Decomposition of Differences in Distribution under Sample Selection and the Gender Wage Gap

Santiago Pereda-Fernández*

July 11, 2022

Abstract

I address the decomposition of the differences between the distribution of outcomes of two groups when individuals self-select themselves into participation. I differentiate between the decomposition for participants and the entire population, highlighting how the primitive components of the model affect each of the distributions of outcomes. Additionally, I introduce two ancillary decompositions that help uncover the sources of differences in the distribution of unobservables and participation between the two groups. The estimation is done using existing quantile regression methods, for which I show how to perform uniformly valid inference. I illustrate these methods by revisiting the gender wage gap, finding that changes in female participation and self-selection have been the main drivers for reducing the gap.

Keywords: Decomposition, Distributional analysis, Gender wage gap, Quantile regression, Sample selection

JEL classification: C13, C14, C31, J16, J31

*Departamento de Economía, Universidad de Cantabria, Avenida de los Castros, s/n, 39005 Santander, Spain. I would like to thank Domenico Depalo, Franco Peracchi, Alexandra Soberón, Paolo Zacchia and seminar participants at Universidad de Cantabria, Università di Bologna and Università di Roma Tor Vergata. All remaining errors are my own. I can be reached via email at santiagopereda@gmail.com.

1 Introduction

Decomposition methods have been widely used to understand the sources of differences between the outcomes of two groups. Following the seminal works by Oaxaca (1973) and Blinder (1973), mean differences between two groups have been attributed to either differences in mean covariates (endowments effect/composition component) or differences in the slope parameters (coefficients effect/structural component). Subsequently, several methods have been proposed to account for differences between other features of the distribution, such as the unconditional quantiles. These decompositions typically rely on the ignorability assumption (selection on observables).

However, individuals present in the decomposition often constitute a self-selected sample. This requires accounting for it to estimate the structural parameters using sample selection methods. Moreover, the decomposition of a feature of the outcomes between two groups could depend on differences the unobservables that depend on the amount of self-selection and participation for each group. This is the case if one considers the distribution of actual outcomes rather than the distribution of potential outcomes that would be observed if everyone participated.

Additionally, sample selection raises the question of which population is the target of the analysis: participants or the entire population. In the first case, the comparison is between those individuals with non-zero outcomes. *E.g.*, one could be interested in understanding the differences in labor earnings of two groups of employed workers. However, it can be argued that including non-participants provides a more comprehensive comparison. This would amount to include unemployed workers and those out of the labor force in the analysis.

This paper makes the following contributions. First, I propose several decompositions according to the target population and the distributional feature of interest. Specifically, I consider decompositions for participants and the entire population for actual outcomes. Second, I propose two ancillary decompositions of the propensity score and the average value of the unobservable variable that affects the outcome, which is interpreted as ability. Third, I estimate the functionals that are used in the decomposition using the Quantile

Regression with Selection (QRS) estimator proposed by Arellano and Bonhomme (2017). I show the asymptotic properties of the estimated components of the decomposition, as well as how to carry uniformly valid inference. Fourth, I obtain new estimates of the evolution of the gender wage gap, finding that the unobserved ability of female workers has increased more than that of males. This, together with the fall in the gender participation gap have been two major factors in reducing the gender wage gap.

In the absence of self-selection, individual actual and potential outcomes coincide. In contrast, actual outcomes equal zero for non-participants when there is self-selection. Hence, there are two potentially relevant target populations to consider: just participants, who have a non-zero outcome, and the entire population, which also includes non-participants. The decomposition for both populations are related and, depending on the context, either could be more appropriate to analyze. I consider both of them, highlighting their differences and similarities in a general framework and under some simplifying assumptions. Furthermore, because differences in the unobservables can account for differences in the actual outcomes, the type of decompositions presented here differ from those that focus on the potential outcomes when every individual is assumed to participate.

Under ignorability, differences in outcomes can be split into differences in covariates and differences in the returns to the covariates. However, when there is sample selection on unobservables, there are two additional components that can explain the differences between the two groups. The first one is the selection component, which reflects how, *caeteris paribus*, the unobserved characteristics of a group may be more or less positively selected relative to those of the other group. The other one is the participation component, which reflects differences in outcomes that can be attributed to the differences in the participation rates between the two groups. This extends the decomposition into three components in a triangular model with a binary treatment considered in Pereda-Fernández (2022).

Most decompositions in a cross-sectional data setting have imposed no selection on unobservables.[1] This problem has been highlighted, *e.g.*, by Kunze (2008) or Huber (2015).

[1] There exist works that explicitly account for time-invariant unobservables in a panel data setting. See Fortin et al. (2011) and references therein for further details.

Some exceptions include Neuman and Oaxaca (2003, 2004) and Mora (2008), who used the Heckit correction to obtain their estimates corrected for selection, or Cukrowska-Torzewska and Lovasz (2016), who used a multinomial correction model. More recently, Huber et al. (2020) used mediation analysis tools to decompose the mean gap with sample selection under the assumption that the effect of the unobservables is homogeneous (*e.g.*, if the model is additively separable).

A common feature of these works is that they are only concerned with decompositions of mean differences. In contrast, the focus gradually expanded to analyze other features, including unconditional distributions. The latter are particularly relevant, as many statistics of interest can be expressed as a function of it. Regarding estimation, there exists a wide menu of available methods, including those based on quantile regression (Machado and Mata, 2005; Melly, 2006; Chernozhukov et al., 2013), on distribution regression Chernozhukov et al. (2013), and on reweighting (DiNardo et al., 1996; Firpo et al., 2018). A thorough review of the advantages and disadvantages of each method is found in Fortin et al. (2011).

A decomposition into four components has previously been considered by Chernozhukov et al. (2019), who also used it in a sample selection framework with a distributional regression estimator. The main difference with respect to the one considered in this paper regards the structural functions employed to model the outcomes. In their paper, they use bivariate Gaussian distributions to model the sample selection locally, whereas the approach in this paper is to model the unobservables with a copula globally. This allows to interpret the unobservables as the conditional ranks in the distribution of potential outcomes. Moreover, it is straightforward to relate differences in the value of these unobservables between the two groups to differences in the copula, the propensity score and the distribution of covariates, which may be of independent interest for the researcher.

