

Appendix A: Changing the order of the decomposition

In the decomposition presented in table 2 and figure 1 we have first estimated the effect of residuals, then the effect of coefficients and finally the effect of characteristics. This order is somewhat arbitrary and we could imagine 5 other orders that are not necessarily less sensible. If different orders give totally different results, this would render the results questionable. Therefore, we have estimated all possible decompositions. The effects of residuals, coefficients and characteristics are plotted in figures A1, A2 and A3 respectively. The capital letters represents the orders of the decomposition as defined by the followings formulas:

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

$$B. \left(\hat{q}(\hat{\beta}^{89}, x^{89}) - \hat{q}(\hat{\beta}^{m73, r89}, x^{89}) \right) + \left(\hat{q}(\hat{\beta}^{m73, r89}, x^{89}) - \hat{q}(\hat{\beta}^{73}, x^{89}) \right) + \left(\hat{q}(\hat{\beta}^{73}, x^{89}) - \hat{q}(\hat{\beta}^{73}, x^{73}) \right)$$

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

$$D. \left(\hat{q}(\hat{\beta}^{89}, x^{89}) - \hat{q}(\hat{\beta}^{89}, x^{73}) \right) + \left(\hat{q}(\hat{\beta}^{89}, x^{73}) - \hat{q}(\hat{\beta}^{m73, r89}, x^{73}) \right) + \left(\hat{q}(\hat{\beta}^{m73, r89}, x^{73}) - \hat{q}(\hat{\beta}^{73}, x^{73}) \right)$$

E.

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

F.

$$\left(\hat{q}(\hat{\beta}^{89}, x^{89}) - \hat{q}(\hat{\beta}^{m73, r89}, x^{89}) \right) + \left(\hat{q}(\hat{\beta}^{m73, r89}, x^{89}) - \hat{q}(\hat{\beta}^{m73, r89}, x^{73}) \right) + \left(\hat{q}(\hat{\beta}^{m73, r89}, x^{73}) - \hat{q}(\hat{\beta}^{73}, x^{73}) \right)$$

The results presented in the paper are those obtain by the order A. We observe that the conclusions of the paper are not sensitive to the choice of the order of the decomposition.

Figure A1: Effects of residuals found by changing the order of the decomposition

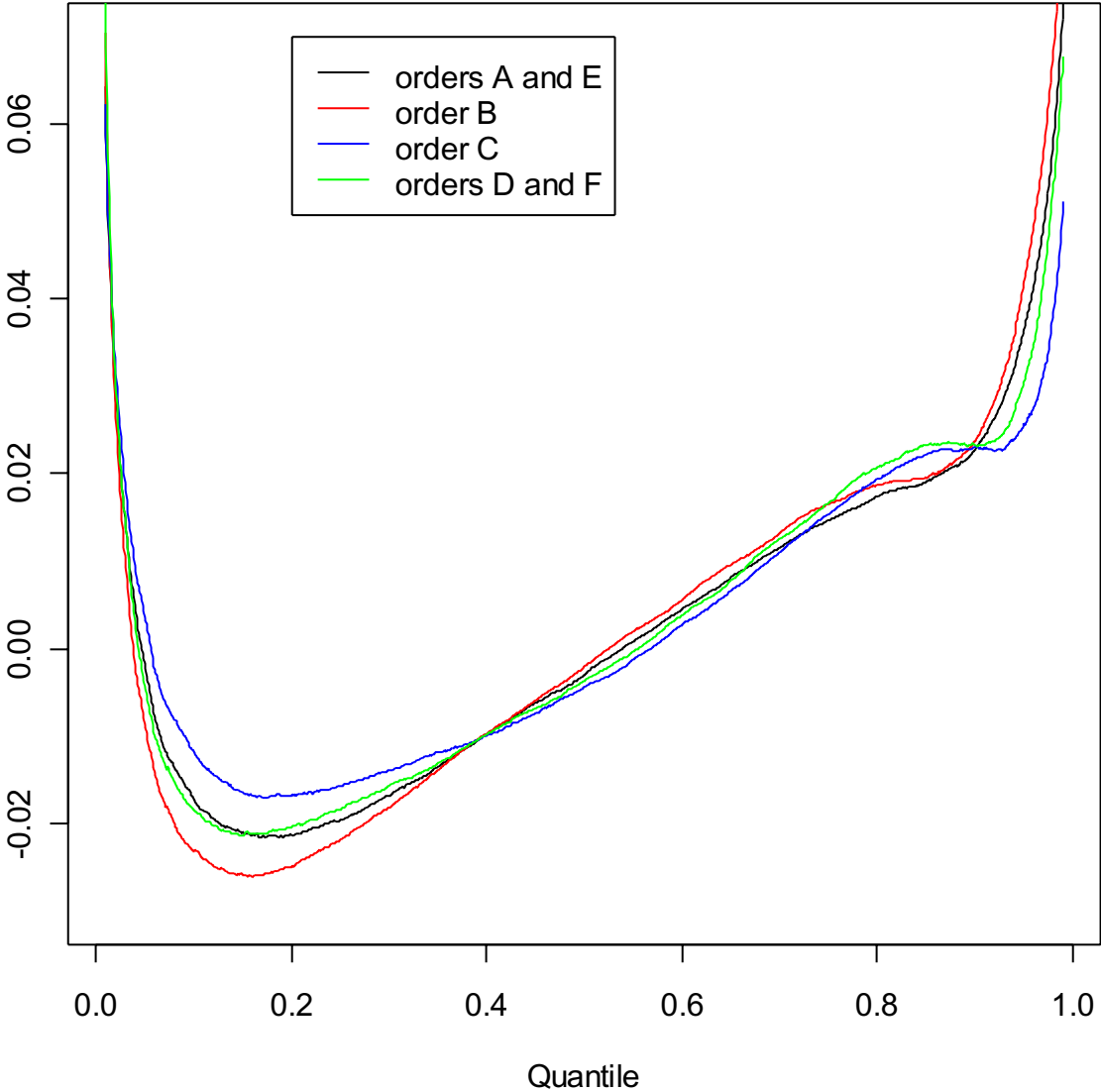


Figure A2: Effects of coefficients found by changing the order of the decomposition

