***	-4.78**					
	educ	6.20***	5.60***	6.20***	6.80***					
MdPLN I	intercept	8.53***	8.05***	8.58***	8.94***					
	exp	3.70***	3.70***	3.70***	3.70***					
	exp^2	-5.00***	-5.00***	-5.00***	-5.00***					
	educ	5.68***	5.68***	5.68***	5.68***					
MdPLN II	intercept	8.58***	8.31***	8.68***	8.73***	***		***	***	*
	exp	3.59***	3.03***	3.45***	4.28***	*	*	*	*	**
	exp^2	-4.87***	-4.11***	-4.67***	-5.82***	^		^		^
	educ	5.31***	4.11***	5.08***	6.73***	**	***	**	***	***
MdPLN III	intercept	8.42***	7.92***	8.66***	8.67***	***		***	***	***
	exp	4.26***	4.70***	3.71***	4.39***		^		*	
	exp^2	-5.66***	-6.01**	-5.18***	-5.73***					
	educ	5.85***	5.45***	4.94***	7.04***		***		***	**
MdPLN III′	intercept	8.40***	7.89***	8.66***	8.66***	***		***	***	***
	exp	4.38***	4.99***	3.74***	4.32***				*	
	exp^2	-5.88***	-6.68**	-5.26***	-5.57***					
	educ	5.97***	5.48***	4.94***	7.23***		***		***	**

Table 4 (cont'd)

2010

Model	Regressor	LS	Quantile regression			Diff. b/w quantile points		Diff. from LS		
			10%	50%	90%	(c)-	(d)-	(b)-	(c)-	(d)-
		(a)	(b)	(c)	(d)	(b)	(c)	(a)	(a)	(a)
Raw data	intercept	8.31***	7.32***	8.73***	8.74***	***		**	***	***
	exp	4.33***	6.52*	3.12***	3.39***				*	^
	exp^2	-5.62***	-8.32	-4.01***	-3.80***					
	educ	6.52***	7.62***	5.00***	7.59***	**	***		***	*
MLN I	intercept	8.31***	7.66***	8.31***	8.95***	***	***	***		***
	exp	4.33***	4.33***	4.33***	4.33***					
	exp^2	-5.62***	-5.62***	-5.62***	-5.62***					
	educ	6.52***	6.52***	6.52***	6.52***					
MLN II	intercept	8.33***	7.47***	8.33***	9.18***	***	***	***		***
	exp	4.18***	4.49***	4.18***	3.88***					
	exp^2	-5.30***	-5.10^	-5.30***	-5.51**					
	educ	6.46***	7.64***	6.46***	5.27***					
MdPLN I	intercept	8.53***	7.97***	8.61***	8.98***	***	***			
	exp	3.40***	3.40***	3.40***	3.40***					
	exp^2	-4.36***	-4.36***	-4.36***	-4.36***					
	educ	5.69***	5.69***	5.69***	5.69***					
MdPLN II	intercept	8.58***	8.19***	8.72***	8.78***	***		***	***	***
	exp	3.43***	3.13***	3.33***	3.82***	^		^		
	exp^2	-4.53***	-4.22***	-4.42***	-4.94***					
	educ	5.30***	4.14***	4.95***	6.88***	**	***	***	***	***
MdPLN III	intercept	8.33***	7.62***	8.63***	8.73***	***	^	***	***	***
	exp	4.00***	4.93***	3.56***	3.52***	^			*	
	exp^2	-4.86***	-5.98*	-4.64***	-3.96**					
	educ	6.53***	6.68***	5.27***	7.45***	^	***		***	*
MdPLN III′	intercept	8.31***	7.56***	8.62***	8.75***	***	*	***	***	***
	exp	4.18***	5.41***	3.64***	3.49***	*		^	*	
	exp^2	-5.23***	-6.99**	-4.80***	-3.95**					
	educ	6.58***	6.82***	5.31***	7.32***	^	***		***	

Table 4 (cont'd)