Other related works are Maasoumi and Wang (2017) and Maasoumi and Wang (2019). In the first one they estimate the differences between functions of the distributions of the two groups for the selected sample and they bound the differences for the entire population. In the second one they analyze the evolution of the gender wage gap in the US using the estimator proposed by Arellano and Bonhomme (2017), including a decomposition of the gap between the two distributions of potential wages for the entire population. There are some differences relative to these two papers. First, I consider differences between the actual distributions and relate them to the different primitives of the model. Second, I consider four different sources of variability, which can help explain which are the determinants of the gap. The latter can be a problem not just if sample selection is ignored, but also if one analyzes the distributions of potential outcomes ignoring participation.

For estimation purposes, I consider a nonseparable model with a univariate unobservable variable in the outcome equation, which naturally points at quantile regression methods. The QRS estimator can estimate the structural function that relates potential wages to covariates and unobservables and the copula that captures the amount of self-selection. This, together with an estimator of the propensity score and the sample distribution of covariates are all the ingredients needed to estimate the different components of the decomposition. Moreover, I discuss under which simplifying assumptions other alternative estimators can be used.

The decomposition methods have been applied to a wide variety of topics, including test scores gaps between genders (Sohn, 2008), schools (Krieg and Storer, 2006) or countries (McEwan and Marshall, 2004), differences in students' enrollment (Borooah and Iyer, 2005), differences in health insurance coverage between different demographic groups (Bustamante et al., 2009), or gender differences in smoking behavior (Bauer et al., 2007). However, most decomposition studies revolved around wage gaps between genders Kunze (2008), race (Barsky et al., 2002), unionization status (Card, 1996), or public-private employees (Depalo and Pereda-Fernández, 2020).

I apply the methods presented in this paper to decompose the gender wage gap. I revisit the estimates by Maasoumi and Wang (2019), which are based on the Current Population Survey for the period 1976-2013. I perform several decompositions, considering different population targets (actual earnings for participants and the full population, as well as potential outcomes for the full population) and several statistics (mean, unconditional distributions and generalized entropy indices). I find that considering just employed workers understates the wage gap relative to consider the entire population. However, the gender gap has substantially diminished over the period considered in both cases. The main contributing factors for this reduction are the increase in female participation, and the increase in the average value of both observed and unobserved characteristics of employed females.

The rest of the paper is organized as follows: Section 2 presents the model and the main functions of interest. The decompositions of interest are presented in Section 3, including particular cases of interest that are nested in the more general model. Section 4 describes the estimation of the decompositions for the general model, describing its asymptotic properties and showing how to conduct inference. In Section 5 I discuss several particular cases, suggesting other feasible methods under such simplifying assumptions. Section 6 revisits the estimates of the evolution of the gender wage gap, and Section 7 concludes.

2 The Model

Consider the following selection model:

$$Y = g_D(X, U) S \tag{1}$$

$$S = \mathbf{1}(\pi_D(Z) - V > 0) \tag{2}$$

where Y denotes the continuous outcome of interest, X a set of predetermined covariates, $Z \equiv (Z_1, X')$ is composed of the instrument Z_1 and the predetermined covariates, S is a binary indicator for participation, D is an observed variable that denotes the group to which the individuals of the population can belong, and U and V are two unobservable random variables. Equations 1-2 conform a sample selection model that allows for a broad class of differences between individuals of different groups. For expositional simplicity, we focus on the leading binary case, *i.e.*, $D = 0, 1$.[2]

This system can be used to model several economic phenomena, such as labor wages,

[2]This framework is reminiscent of those considered for the study of the marginal treatment effect (MTE; Björklund and Moffitt, 1987). Such framework can be rationalized by a generalized Roy model with imperfect information (Pereda-Fernández, 2022).

denoted by Y. These could be different for people belonging to different demographic groups, like gender (D). Equation 2 reflects the fact that only employed workers ($S = 1$) have a wage. These are modeled by Equation 1, and they depend on their observed and unobserved characteristics, respectively given by X and U. The latter, together with V are the unobservables of the model, which can be interpreted as the ability of the worker and the unobserved propensity to be unemployed, respectively.[3]

Following Heckman and Vytlacil (2005), the distribution of V can be normalized to be uniformly distributed on the unit interval. This is convenient, as π_d can now be interpreted as the propensity score.[4] Moreover, if the same normalization is applied to the distribution of U, Equation 1 uses the Skorohod representation, allowing us to interpret $g_d(x, u)$ as the structural quantile function (SQF). These normalizations offer two advantages: the joint distribution of the unobservables conditional on $D = d$ and $X = x$, denoted by $C_{d,x}(u, v) \equiv \mathbb{P}(U \leq u, V \leq v | D = d, X = x)$, can be interpreted as a copula and, by taking the inverse of the SQF with respect to U, one obtains the conditional distribution of potential outcomes if all were participants.[5]

The copula also conveys some important information for the policy maker regarding some counterfactual scenarios. It captures the amount of self-selection on unobservables, linking the participation decision to the outcome. Negative amounts of correlation are associated with positive selection of individuals into participation. Consequently, the more negative the amount of correlation, the lower the potential outcomes of non-participants relative to participants.

[3] One could also consider work intensity as an additional dependent variable in the model. *E.g.*, one could consider weekly hours or worked weeks, as in Fernández-Val et al. (2022). Such extensions could also be considered by appropriately modeling the unobservable variables, possibly adding some extra ones. Such comprehensive models could be informative about features that are not captured by a simpler model, at the cost of making them computationally more complicated. Their study is left for future work.

[4] Note that the propensity score is sensitive to the instrument used. Therefore, using different instruments would lead to different results. See, *e.g.*, Mogstad et al. (2021) for further details on the interpretation of treatment effects in with multiple instruments, a framework related to the one considered in this paper.

[5] This constitutes a conceptual difference relative to Chernozhukov et al. (2019). They model the relation between the outcome variable and the participation decision using a bivariate model that is observationally equivalent. However, it is not possible to interpret the distribution of the unobservables as a copula with their representation. Hence, the interpretation of their unobservables is less clear, as they are locally defined.