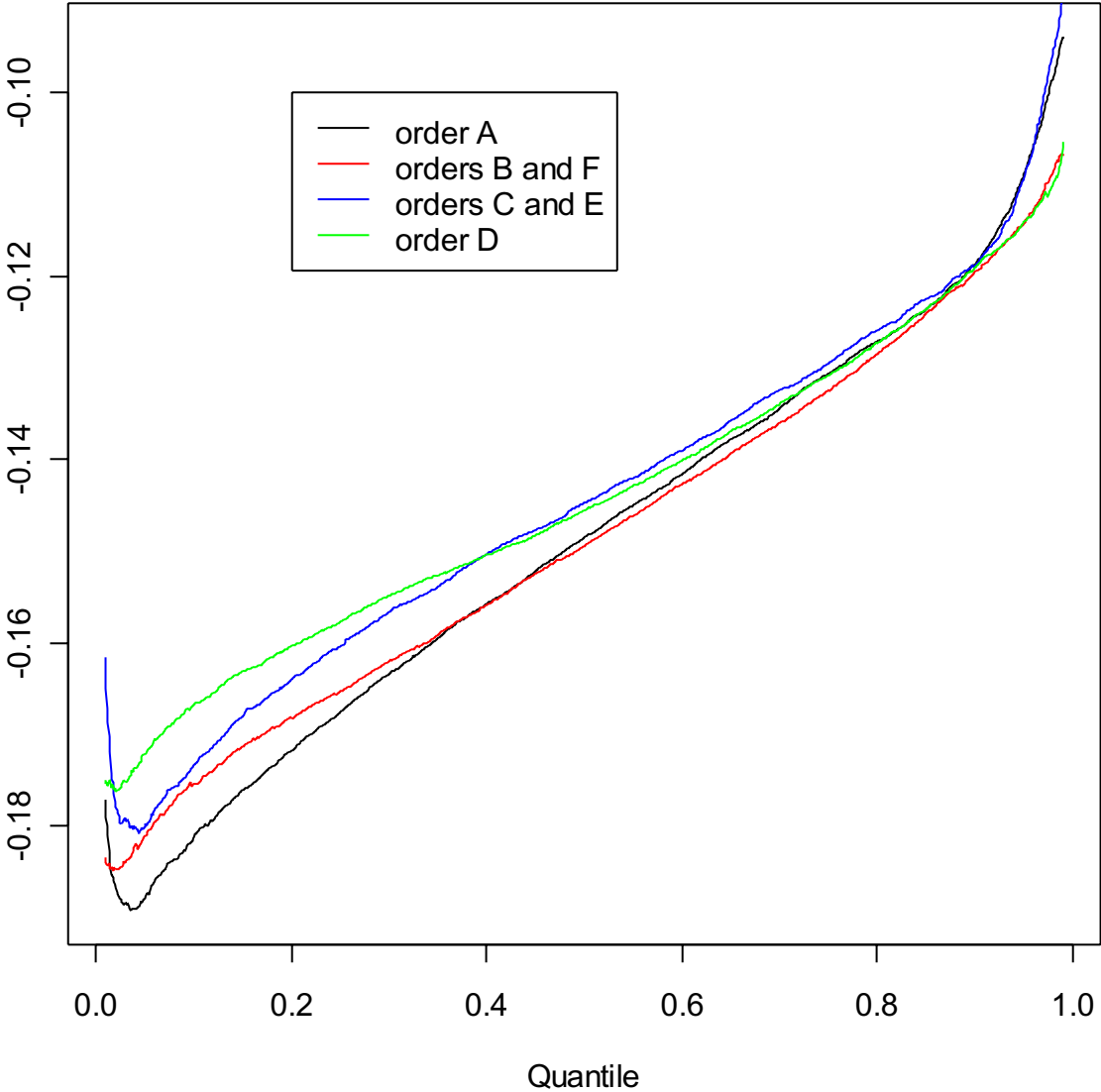
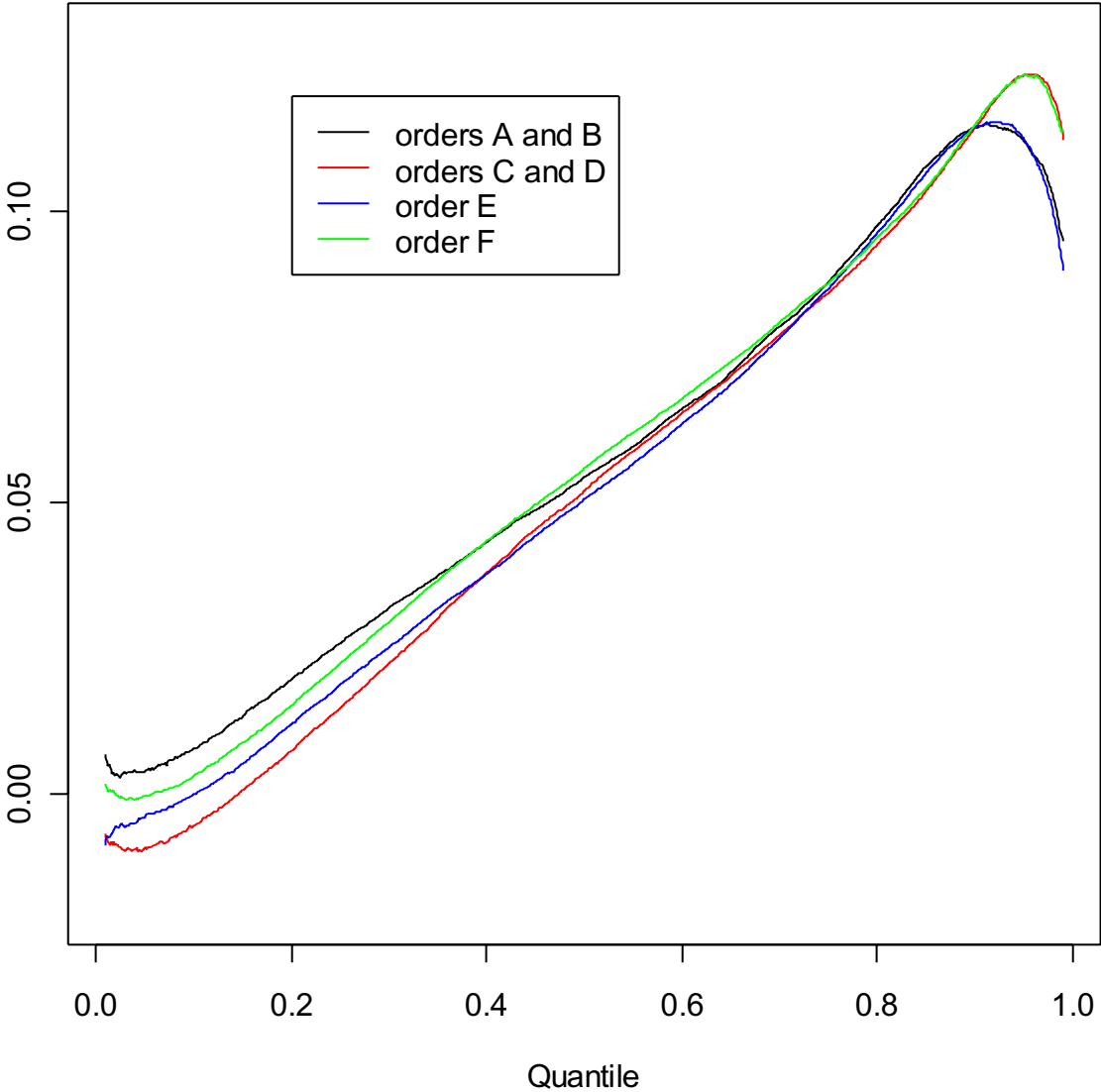