2012

Model	Regressor	LS	Quantile regression			Diff. b/w quantile points		Diff. from LS		
			10%	50%	90%	(c)-	(d)-	(b)-	(c)-	(d)-
		(a)	(b)	(c)	(d)	(b)	(c)	(a)	(a)	(a)
Raw data	intercept	8.00***	6.81***	8.40***	8.69***	***		***	*	***
	exp	5.88***	9.05**	4.92***	3.79***					*
	exp^2	-8.22***	-13.87^	-7.10***	-3.93*					*
	educ	6.94***	8.85***	5.57***	7.06***	*		^		
MLN I	intercept	8.00***	7.32***	8.00***	8.68***	***	***	***		***
	exp	5.88***	5.88***	5.88***	5.88***					
	exp^2	-8.22***	-8.22***	-8.22***	-8.22***					
	educ	6.94***	6.94***	6.94***	6.94***					
MLN II	intercept	8.00***	6.98***	8.00***	9.02***	***	***	***		***
	exp	5.79***	8.67***	5.79***	2.92***	**	**	**		**
	exp^2	-8.04***	-13.82**	-8.04***	-2.26	*	*	*		*
	educ	7.03***	7.40***	7.03***	6.65***					
MdPLN I	intercept	8.26***	7.64***	8.36***	8.75***	*	***			
	exp	4.24***	4.24***	4.24***	4.24***					
	exp^2	-5.26***	-5.26***	-5.26***	-5.26***					
	educ	6.31***	6.31***	6.31***	6.31***					
MdPLN II	intercept	8.37***	7.87***	8.52***	8.64***	***	*	***	***	***
	exp	4.00***	3.87***	3.96***	4.18***					
	exp^2	-4.98***	-5.01***	-5.04***	-4.82***					
	educ	5.71***	4.87***	5.33***	7.11***	*	***	***	***	***
MdPLN III	intercept	8.03***	6.93***	8.43***	8.67***	***	***	***	***	***
	exp	5.71***	9.04***	4.46***	3.86***	**		**	***	**
	exp^2	-7.90***	-14.19***	-5.84***	-4.17***	*	^	*	*	*
	educ	6.87***	7.84***	5.59***	7.14***	*	***		***	
MdPLN III′	intercept	8.00***	6.95***	8.42***	8.64***	***	**	***	***	***
	exp	5.85***	9.00***	4.44***	4.10***	**		**	***	*
	exp^2	-8.17***	-14.16***	-5.75***	-4.64***	*		*	*	*
	educ	6.95***	7.75***	5.61***	7.23***	^	***		**	

Table 4 (cont'd)

2014

Model	Regressor	LS	Quantile regression			Diff. b/w quantile points		Diff. from LS		
			10%	50%	90%	(c)-	(d)-	(b)-	(c)-	(d)-
		(a)	(b)	(c)	(d)	(b)	(c)	(a)	(a)	(a)
Raw data	intercept	8.19	6.90	8.71	8.85					
	exp	5.04	9.57	3.92	3.49					
	exp^2	-6.77	-14.94	-5.54	-4.11					
	educ	6.55	8.14	4.32	6.50					
MLN I	intercept	8.19	7.58	8.19	8.79					
	exp	5.04	5.04	5.04	5.04					
	exp^2	-6.77	-6.77	-6.77	-6.77					
	educ	6.55	6.55	6.55	6.55					
MLN II	intercept	8.19	7.18	8.19	9.20					
	exp	4.83	7.19	4.83	2.47					
	exp^2	-6.23	-10.55	-6.23	-1.92					
	educ	6.64	7.91	6.64	5.37					
MdPLN I	intercept	8.52	7.95	8.61	8.97					
	exp	3.31	3.31	3.31	3.31					
	exp^2	-3.90	-3.90	-3.90	-3.90					
	educ	5.52	5.52	5.52	5.52					
MdPLN II	intercept	8.61	8.19	8.76	8.83					
	exp	3.13	2.94	3.10	3.34					
	exp^2	-3.71	-3.69	-3.79	-3.56					
	educ	4.98	3.97	4.61	6.45					
MdPLN III	intercept	8.21	7.12	8.64	8.87					
	exp	4.65	7.47	3.61	2.98					
	exp^2	-5.83	-10.46	-4.49	-2.86					
	educ	6.59	8.06	4.95	6.43					
MdPLN III′	intercept	8.20	7.16	8.61	8.88					
	exp	4.87	7.76	3.87	2.91					
	exp^2	-6.37	-11.49	-4.99	-2.68					
	educ	6.57	7.79	4.93	6.50					

Table 5 Effects of longer work experience and higher education on the mean and dispersion of wages estimated from LN- and dPLN-regression models fitted to Italian male wages