To show that these normalizations are without loss of generality, consider an alternative data generating process determined by functions $\tilde{g}_D$ and $\tilde{\pi}_D$, as well as the distribution of the unobservables $\tilde{F}_{\tilde{U},\tilde{V}|D,X}$. Their observational equivalence is established by Lemma 1:

Lemma 1. *Let* $Y = \tilde{g}_D\left(X, \tilde{U}\right) S$ *and* $S = \mathbf{1}\left(\tilde{\pi}_D(Z) - \tilde{V} > 0\right)$, *where the distribution of the unobservables is given by* $\tilde{F}_{\tilde{U},\tilde{V}|D,X}\left(\tilde{U}, \tilde{V}|D, X\right)$, *with marginal distributions* $\tilde{F}_{\tilde{U}|D,X}$ *and* $\tilde{F}_{\tilde{V}|D,X}$, *that may depend non-trivially on* X. *Then, there exist* g_D, π_D *such that the model given by Equations 1-2, where* $U|D,X \sim U[0,1]$ *and* $V|D,X \sim U[0,1]$, *generates the same distribution of* (Y, S, D, Z).

The identification conditions are listed in Arellano and Bonhomme (2017).[6] Apart from some standard assumptions (exclusion restriction, continuous outcomes, propensity score strictly within the unit interval and a well-defined continuous copula for the unobservables), they require either identification at infinity, or a continuous instrument and that the copula be real analytic with respect to its second argument.[7] One could relax the latter by imposing a parametric assumption, a possibility considered in Appendix C. Additionally, these assumptions can be relaxed if the copula is homogeneous with respect to the covariates and one uses variation in the covariates as a source of exogenous variation.

3 Decompositions

Decompositions typically vary according to the distributional feature that is decomposed, such as the difference of the means or of the unconditional quantiles. Sample selection introduces two additional margins of choice regarding which decomposition to make. First, because only a fraction of the population participates, one could consider a decomposition

[6]Note that the identification is based on instrumental variables, rather than on a control function approach. A similar model was considered using control functions was considered by Fernández-Val et al. (2022). The two models are non-nested, and in particular Fernández-Val et al. (2022) works under the assumption that the unobservables of Z, including X. This contrasts with the assumptions in Arellano and Bonhomme (2017), which allows for non-trivial dependence between the unobservables and X. A comparison of the conditions required for identification with instrumental variables and control functions is provided, *e.g.*, by Heckman and Vytlacil (2007).

[7]The set of assumptions required for identification by Chernozhukov et al. (2019) is different, though they are related. They provide a comparison of both sets of assumptions in their paper.

for participants or for the entire population, assigning a value of zero for non-participants. Second, the distribution of the unobservables for participants and non-participants differ, making the distributions of actual and potential outcomes differ, too. The focus here is on actual outcomes of both participants and the full population.

3.1 Primitives of the Decompositions

The decomposition of the outcomes for either participants or the entire population requires to account for the following primitive functions: the SQF, $g_d(x,u)$, the propensity score, $\pi_d(z)$, the marginal distribution of the covariates, $F_Z^d(z)$, and the conditional copula of the unobservables, $C_{d,x}(u,v)$, for $d=0,1$. This contrasts with decompositions of potential outcomes, which account for selection only to estimate the structural parameters, and then base the decomposition on the SQF and the propensity score.

The leading feature that is decomposed is the mean outcome. This requires assigning a value to the counterfactual mean outcome for the population of interest when a combination of the distribution of the covariates or the estimated structural functions are changed to match those of the other group. Following the previous discussion, the mean outcome for participants with the distribution of the observables of group h, the SQF of group k, the copula of group l and the propensity score of group m is given by

$$\mathbb{E}\left[Y^{hklm}|S=1\right] \equiv \int_{\mathcal{Z}} \int_0^1 g_k(x,u)\, dG_{l,x}(u,\pi_m(z))\, dF_Z^h(z) \tag{3}$$

where $G_{d,x}(u,v) \equiv \mathbb{P}(U \leq u|D=d, X=x, V \leq v) = \frac{C_{d,x}(u,v)}{v}$ denotes the copula conditional on participation. This is the channel through which the unobservables affect the mean outcome. Similarly, the counterfactual value for the entire population is given by

$$\mathbb{E}\left[Y^{hklm}\right] \equiv \int_{\mathcal{Z}} \int_0^1 g_k(x,u)\, dC_{l,x}(u,\pi_m(z))\, dF_Z^h(z) \tag{4}$$

More generally, one can obtain equivalent expressions for the distributions for each group, that constitute the building blocks for other functionals of interest, notably unconditional quantiles. The unconditional cumulative distribution functions for participants and the entire population are respectively given by

$$F_{Y|S=1}^{hklm}(y) \equiv \int_{\mathcal{Z}} \int_0^1 \mathbf{1}(g_k(x,u) \le y)\, dG_{l,x}(u, \pi_m(z))\, dF_Z^h(z) \tag{5}$$

$$F_Y^{hklm}(y) \equiv \int_{\mathcal{Z}} \left[\int_0^1 \mathbf{1}(g_k(x,u) \le y)\, dC_{l,x}(u, \pi_m(z)) + (1 - \pi_m(z)) \right] dF_Z^h(z) \tag{6}$$

Consequently, the unconditional quantile functions are given by[8]

$$Q_{Y|S=1}^{hklm}(\tau) \equiv \inf \left\{ y : \tau \le F_{Y|S=1}^{hklm}(y) \right\} \tag{7}$$

$$Q_Y^{hklm}(\tau) \equiv \inf \left\{ y : \tau \le F_Y^{hklm}(y) \right\} \tag{8}$$

3.2 Main Decompositions

The decomposition in a sample selection model is more intricate than in the regular one. To see this, note that the covariates can have an impact on the final outcome through three different channels. First, they affect the distribution of potential outcomes through the SQF, the only channel present when there is no sample selection. Second, they affect the propensity to participate, making some individuals more likely to participate. Third, they affect the amount of self-selection through the distribution of the unobservables, which is indexed by x in general. In other words, some characteristics may display some correlation with the unobservables, either reinforcing or mitigating the differences between the two groups.