Figure A3: Effects of characteristics found by changing the order of the decomposition



Appendix B: Changing the number of quantile regression

The results presented in the paper are based on the estimation of 200 quantile regressions in the first step. The asymptotic results of Melly (2004) are valid only if the number of quantile regressions goes to infinity as the number of observations goes to infinity. Thus, the highest possible number of quantile regressions should be run in the first step. On the other hand, Gossling, Machin and Meghir (2000) propose a similar procedure but estimate only 9 quantile regressions. A possible concern is that the model could be overparametrized with too many quantile regressions. Therefore, as a robustness test, we compare the decomposition results with different numbers of quantile regressions: 10, 100, 200 and 400. Figures B1, B2 and B3 plot the effects of residuals, coefficients and characteristics based on different numbers of quantile regressions.

The results are very similar and we can almost not distinguish the results with 100, 200 or 400 different quantile regressions. The results with 10 quantile regressions are more noisy but there are not fundamentally different. It follows from these results that too many quantile regressions are not a problem. There is no overparametrization since each quantile regression estimates a different quantity.

Figure B1: Effects of residuals with different numbers of quantile regressions

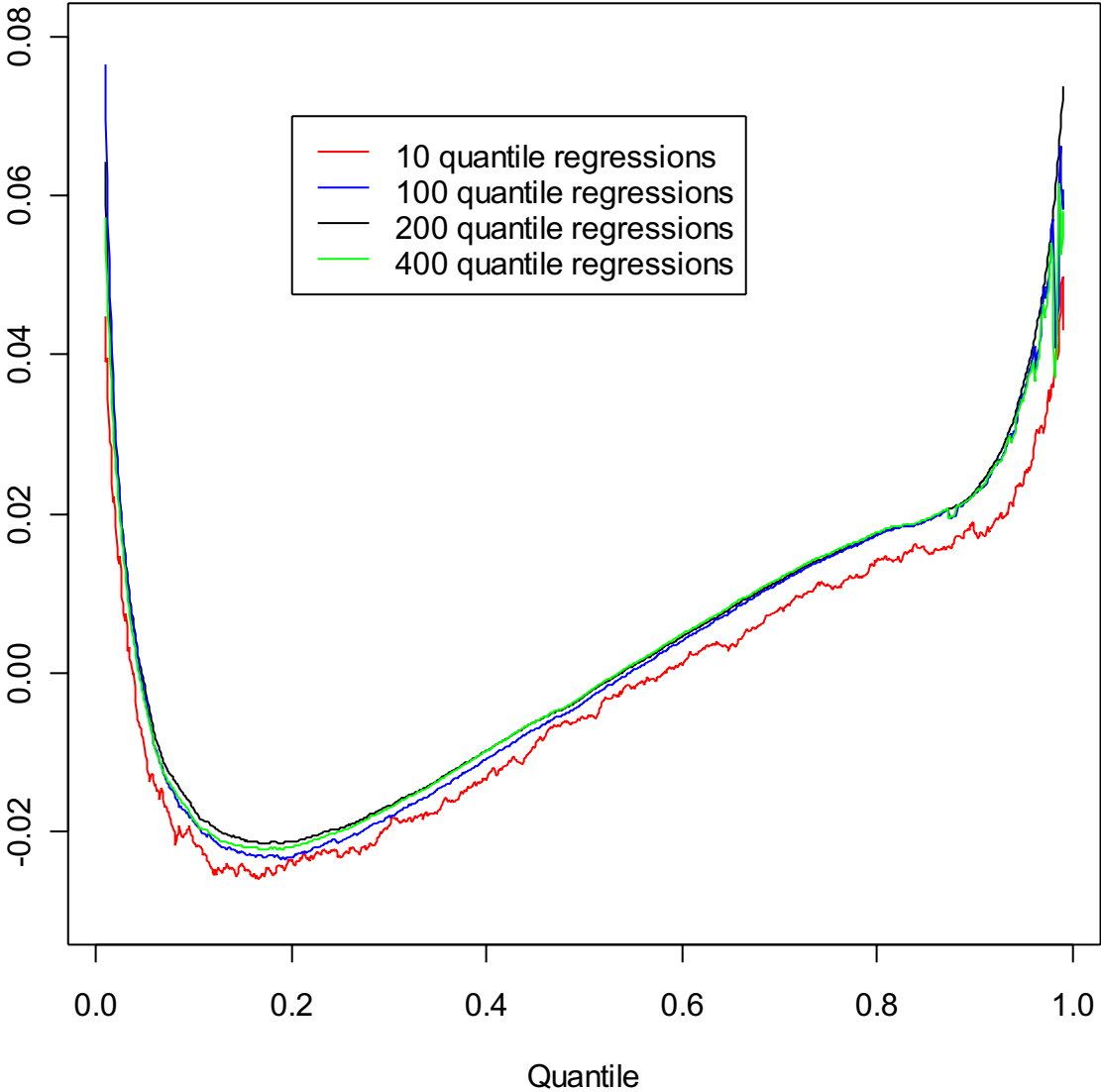


Figure B2: Effects of coefficients with different numbers of quantile regressions