2008, 2010 and 2012 on average

Para.	Statistics	Potential work experience years				Education years			
		MLN	MdPL			MLN	MdPL		
		II	II	III	III′	II	II	III	III′
Total	Mean	1.69***	1.56***	1.81***	1.86***	6.42***	5.93***	6.73***	6.94***
	Gini	-0.63*	0.12	-0.09	0.01	-0.35	2.47***	2.33***	3.01**
	MLD	-1.56**	0.25	-0.78	-0.67	-0.75	4.56***	2.91^	4.10^
	T1	-1.28*	0.70^	0.48	0.94	-0.75	6.63***	7.32**	10.05*
	SCV	-1.25^	1.67^	2.56	n.a.	-0.88	11.55**	16.81^	n.a.
μ	Mean	1.81***	1.46***	0.89***	0.94***	6.56***	5.44***	3.35***	3.20***
	Gini	-0.28***	-0.37***	-0.26**	-0.22*	0.00	0.00	0.00	0.00
	MLD	-0.72***	-0.66***	-0.50**	-0.43*	0.00	0.00	0.00	0.00
	T1	-0.53***	-0.61***	-0.42**	-0.35*	0.00	0.00	0.00	0.00
	SCV	-0.43**	-0.58***	-0.36**	n.a.	0.00	0.00	0.00	n.a.
σ	Mean	-0.12	0.10***	0.08^	0.06^	-0.14	0.49***	0.50**	0.38*
	Gini	-0.36	0.49***	0.40*	0.31*	-0.35	2.47***	2.43**	1.86**
	MLD	-0.84	0.91***	0.72^	0.54^	-0.75	4.56***	4.35**	3.28*
	T1	-0.75	1.31**	1.11^	0.80	-0.75	6.63***	6.68*	4.93*
	SCV	-0.82	2.25*	2.23	n.a.	-0.88	11.55**	13.26	n.a.
α & β	Mean			0.84***	0.85***			2.88***	3.36***
	Gini			-0.22	-0.07			-0.10	1.15
	MLD			-1.00*	-0.78			-1.44	0.82
	T1			-0.21	0.49			0.64	5.12
	SCV			0.68	n.a.			3.55^	n.a.
α	Mean			0.43***	0.48***			1.52***	2.13**
	Gini			0.38**	0.61**			1.50***	2.94*
	MLD			0.64**	1.11*			2.55***	5.28*
	T1			0.89**	1.72*			3.48***	8.26^
	SCV			1.49**	n.a.			5.70**	n.a.
β	Mean			0.41***	0.37***			1.36***	1.22***
	Gini			-0.61***	-0.68***			-1.60***	-1.79***
	MLD			-1.64***	-1.89***			-3.98***	-4.45***
	T1			-1.10***	-1.23***			-2.84***	-3.15***
	SCV			-0.80***	n.a.			-2.15***	n.a.

Notes: Estimates are presented in percent. See the footnote under Table 1 for the statistical significance levels

Table 5 (cont'd)

2008

Para.	Statistics	Potential work experience years: MLN II	Potential work experience years: MdPLN II	MdPLN III	MdPLN III′	Education years: MLN II	Education years: MdPLN II	MdPLN III	MdPLN III′
Total	Mean	1.62***	1.51***	1.77***	1.84***	6.40***	5.81***	6.38***	6.64***
	Gini	-0.76**	0.19	0.04	0.26	0.79	2.77*	3.12*	4.04*
	MLD	-1.79**	0.45	-0.30	0.12	1.64	5.46*	5.27^	7.08^
	Theil	-1.51**	1.05	1.12	1.90	1.64	7.96^	9.68^	12.62^
	SCV	-1.46*	2.48	4.13	22.45	1.86	14.71	21.62	95.05
μ	Mean	1.75***	1.40***	0.79***	0.78***	6.20***	5.31***	3.46***	3.30***
	Gini	-0.29***	-0.43***	-0.39**	-0.38**	0.00	0.00	0.00	0.00
	MLD	-0.76***	-0.78***	-0.76*	-0.72*	0.00	0.00	0.00	0.00
	Theil	-0.52***	-0.72***	-0.65*	-0.60*	0.00	0.00	0.00	0.00
	SCV	-0.35*	-0.70***	-0.59*	-0.54	0.00	0.00	0.00	0.00
σ	Mean	-0.12	0.11^	0.10	0.07	0.20	0.51*	0.47	0.34
	Gini	-0.47	0.62*	0.53	0.36	0.79	2.77*	2.39^	1.73^
	MLD	-1.02	1.23^	1.03	0.68	1.64	5.46*	4.65	3.30^
	Theil	-0.99	1.77^	1.60	1.05	1.64	7.96^	7.24	5.08
	SCV	-1.11	3.17	3.53	3.75	1.86	14.71	15.98	18.23
α & β	Mean			0.88***	0.99***			2.45***	2.99***
	Gini			-0.09	0.29			0.73	2.32
	MLD			-0.57	0.16			0.61	3.78
	Theil			0.16	1.45			2.44	7.53
	SCV			1.18	19.24			5.64	76.82
α	Mean			0.47***	0.62**			1.54**	2.22**
	Gini			0.47*	0.90*			1.84**	3.44*
	MLD			0.83*	1.68*			3.27**	6.45*
	Theil			1.13*	2.47*			4.31*	9.38*
	SCV			1.87^	19.86			6.96*	77.93
β	Mean			0.41***	0.37***			0.91*	0.78*
	Gini			-0.56***	-0.61***			-1.11**	-1.12*
	MLD			-1.40***	-1.52***			-2.66*	-2.67*
	Theil			-0.97***	-1.02**			-1.87**	-1.84*
	SCV			-0.68**	-0.62			-1.32*	-1.11