Consequently, to better assess how the covariates affect the propensity score and the selection on unobservables channels, it is useful to report two ancillary decompositions, presented in Section 3.3. Moreover, the decomposition is path dependent. I present one of the possible orders for the decomposition of mean differences for participants and describe

[8]Note that if Y is continuous and there is no bunching at any particular value, Equation 7 equals the inverse of the cumulative distribution function, but the same is not true for Equation 8, which has a mass point at $Y = 0$.

each component in turn.[9]

Note that, by definition, $\mathbb{E}[Y|D=1,S=1]=\mathbb{E}[Y^{1111}|S=1]$ and $\mathbb{E}[Y|D=0,S=1]=\mathbb{E}[Y^{0000}|S=1]$. The difference between these two can be decomposed as:

$$\begin{aligned}\mathbb{E}[Y|D=1,S=1]-\mathbb{E}[Y|D=0,S=1] &= \underbrace{\mathbb{E}\left[Y^{1111}|S=1\right]-\mathbb{E}\left[Y^{0111}|S=1\right]}_{\text{endowments component}} \\ &+ \underbrace{\mathbb{E}\left[Y^{0111}|S=1\right]-\mathbb{E}\left[Y^{0011}|S=1\right]}_{\text{coefficients component}} \\ &+ \underbrace{\mathbb{E}\left[Y^{0011}|S=1\right]-\mathbb{E}\left[Y^{0001}|S=1\right]}_{\text{selection component}} \\ &+ \underbrace{\mathbb{E}\left[Y^{0001}|S=1\right]-\mathbb{E}\left[Y^{0000}|S=1\right]}_{\text{participation component}} \qquad (9)\end{aligned}$$

Equation 9 decomposes the mean difference between the two groups into four components. The first two are those present in decompositions under the ignorability assumption. The endowments component captures how differences in the covariates between the two groups lead to differences in mean outcomes. The other term present in this type of decompositions is the coefficients component. It reflects how differences in the SQF between the two groups are related to differences in mean outcomes, which may be partly driven by discrimination. For example, in the Oaxaca-Blinder decomposition, this term equals the difference in the OLS coefficients between the two groups, scaled by the average covariates of one of them.

The remaining two terms arise in a sample selection framework. The participation component has the easiest interpretation, as it relates differences in the probability of participating for both groups to differences in mean outcomes. To get some intuition, assume that more able individuals (*i.e*, those with high U) are also more prone to participate (*i.e.*, low V). Then, as the propensity score increases from zero to one, the average ability of participants gets smaller towards its mean value, reducing the mean wage for participants.

The selection component links differences in the amount of self-selection into participation

[9] Specifically, there are a total of 24 decompositions, and the size of each component may vary in each of them. See Fortin et al. (2011) for further details on the limits of path dependent decompositions.

to differences in mean outcomes. The interpretation is slightly different from that of the participation component, as it affects the distribution of the unobservables without affecting the participation. Therefore, *caeteris paribus*, the higher the level of selection, the higher the average ability of participants and, consequently, the higher their mean outcome.

Note that because of the linearity of the expectation operator, each component can be written in terms of the difference of the primitive functions. For instance, the coefficients component can be written as

$$\mathbb{E}\left[Y^{0111}|S=1\right]-\mathbb{E}\left[Y^{0011}|S=1\right]=\int_{\mathcal{Z}}\int_0^1\left(g_1\left(x,u\right)-g_0\left(x,u\right)\right)dG_{1,x}\left(u,\pi_1\left(z\right)\right)dF_Z^0\left(z\right)$$

This convenient property does not generally hold, and in particular for the decomposition of the unconditional quantiles. Moreover, in some particular cases, the expression for some components can be simplified, as shown in Section 5. The decomposition for the entire population is analogous to the one for participants:

$$\begin{aligned}\mathbb{E}\left[Y|D=1\right]-\mathbb{E}\left[Y|D=0\right]=&\underbrace{\mathbb{E}\left[Y^{1111}\right]-\mathbb{E}\left[Y^{0111}\right]}_{\text{endowments component}}+\underbrace{\mathbb{E}\left[Y^{0111}\right]-\mathbb{E}\left[Y^{0011}\right]}_{\text{coefficients component}}\\&+\underbrace{\mathbb{E}\left[Y^{0011}\right]-\mathbb{E}\left[Y^{0001}\right]}_{\text{selection component}}+\underbrace{\mathbb{E}\left[Y^{0001}\right]-\mathbb{E}\left[Y^{0000}\right]}_{\text{participation component}}\end{aligned} \tag{10}$$

The interpretation of the different components is similar, with one notable difference: because the outcome for non-participants equals zero by definition, the mean outcome for the entire population is the mean outcome for participants multiplied by the propensity score. Therefore, the endowments and participation components operate through two channels: first, by changing the proportion of participants; second, by affecting the average outcome of participants. In contrast, the coefficients and selection components operate exclusively through the second channel.

The last two considered decompositions are those of the unconditional quantiles of the outcome for both target populations.[10] They consist of the same four components as the

[10] Additional decompositions can be constructed analogously. *E.g.*, if one is interested in differences

mean decompositions, although there are some important differences. From a mathematical standpoint, the unconditional quantile function is not a linear operator. Therefore, their expressions are more convoluted even under some simplifying assumptions. From a policy perspective, they are more informative, as they allow to assess which segments of the population present the largest differences. This makes them particularly relevant in cases in which the differences are of opposite signs for different quantiles of the distribution.