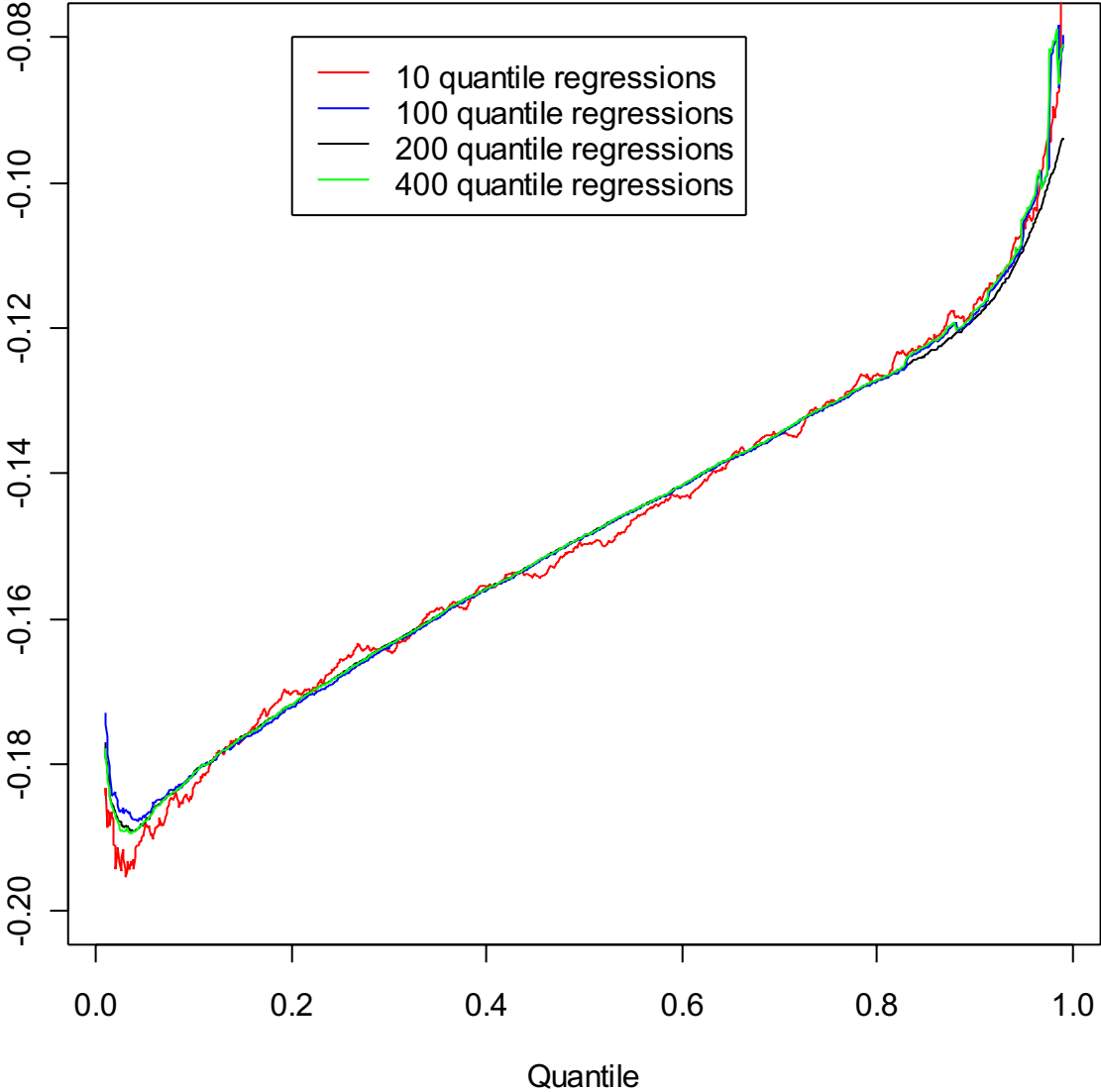
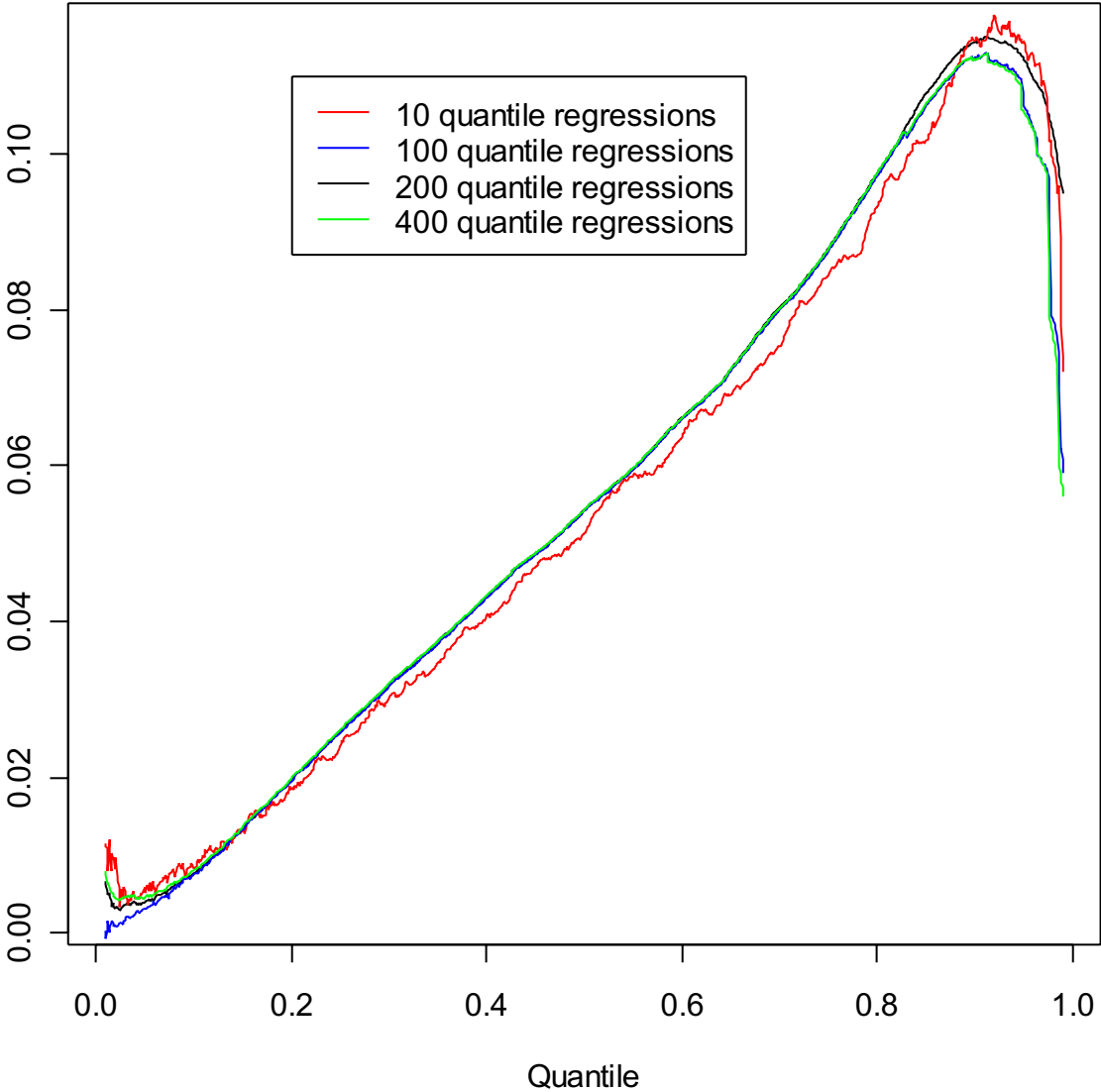


Figure B3: Effects of characteristics with different numbers of quantile regressions



Appendix C: Changes in the distribution of wages over sub-periods

Some insights can be gained by considering smaller periods than the 16 years long period from 1973 to 1989. In this appendix, we present results for the three sub-periods 1973-1979, 1979-1984 and 1984-1989. Exactly the same procedure was used for each period as for the results presented in the paper. Figures C1, C2 and C3 plot the results.