Table 5 (cont'd)

2010

Para.	Statistics	Potential work experience years				Education years			
		MLN	MdPLN			MLN	MdPLN		
		II	II	III	III′	II	II	III	III′
Total	Mean	1.55***	1.40***	1.73***	1.74***	6.00***	5.84***	6.98***	6.95***
	Gini	-0.68*	0.12	-0.02	-0.06	-1.42	2.69***	2.84*	2.43*
	MLD	-1.51*	0.25	-0.57	-0.80	-2.99	4.91**	3.71	2.51
	T1	-1.39*	0.65	0.68	0.62	-2.98	7.24**	9.18^	8.27*
	SCV	-1.50^	1.52	2.81	3.07^	-3.43	12.62*	21.10	20.51*
μ	Mean	1.70***	1.32***	0.75***	0.77***	6.46***	5.30***	3.14***	3.26***
	Gini	-0.21***	-0.30***	-0.29***	-0.27***	0.00	0.00	0.00	0.00
	MLD	-0.53***	-0.51***	-0.54***	-0.51***	0.00	0.00	0.00	0.00
	T1	-0.42**	-0.46***	-0.42**	-0.40**	0.00	0.00	0.00	0.00
	SCV	-0.37*	-0.40***	-0.31*	-0.30*	0.00	0.00	0.00	0.00
σ	Mean	-0.15	0.08**	0.08	0.07	-0.46	0.54***	0.65	0.60^
	Gini	-0.47	0.42**	0.37	0.32	-1.42	2.69***	3.03^	2.85^
	MLD	-0.98	0.76*	0.67	0.57	-2.99	4.91**	5.48	5.04^
	T1	-0.98	1.11*	1.05	0.88	-2.98	7.24**	8.60	7.79
	SCV	-1.13	1.91*	2.10	1.70	-3.43	12.62*	17.23	15.17
α & β	Mean			0.91***	0.89***			3.19***	3.08***
	Gini			-0.11	-0.11			-0.20	-0.42
	MLD			-0.70	-0.86			-1.77	-2.54
	T1			0.05	0.15			0.58	0.48
	SCV			1.01	1.67			3.87	5.34
α	Mean			0.45*	0.47^			1.66**	1.55**
	Gini			0.47^	0.60^			1.60*	1.88*
	MLD			0.79^	1.02^			2.69*	3.23*
	T1			1.06^	1.45			3.68*	4.63^
	SCV			1.71^	2.66			6.09^	8.63
β	Mean			0.46**	0.42***			1.53**	1.54***
	Gini			-0.58***	-0.71***			-1.80***	-2.31***
	MLD			-1.49***	-1.88***			-4.46***	-5.77***
	T1			-1.01***	-1.29***			-3.10***	-4.15***
	SCV			-0.70***	-0.99***			-2.21***	-3.29**

Table 5 (cont'd)