Formally, the decomposition of the unconditional quantile distribution for participants at quantile τ equals

$$
\begin{aligned}
Q_{Y|D=1,S=1}(\tau) - Q_{Y|D=0,S=1}(\tau) = {} & \underbrace{Q^{1111}_{Y|S=1}(\tau) - Q^{0111}_{Y|S=1}(\tau)}_{\text{endowments component}} + \underbrace{Q^{0111}_{Y|S=1}(\tau) - Q^{0011}_{Y|S=1}(\tau)}_{\text{coefficients component}} \\
& + \underbrace{Q^{0011}_{Y|S=1}(\tau) - Q^{0001}_{Y|S=1}(\tau)}_{\text{selection component}} + \underbrace{Q^{0001}_{Y|S=1}(\tau) - Q^{0000}_{Y|S=1}(\tau)}_{\text{participation component}} \qquad (11)
\end{aligned}
$$

while the decomposition of the unconditional quantile distribution for the entire population at quantile τ is given by

$$
\begin{aligned}
Q_{Y|D=1}(\tau) - Q_{Y|D=0}(\tau) = {} & \underbrace{Q^{1111}_{Y}(\tau) - Q^{0111}_{Y}(\tau)}_{\text{endowments component}} + \underbrace{Q^{0111}_{Y}(\tau) - Q^{0011}_{Y}(\tau)}_{\text{coefficients component}} \\
& + \underbrace{Q^{0011}_{Y}(\tau) - Q^{0001}_{Y}(\tau)}_{\text{selection component}} + \underbrace{Q^{0001}_{Y}(\tau) - Q^{0000}_{Y}(\tau)}_{\text{participation component}} \qquad (12)
\end{aligned}
$$

3.3 Ancillary Decompositions

The previous discussion highlighted the multiple channels through which the covariates can affect the final outcome in a sample selection framework. To better assess their role, researchers could perform two additional decompositions of the propensity score for the entire population and the average value of the unobservables for participants. These decompositions

in inequality, one could construct equivalent counterfactual values for the Gini index and apply the decomposition. As long as the statistic of interest can be expressed in terms of the same primitives of the model, and under some regularity conditions, they will be well-behaved. See, *e.g.*, Chernozhukov et al. (2013) for further details.

could be presented alongside the main decomposition, complementing the analysis.[11]

The first one is the participation decomposition. This is a regular means decomposition without selection using the propensity score as the dependent variable. As such, differences in the propensity score would be split into the usual endowments and coefficients components.

The second one is the self-selection decomposition. This is more reminiscent to the decompositions of the mean outcomes. To see this, note that the average value of U for participants with the distribution of the observables of group h, the copula of group l and the propensity score of group m is given by

$$\mathbb{E}\left[U^{hlm}|S=1\right] \equiv \int_{\mathcal{Z}} \int_0^1 u dG_{l,x}\left(u, \pi_m\left(z\right)\right) dF_Z^h\left(z\right) \tag{13}$$

Then, the difference of this statistic between the two groups can be decomposed as:

$$\begin{aligned}\mathbb{E}\left[U|D=1,S=1\right] - \mathbb{E}\left[U|D=0,S=1\right] = & \underbrace{\mathbb{E}\left[U^{111}|S=1\right] - \mathbb{E}\left[U^{011}|S=1\right]}_{\text{endowments component}} \\ & + \underbrace{\mathbb{E}\left[U^{011}|S=1\right] - \mathbb{E}\left[U^{001}|S=1\right]}_{\text{selection component}} \\ & + \underbrace{\mathbb{E}\left[U^{001}|S=1\right] - \mathbb{E}\left[U^{000}|S=1\right]}_{\text{participation component}} \end{aligned} \tag{14}$$

Similarly to the main decompositions, the endowments component captures differences in the copula and the propensity score due to differences in the covariates. On the other hand, the selection and participation components capture differences in the unobserved ability between the two groups due to differences in the copula and the propensity score, respectively.

[11] It is worth stressing that, because the variable that is decomposed is not the main outcome, the components of the ancillary decompositions are not comparable to those of the main decompositions, although they are related to each other.

4 Estimation

For expositional convenience, I present the estimators of the decompositions for participants. The estimators of the decomposition for the entire population are similarly constructed using the analogy principle. Their exact expressions and their asymptotic properties are presented in Appendix D. Throughout the entire section, the following assumptions are maintained:

Assumption 1. *$(Y_i, S_i, D_i, Z_i')'$ are iid for $i = 1, ..., n$, defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and take values in a compact set.*

Assumption 2. *The sample size for the d-th group is non-decreasing in n, such that $^{n}/_{n_d} \to p_d \in [0, \infty) \, \forall d$ as $n \to \infty$.*

Assumption 3. *Let $\pi_d (Z) \equiv \pi_d (Z; \gamma_d)$, with $\dim (\gamma_d) < \infty$. $\pi_d (Z; \gamma_d)$ is continuously differentiable with respect to γ_d. Moreover, there exists an asymptotically linear estimator $\hat{\gamma} \equiv (\hat{\gamma}_0', \hat{\gamma}_1')'$ that admits the following representation: $\hat{\gamma} - \gamma = -B^{-1} \frac{1}{n_d} \sum_{i=1}^{n} r_d (s_i, z_i; \gamma) + o_P \left(\frac{1}{\sqrt{n}} \right)$.*

Assumption 4. *Y has conditional density that is bounded from above and away from zero, a.s. on a compact set $\mathcal{Y}$. The density is given by $f_{Y|S=1,D,Z} (y)$ for $D = 0, 1$.*

Assumption 1 states the sampling process of the data. Assumption 2 further restricts it to ensure that the number of individuals in each group converges to a fixed proportion with respect to the sample size. The propensity score is required to satisfy some mild regularity conditions given by Assumption 3. It ensures that its asymptotic distribution is well-behaved, and it is satisfied by many estimation methods, such as maximum likelihood. Finally, Assumption 4 ensures that the dependent variable has a finite conditional density for the entirety of its support.