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

Figure C1: Decomposition of changes in distribution between 1973 and 1979

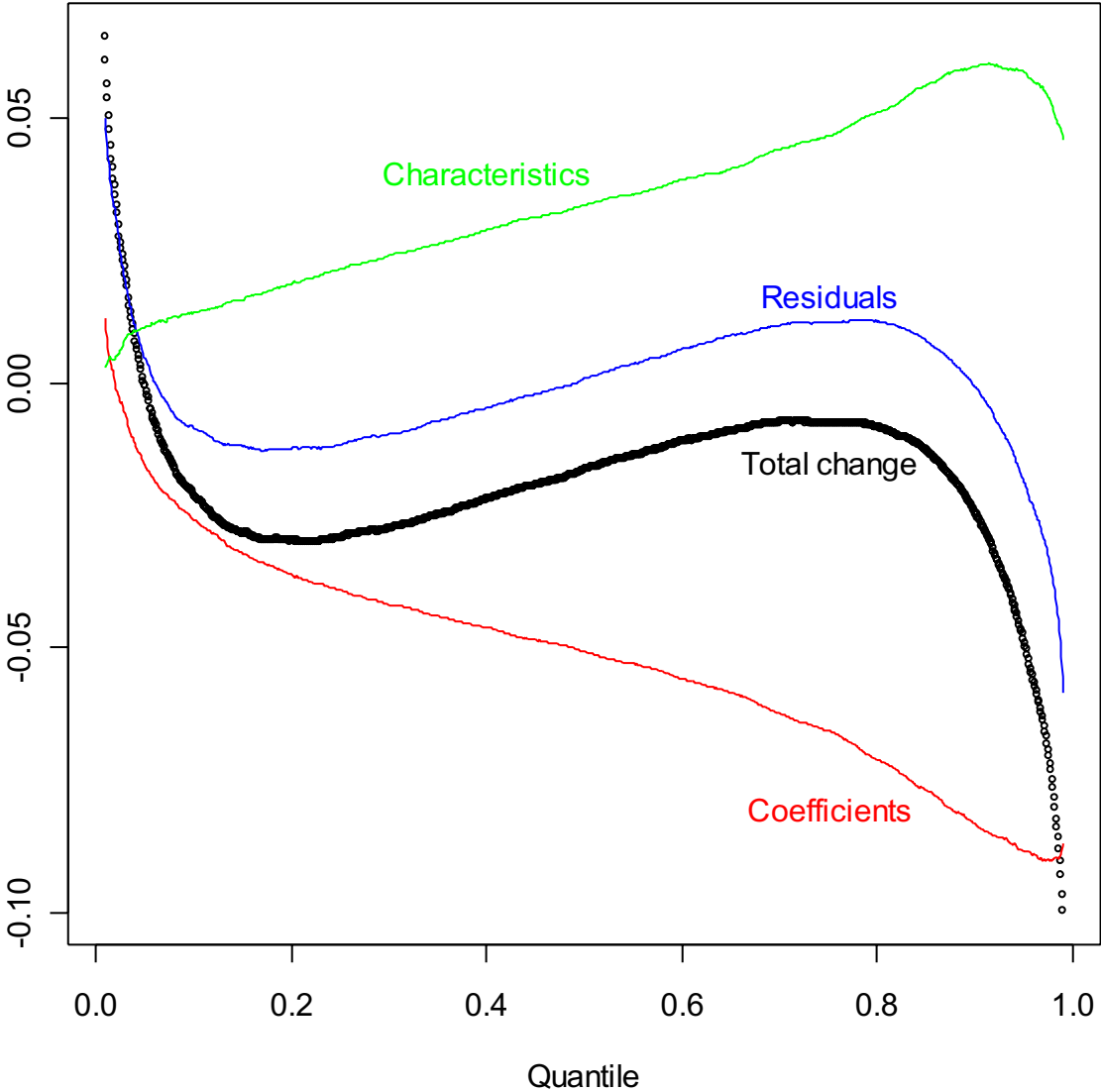


Figure C2: Decomposition of changes in distribution between 1979 and 1984

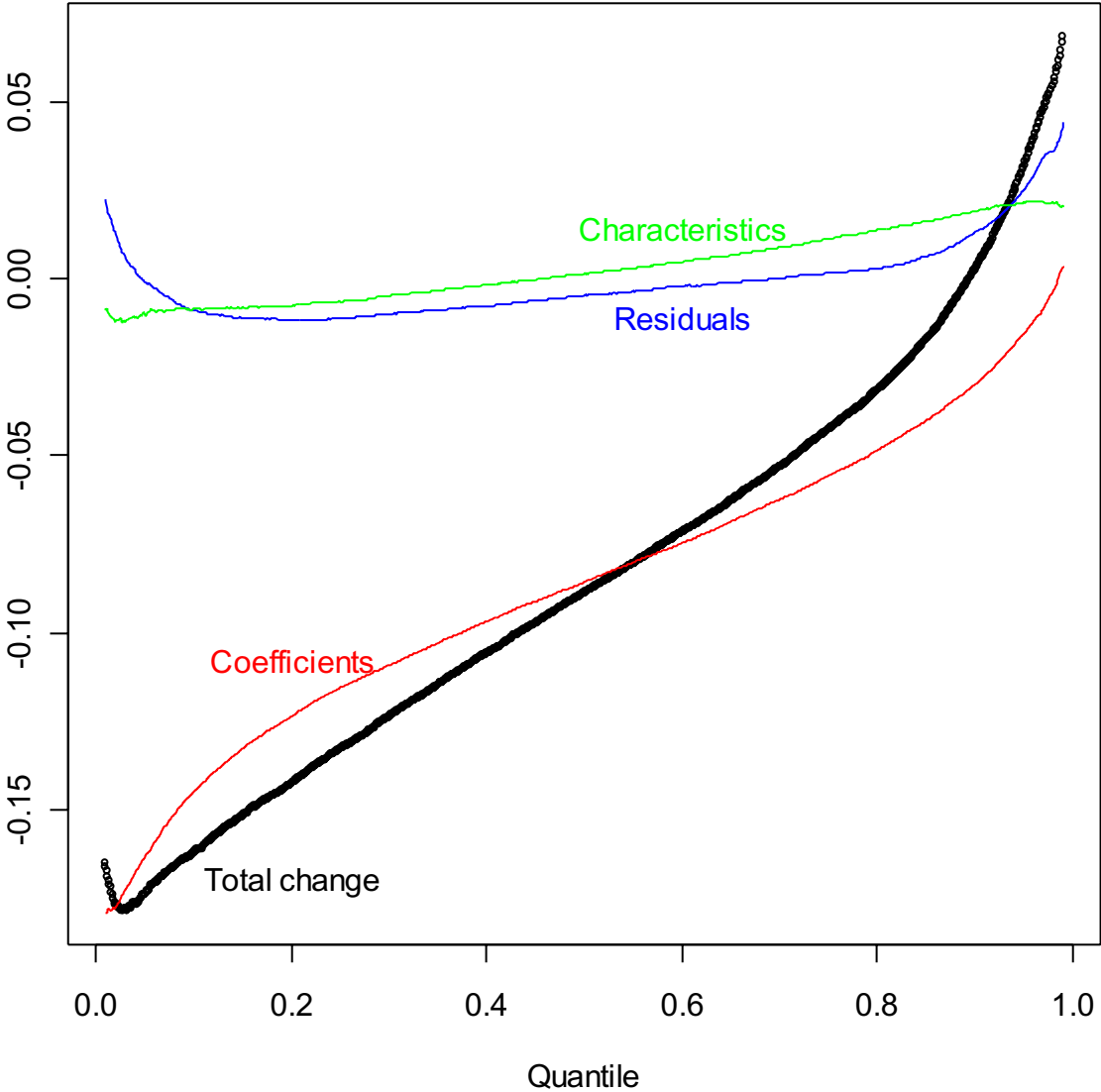
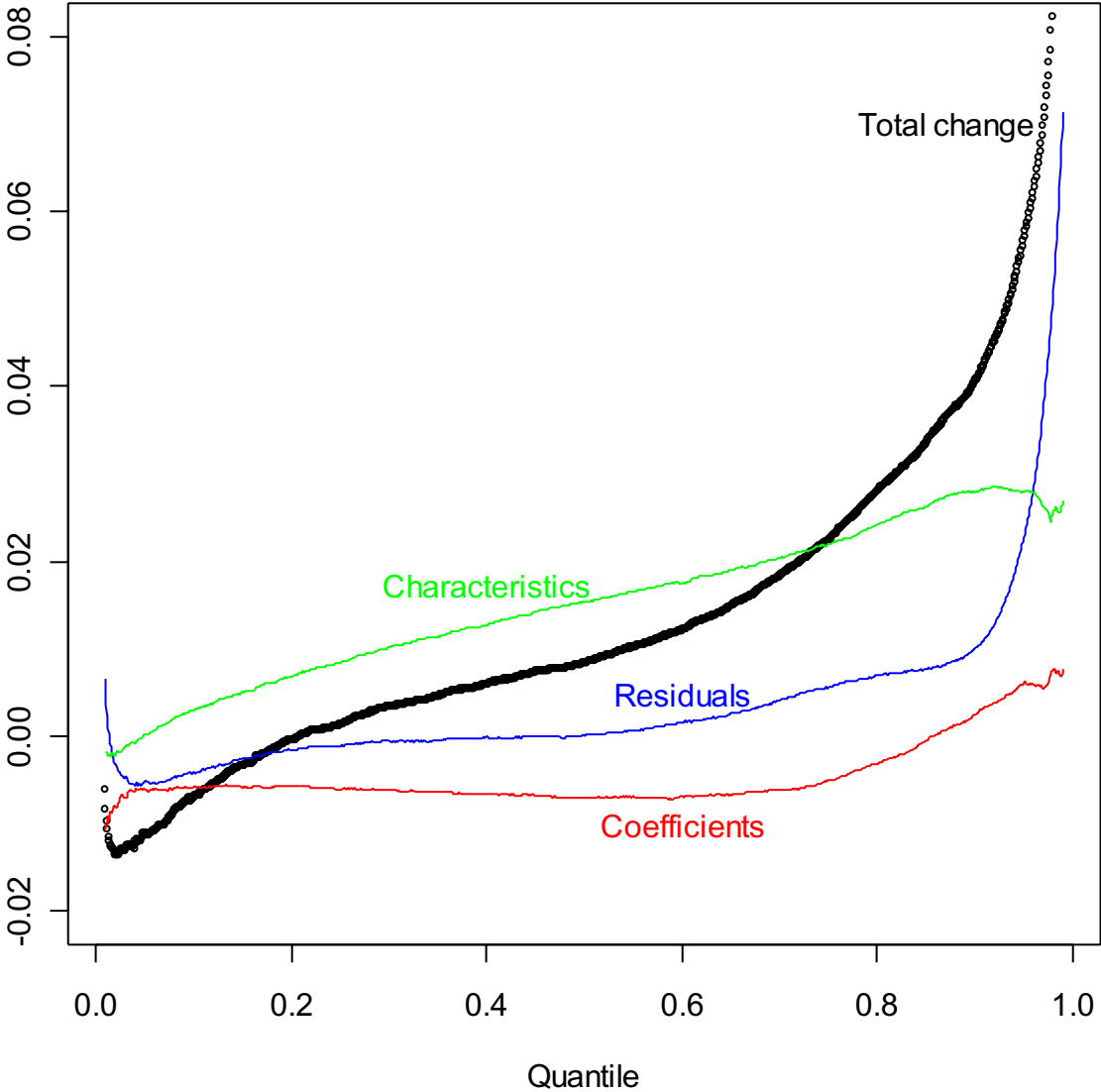


Figure C3: Decomposition of changes in distribution between 1984 and 1989



Appendix D: Monte Carlo comparison of three decomposition methods

In order to compare the quantile regression decomposition with the JMP and DFL decompositions, we have performed a Monte Carlo simulation. We consider two years, r and o , and we want to decompose the changes in the median and in the difference between the 90th and the 10th percentiles of the distribution of the dependent variable y . The wage structure in year r is the reference one. We consider two different data generating processes (DGP) for y^r and y^o .