2012

Para.	Statistics	Potential work experience years				Education years			
		MLN	MdPL			MLN	MdPL		
		II	II	III	III′	II	II	III	III′
Total	Mean	1.90***	1.75***	1.93***	1.99***	6.87***	6.13***	6.84***	7.23***
	Gini	-0.46	0.05	-0.29	-0.17	-0.42	1.96***	1.05	2.55
	MLD	-1.38	0.06	-1.47^	-1.33	-0.90	3.32***	-0.24	2.72
	T1	-0.94	0.39	-0.35	0.31	-0.90	4.69***	3.10^	9.27
	SCV	-0.80	1.00	0.74	n.a.	-1.06	7.32**	7.72*	n.a.
μ	Mean	2.00***	1.66***	1.14***	1.26***	7.03***	5.71***	3.45***	3.03**
	Gini	-0.33**	-0.39***	-0.11	-0.03	0.00	0.00	0.00	0.00
	MLD	-0.86***	-0.68***	-0.21	-0.05	0.00	0.00	0.00	0.00
	T1	-0.65*	-0.66***	-0.19	-0.04	0.00	0.00	0.00	0.00
	SCV	-0.57	-0.65***	-0.17	n.a.	0.00	0.00	0.00	n.a.
σ	Mean	-0.09	0.09**	0.06	0.05	-0.16	0.42***	0.39*	0.21
	Gini	-0.13	0.44**	0.30	0.25	-0.42	1.96***	1.88*	1.00
	MLD	-0.52	0.74**	0.47	0.37	-0.90	3.32***	2.91*	1.50
	T1	-0.29	1.05**	0.67	0.48	-0.90	4.69***	4.20*	1.93
	SCV	-0.22	1.65*	1.05	n.a.	-1.06	7.32**	6.58^	n.a.
α & β	Mean			0.73*	0.68^			3.00***	3.99*
	Gini			-0.48	-0.39			-0.83	1.56
	MLD			-1.73^	-1.66			-3.15	1.22
	T1			-0.84	-0.14			-1.10	7.34
	SCV			-0.14	n.a.			1.14	n.a.
α	Mean			0.36^	0.36			1.37^	2.64
	Gini			0.20	0.34			1.07^	3.49
	MLD			0.30	0.61			1.68	6.15
	T1			0.49	1.23			2.45	10.78
	SCV			0.88	n.a.			4.05	n.a.
β	Mean			0.36	0.33			1.63*	1.35^
	Gini			-0.67**	-0.73*			-1.90**	-1.93
	MLD			-2.02**	-2.27*			-4.83**	-4.93
	T1			-1.33**	-1.36*			-3.55**	-3.45
	SCV			-1.03*	n.a.			-2.91**	n.a.

Table 5 (cont'd)

2014

Para.	Statistics	Potential work experience years MLN II	MdPLN II	MdPLN III	MdPLN III′	Education years MLN II	MdPLN II	MdPLN III	MdPLN III′
Total	Mean	1.78	1.54	1.83	2.07	6.18	5.49	6.47	6.92
	Gini	-0.69	0.41	-0.29	0.56	-1.62	2.75	1.08	2.56
	MLD	-1.80	0.74	-1.63	0.01	-3.40	4.83	-1.03	1.77
	Theil	-1.44	1.32	-0.15	3.42	-3.38	7.08	3.95	10.72
	SCV	-1.37	2.54	1.60	n.a.	-3.84	11.85	11.35	n.a.
μ	Mean	1.92	1.41	0.62	0.48	6.64	4.98	1.72	1.82
	Gini	-0.30	-0.28	-0.14	-0.11	0.00	0.00	0.00	0.00
	MLD	-0.78	-0.47	-0.26	-0.19	0.00	0.00	0.00	0.00
	Theil	-0.62	-0.46	-0.22	-0.18	0.00	0.00	0.00	0.00
	SCV	-0.56	-0.44	-0.19	n.a.	0.00	0.00	0.00	n.a.
σ	Mean	-0.14	0.13	0.05	-0.03	-0.46	0.51	0.38	0.25
	Gini	-0.39	0.69	0.26	-0.15	-1.62	2.75	2.13	1.42
	MLD	-1.02	1.21	0.41	-0.22	-3.40	4.83	3.27	2.13
	Theil	-0.82	1.77	0.60	-0.27	-3.38	7.08	4.85	2.60
	SCV	-0.81	2.98	0.98	n.a.	-3.84	11.85	7.93	n.a.
α & β	Mean			1.16	1.62			4.37	4.85
	Gini			-0.41	0.81			-1.04	1.15
	MLD			-1.78	0.43			-4.30	-0.36
	Theil			-0.53	3.87			-0.90	8.12
	SCV			0.80	n.a.			3.42	n.a.
α	Mean			0.54	1.17			1.94	2.80
	Gini			0.53	1.79			1.87	4.51
	MLD			0.83	3.19			2.91	7.99
	Theil			1.21	5.67			4.29	14.01
	SCV			2.04	n.a.			7.28	n.a.
β	Mean			0.62	0.45			2.42	2.05
	Gini			-0.94	-0.98			-2.91	-3.36
	MLD			-2.60	-2.77			-7.21	-8.35
	Theil			-1.74	-1.80			-5.19	-5.89
	SCV			-1.24	n.a.			-3.86	n.a.