4.1 Main Decompositions

Let $\upsilon_{\ell}(z,\tau,\eta,f) \equiv \left(g_k(x,\tau), c_{l,x}(\tau,\eta), \pi_m(z), \int_{\mathcal{Z}} f dF_Z^h\right)$ be the vector that contains the structural functions, where $\ell \equiv (h,k,l,m)$ is used to keep notation compact.[12] This vector contains the four relevant components to compute the effects of decompositions 9-12. Its estimator, $\hat{\upsilon}_{\ell}$, needs to satisfy the following condition:

Condition 1. *The estimator of the components of the decomposition for the general model, $\hat{\upsilon}_{\ell}(z,\tau,\eta,f)$, satisfies the law $\sqrt{n}\left(\hat{\upsilon}_{\ell}(z,\tau,\eta,f) - \upsilon_{\ell}(z,\tau,\eta,f)\right) \Rightarrow \mathbb{Z}_{\upsilon_{\ell}}(z,\tau,\eta,f)$ for all ℓ, where $\mathbb{Z}_{\upsilon_{\ell}}(z,\tau,\eta,f)$ is a zero-mean Gaussian process.*

I proceed to derive the asymptotic distribution of the components of the decomposition for an estimator that satisfies Condition 1. Next, I present an adaptation of the estimator proposed by Arellano and Bonhomme (2017) and verify that it satisfies this condition. Moreover, I discuss how to implement it, as well as how to conduct inference.

Given $\hat{\upsilon}_{\ell}(z,\tau,\eta,f)$, the counterfactual mean outcome is estimated as:

$$\hat{\mathbb{E}}\left[Y^{\ell}|S=1\right] = \frac{1}{n_h}\sum_{i=1}^{n}\int_{\varepsilon}^{1-\varepsilon} \hat{g}_k(x_i,u)\, d\hat{G}_{l,x}(u,\hat{\pi}_m(z_i))\, \mathbf{1}(d_i=h) \tag{15}$$

where $n_h \equiv \sum_{i=1}^{n}\mathbf{1}(d_i=h)$ is the number of workers in group $h=0,1$. Each of the effects is computed as $\hat{\mathbb{E}}\left[Y^{\ell}|S=1\right] - \hat{\mathbb{E}}\left[Y^{\ell'}|S=1\right]$ for the appropriate choice of ℓ,ℓ'. Similarly, the estimator of the unconditional distribution equals:

$$\hat{Q}_{Y|S=1}^{\ell}(\tau) = \inf\left\{y : \tau \leq \hat{F}_{Y|S=1}^{\ell}(y)\right\} \tag{16}$$

where $\hat{F}_{Y|S=1}^{\ell}(y) = \frac{1}{n_h}\sum_{i=1}^{n}\left[\varepsilon + \int_{\varepsilon}^{1-\varepsilon}\mathbf{1}(\hat{g}_k(x_i,u)\leq y)\, d\hat{G}_{l,x}(u,\hat{\pi}_m(z_i))\right]\mathbf{1}(d_i=h)$.

To keep notation compact, denote the difference between two counterfactual values of a statistic indexed by ℓ and ℓ' by $\Delta^{\ell,\ell'}(\cdot)$. The asymptotic distribution of the components of the means and unconditional quantile decomposition are established in the following theorem:

[12] Strictly speaking, the last component of this vector is not of the structural functions, which would be F_Z^h. However, the former is more convenient to derive the asymptotic theory, so I refer to υ_{ℓ} as the vector of structural functions.

Theorem 1. *Let the estimator* $\hat{v}_\ell(z,\tau,\eta)$ *satisfy Condition 1. Under Assumptions 1-4, the following hold for all* (ℓ,ℓ')*:*

$$\sqrt{n}\Delta^{\ell,\ell'}\left(\hat{\mathbb{E}}\left[Y|S=1\right]-\mathbb{E}\left[Y|S=1\right]\right)\Rightarrow\mathbb{Z}_{\Delta^{\ell\ell'}Y|S=1}$$

where $\mathbb{Z}_{\Delta^{\ell\ell'}Y|S=1}$ *is a zero-mean Gaussian random variable, defined in Appendix A, and*

$$\sqrt{n}\Delta^{\ell,\ell'}\left(\hat{Q}_{Y|S=1}(\tau)-Q_{Y|S=1}(\tau)\right)\Rightarrow\mathbb{Z}_{Q|S=1,\ell\ell'}(\tau)$$

where $\mathbb{Z}_{Q|S=1,\ell\ell'}(\tau)$ *is a zero-mean Gaussian process, defined in Appendix A.*

The estimator proposed by Arellano and Bonhomme (2017), which I adapt it to account for different groups, satisfies Condition 1. To derive the asymptotic distribution of the estimator, I work with the following assumptions:

Assumption 5. $g_d(x,\tau)=x'\beta_d(\tau)$ *for* $d=0,1$*, where* β_d *is continuous and such that* $g_d(x,\tau)$ *is increasing in its last argument.*

Assumption 6. $C_{d,x}(u,v)\equiv C_{d,x}(u,v;\theta_d)$*, with* $\dim(\theta_d)<\infty$ *for* $d=0,1$. $C_{d,x}(u,v;\theta_d)$ *is uniformly continuous and differentiable with respect to its arguments a.e. Its density,* $c_{d,x}(u,v;\theta_d)$*, is well-defined and finite.*

Assumption 7. *Let* $\beta(\tau)\equiv\left(\beta_1(\tau)',\beta_0(\tau)'\right)'$ *and* $\theta\equiv(\theta_1',\theta_0')'$. $\forall\tau\in\mathcal{T}$, $\left(\beta(\tau)',\theta',\gamma'\right)'\in int\mathcal{B}\times\Theta\times\Gamma$*, where* $\mathcal{B}\times\Theta\times\Gamma$ *is compact and convex, and* $\mathcal{T}=[\varepsilon,1-\varepsilon]$*, for some constant* ε *that is used to avoid the estimation of extreme quantiles.*

Assumption 8. *Matrices of derivatives of the moments* $J_{\beta d}(\tau)$, $\tilde{J}_{\beta d}(\tau)$, $J_{\gamma d}(\tau)$, $\tilde{J}_{\gamma d}(\tau)$, $J_{\theta d}(\tau)$, $\tilde{J}_{\theta d}(\tau)$ *for* $d=0,1$*, as defined in Appendix A, are continuous and have full rank, uniformly over* $\mathcal{B}\times\Theta\times\Gamma\times\mathcal{T}$ *and* $d=0,1$.