The first is the model that JMP have in mind:

$$\begin{aligned}y_r &= 1 + x_r + u_r, \quad x_r \sim N(4,1), \quad u_r \sim N(0,0.5) \\y_o &= 0.8 + 1.1x_o + u_o, \quad x_o \sim N(5,1), \quad u_o \sim N(0,0.6).\end{aligned}$$

The level of the independent variable increased between r and o but it does not influence the variability of wages since the error term is independent of x . The return to the covariate increased but the constant decreased. Finally, the variance of the error term increased.

The second DGP is a standard location-scale model with an error term hit by a linear heteroscedastic scale:

$$\begin{aligned}y_r &= 1 + x_r + (0.5 + 0.125x_r)u_r, \quad x_r \sim N(4,1), \quad u_r \sim N(0,0.5) \\y_o &= 0.8 + 1.1x_o + (0.5 + 0.125x_o)u_o, \quad x_o \sim N(5,1), \quad u_o \sim N(0,0.5).\end{aligned}$$

The changes in characteristics and coefficients are exactly the same as in the first DGP. However, now the increase in the level of characteristics increases the variability of wages because of the heteroscedastic error term. The distribution of residuals conditionally on x does not change between both periods.

To implement the JMP decomposition we assume an independent error term and we do not try to condition on x when we estimate the residuals distribution. Most of the applications using the JMP approach follow a similar procedure. However, it is not clear if it was really

implemented in this way by JMP. The DFL decomposition was implemented as described in the original paper with a probit. A fourth and a fifth order polynomial in x was used as regressor for the probit estimation respectively for the first and the second DGP. These orders were chosen by minimizing the mean squared error (MSE). The quantile regression decomposition presented in the second section of the paper was implemented with 200 different quantile regressions uniformly distributed between 0 and 1.

The results are given in table D1 and D2. For each estimate, we give the true value, the mean, the standard error and the MSE obtained by 1000 simulations with 2000 observations. The results show that the decomposition of changes in the median of the distribution are very similar between the estimator for both DGPs. The standard errors of the DFL estimates are the highest ones, as expected since this estimator is the least restrictive one and do not assume a linear relationship between y and x . The standard errors of the JMP estimates are the lowest ones for the estimation of the effects of residuals and coefficients, as was expected since the error term is normally distributed. However, the estimates of the decomposition using quantile regression have smaller standard errors for the estimation of the effects of characteristics.

If we consider the decomposition of the 90-10 gap for the first DGP, we observe that all estimators are consistent and the ordering of the standard error is the same as for the decomposition of the changes in the median. For the second DGP, however, the results differ between the JMP decomposition and the two others. The JMP wrongly attributes to residuals the part of the difference due to the characteristics. With this DGP the residuals do not change between both periods and it is clearly wrong to attribute one part of the increase in inequality to residuals. This error arises from the heteroscedastic error term. The change in the distribution of characteristics puts more weight to the observations with higher within-group inequality. It is the type of situations that we find in the application. The quantile regression and the DFL decomposition are consistent with this type of error term and rightly find that

characteristics account for a part of the increase in inequality. Note however that the standard errors of the components of the DFL decomposition are such high that the MSE of the JMP estimates are smaller although they are inconsistent.

To conclude, these simulations show that the JMP decomposition is the most restrictive decomposition but the most efficient one if all restrictions are satisfied. In particular, assuming independent residuals strongly affects the conclusions in the presence of heteroscedasticity. In this case, the quantile regression decomposition is a consistent alternative. Finally, to be honest with the DFL decomposition, we could imagine a third DGP with a nonlinear functional form. In this case, the DFL would be the only consistent method. The choice of the estimator in applications must be guided by the plausibility of the assumptions. There is no estimator that dominates the others in all situations.

Table D1: Monte Carlo simulation for the first data generating process

Changes in	Method	Effects of											
		residuals				coefficients				characteristics			
		true	mean	σ	mse	true	mean	σ	mse	true	mean	σ	mse
Median	QR	0	-0.04	1.54	0.02	30	30.10	2.63	0.07	100	100.01	3.67	0.13
	JMP	0	0.03	1.01	0.01	30	29.89	2.26	0.05	100	99.96	4.52	0.20
	DFL					30	30.27	3.46	0.12	100	99.83	4.59	0.21
Interdecile range	QR	11.35	12.15	3.25	0.11	23.11	22.48	5.00	0.25	0	0.11	5.43	0.30
	JMP	11.35	11.82	2.68	0.07	23.11	22.67	4.25	0.18	0	0.13	7.17	0.51
	DFL					34.47	34.60	9.66	0.93	0	0.18	10.70	1.14

Note: all numbers have been multiplied by 100; 1000 replications; 2000 observations; details in text.

Table D2: Monte Carlo simulation for the second data generating process

Changes in	Method	Effects of											
		residuals				coefficients				characteristics			
		true	mean	σ	mse	true	mean	σ	mse	true	mean	σ	mse
Median	QR	0	-0.03	1.65	0.03	30	30.27	2.71	0.07	100	99.86	3.57	0.13
	JMP	0	-0.41	0.85	0.01	30	30.21	2.27	0.05	100	100.32	4.49	0.20
	DFL					30	30.47	3.82	0.15	100	99.76	4.91	0.24
Interdecile range	QR	0	0.07	3.14	0.10	22.53	22.75	4.83	0.23	7.47	7.53	5.22	0.27
	JMP	0	7.25	2.22	0.57	22.53	22.87	4.34	0.19	7.47	0.11	7.16	1.05
	DFL					22.53	23.37	10.34	1.08	7.47	6.72	11.13	1.24

Note: all numbers have been multiplied by 100; 1000 replications; 2000 observations; details in text.