Assumption 9. *Denote the support of* $\pi_d(Z)$ *conditional on* $X=x$ *by* $\mathcal{P}_{d,x}$. $\forall x\in\mathcal{X}$ *and* $d=0,1$, $\mathcal{P}_{d,x}\in[0,1]$ *is an open interval.*

Assumption 5 is made for convenience. Despite restricting the shape of the distribution of Y, it does not force conditional quantile curves to be parallel to each other for different values of the covariates. Moreover, it does not suffer from the curse of dimensionality as other more flexible alternatives, such as the partially linear model (Lee, 2003). Assumption 6 is also imposed for practical reasons, and it allows to use the most common parametric copulas, such as the Clayton, the Gaussian, or a Bernstein copula of a fixed order.[13] Importantly, it allows the copula to be dependent on the covariates. Assumption 7 is a regularity condition and Assumption 8 ensures that the moments needed to derive the asymptotic distribution of the estimator have full rank. Lastly, Assumption 9 ensures that the participation decision is not deterministic for any individual.

The following steps describe how to compute the estimator, and how to implement it to obtain an estimate of the decompositions:

1. Estimate the propensity score by $\hat{\pi}_d(z_i) \equiv \pi_d(z_i, \hat{\gamma}_d)$.

2. Fix a value of $t \in \Theta$. For $d = 0, 1$ and $\tau \in \mathcal{T}$, compute $\hat{\beta}_d(\tau; t)$ as

$$\hat{\beta}_d(\tau; t) \equiv \arg\min_{b \in \mathcal{B}} \sum_{i=1}^{n} \mathbf{1}(d_i = d)\, s_i \rho_{G_{d,x}(\tau, \hat{\pi}_d(z_i); t)}(y_i - x_i' b) \tag{17}$$

   where $\rho_u(x) \equiv xu\mathbf{1}(x \geq 0) - (1-u)\, x\mathbf{1}(x < 0)$ denotes the check function.

3. Estimate the copula parameters for $d = 0, 1$ by minimizing over $t \in \Theta$:

$$\hat{\theta}_d \equiv \arg\min_{t \in \Theta} \left\| \sum_{i=1}^{n} \int_{\varepsilon}^{1-\varepsilon} \mathbf{1}(d_i = d)\, s_i \varphi(\tau, z_i) \left[\mathbf{1}\left(y_i \leq x_i' \hat{\beta}_d(\tau; t)\right) - G_{d,x}(\tau, \hat{\pi}(z_i); t)\right] d\tau \right\| \tag{18}$$

   where $\varphi(\tau, z_i)$ is an instrument function.[14]

[13] Note that Assumption 6 does not provide a way to select the most appropriate copula given some data. This could be done, *e.g.*, by choosing the copula that minimizes the criterion function given in Equation 18 whenever the candidate copulas have the same number of parameters. Alternatively, it could be selected using cross validation as in Pereda-Fernández (2022).

[14] *E.g.*, a polynomial of the propensity score (Arellano and Bonhomme, 2017).

4. The slope parameters are obtained by $\hat{\beta}_d(\tau) \equiv \hat{\beta}_d\left(\tau; \hat{\theta}_d\right)$ for $d = 0, 1$.

5. The SQF and the copula are respectively given by $\hat{g}_d(x, \tau) = x'\hat{\beta}_d(\tau)$ and $\hat{C}_{d,x}(\tau, \pi) = C_{d,x}\left(\tau, \pi; \hat{\theta}_d\right)$.

Step 1 is standard and can be done, *e.g.*, by logit or probit. Step 2 is a rotated quantile regression conditional on a particular value of the copula. Note that in practice one needs to set a grid of values of τ, such as $\tau = \{0.01, ..., 0.99\}$. The third step is computationally expensive, as it involves the minimization over a non-convex space. The most common parametric copulas depend on few parameters, so that a grid search may be feasible. For Bernstein copulas, one can use the algorithm proposed in Pereda-Fernández (2022). The last two steps are immediate, and they yield the slope parameters, which are those estimated in step 2 for the value of the copula estimated in step 3, the SQF and the copula.

The following theorem establishes the uniform asymptotic distribution of the Arellano and Bonhomme (2017) estimator:

Theorem 2. *Let* $\hat{\vartheta}_d(\tau) \equiv \left(\hat{\beta}_d(\tau)', \hat{\theta}_d', \hat{\gamma}_d'\right)'$. *Under Assumptions 1-9, their joint asymptotic distribution is given by* $\sqrt{n}\left(\hat{\vartheta}_d(\tau) - \vartheta_d(\tau)\right) \Rightarrow \mathbb{Z}_{\vartheta_d}(\tau)$, *where* $\mathbb{Z}_{\vartheta_d}(\tau)$ *is a zero-mean Gaussian process with covariance function* $\Sigma_{\vartheta_d}(\tau, \tau')$, *where:*

$$\Sigma_{\vartheta_d}(\tau, \tau') = \sqrt{p_d p_{d'}} H_d(\tau) \Sigma_{R_d}(\tau, \tau') H_{d'}(\tau')'$$
$$H_d(\tau) = F_d^I(\tau)\left[C_d(\tau) + \left(I - \int_\varepsilon^{1-\varepsilon} D_d(u) F_d^I(u) du\right)^{-1} \int_\varepsilon^{1-\varepsilon} D_d(u) F_d^I(u) C_d(u) du\right]$$
$$\Sigma_{R_d}(\tau, \tau') = \mathbb{E}\left[\mathbb{Z}_{R_d}(\tau) \mathbb{Z}_{R_{d'}}(\tau')'\right]$$

and functions $C_d(\tau)$, $D_d(\tau)$, $F_d^I(\tau)$ *and* $\mathbb{Z}_{R_d}(\tau)$ *are defined in Appendix A.*

It only remains to verify that it satisfies Condition 1. This is done in the following corollary:

Corollary 1. *Let* $\hat{g}_d(x, \tau) = x'\hat{\beta}_d(\tau)$, $\hat{c}_{d,x}(\tau, \eta) = c_{d,x}\left(\tau, \eta; \hat{\theta}_d\right)$, $\hat{\pi}_d(z) = \pi_d(z; \hat{\gamma}_d)$ *and* $\hat{F}_Z^d(z) = \frac{1}{n_d}\sum_{i=1}^n \mathbf{1}(d_i = d)\mathbf{1}(Z_i \leq z)$. *They satisfy Condition 1.*

4.2 Inference

The expressions of the asymptotic variance of the different estimators are complex and depend on several density functions. Therefore, using resampling methods to obtain standard errors is preferable to obtaining closed-form expressions. In this paper I consider the weighted bootstrap (Ma and Kosorok, 2005). The following assumption defines the weights used to obtain all the bootstrap estimates:

Assumption 10. *Let W_i be an iid sample of positive weights, such that $\mathbb{E}(W_i) = 1$, $Var(W_i) = \omega_0 > 0$ and is independent of $(Y_i, D_i, S_i, Z_i')'$ for $i = 1, ..., n$.*

For the Arellano and Bonhomme (2017) estimator, the weighted bootstrap is implemented as follows:

- For each repetition $j = 1, ..., J$, draw the weights $w_{i,j}$ for $i = 1, ..., n$ that satisfy Assumption 10.
- Estimate the propensity score using the weights for each observation. Denote the estimate by $\hat{\pi}^*_{d,j}(z_i) \equiv \pi_d\left(z_i, \hat{\gamma}^*_{d,j}\right)$.
- Estimate the slope and copula parameters by adding the weights to Equations 17-18:

$$\hat{\beta}^*_{d,j}(\tau; t) \equiv \arg\min_{b\in\mathcal{B}} \sum_{i=1}^{n} w_{i,j}\mathbf{1}(d_i = d)\, s_i \rho_{G_{d,x}\left(\tau, \hat{\pi}^*_{d,j}(z_i); t\right)}(y_i - x_i' b)$$

$$\hat{\theta}^*_{d,j} \equiv \arg\min_{t\in\Theta} \left\| \sum_{i=1}^{n} \int_{\varepsilon}^{1-\varepsilon} w_{i,j}\mathbf{1}(d_i = d)\, s_i \varphi(\tau, z_i) \left[\mathbf{1}\left(y_i \leq x_i' \hat{\beta}^*_{d,j}(\tau; t)\right) - G_{d,x}\left(\tau, \hat{\pi}^*_{d,j}(z_i); t\right)\right] d\tau \right\|$$

$$\hat{\beta}^*_{d,j}(\tau) \equiv \hat{\beta}^*_{d,j}\left(\tau; \hat{\theta}^*_{d,j}\right)$$

- Estimate the counterfactual mean outcomes and unconditional distributions as:

$$\hat{\mathbb{E}}^*\left[Y_j^{\ell} | S = 1\right] = \frac{1}{n_h} \sum_{i=1}^{n} w_{i,j} \int_{\varepsilon}^{1-\varepsilon} \hat{g}^*_{k,j}(x_i, u)\, d\hat{G}^*_{l,x,j}\left(u, \hat{\pi}^*_{m,j}(z_i)\right) \mathbf{1}(d_i = h)$$

$$\hat{F}^{\ell,*}_{Y_j|S=1}(y) = \frac{1}{n_h} \sum_{i=1}^{n} w_{i,j} \left[\varepsilon + \int_{\varepsilon}^{1-\varepsilon} \mathbf{1}\left(\hat{g}^*_{k,j}(x_i, u) \leq y\right) d\hat{G}_{l,x,j}\left(u, \hat{\pi}^*_{m,j}(z_i)\right)\right] \mathbf{1}(d_i = h)$$

where $\hat{g}^*_{d,j}(x,u) \equiv x'\hat{\beta}^*_{d,j}(u)$ and $\hat{G}^*_{d,x,j}(u,\pi) \equiv G_{d,x}\left(u,\pi;\hat{\theta}^*_{d,j}\right)$.

- The remaining counterfactual unconditional quantile function and the components of the decomposition are computed as described in the text. They are denoted by $\hat{Q}^{\ell,*}_{Y_j|S=1}(\tau)$, $\Delta^{\ell,\ell'}\hat{\mathbb{E}}^*[Y_j|S=1]$, and $\Delta^{\ell,\ell'}\hat{Q}^*_{Y_j|S=1}(\tau)$.

- Once all J estimates have been obtained, estimate the variance of each of the statistics as $\frac{q_{0.75}(\tau)-q_{0.25}(\tau)}{z_{0.75}-z_{0.25}}$, where z_p is the p-th quantile of the standard normal distribution, and $q_p(\tau)$ is the p-th quantile of the distribution of the statistic, for $j=1,...,J$.

This bootstrap estimator of the variance is based on the one presented in Chernozhukov et al. (2013). Even though it is possible to use the variance of the estimator across repetitions of the bootstrap to obtain the standard errors, it would require additional conditions for it to be valid (Kato, 2011). This estimator only requires that the bootstrap converges in distribution to the asymptotic distribution of the sample estimator, which is established in the following theorem:

Theorem 3. *Under Assumptions 1-10, the weighted bootstrap estimators are denoted by* $\hat{\vartheta}^*_{d,j}(\tau)$, $\Delta^{\ell,\ell'}\hat{\mathbb{E}}^*[Y_j|S=1]$, *and* $\Delta^{\ell,\ell'}\hat{Q}^*_{Y_j|S=1}(\tau)$. *They consistently estimate the limiting laws of* $\hat{\vartheta}_d(\tau)$, $\Delta^{\ell,\ell'}\hat{\mathbb{E}}[Y|S=1]$, *and* $\Delta^{\ell,\ell'}\hat{Q}_{Y|S=1}(\tau)$. *Moreover,*

$$\sqrt{\frac{n}{\omega_0}}\left(\hat{\vartheta}^*_{d,j}(\tau)-\hat{\vartheta}_d(\tau)\right) \Rightarrow \mathbb{Z}_{\vartheta_d}(\tau)$$

$$\sqrt{\frac{n}{\omega_0}}\left(\Delta^{\ell,\ell'}\hat{\mathbb{E}}^*[Y_j|S=1]-\Delta^{\ell,\ell'}\hat{\mathbb{E}}[Y|S=1]\right) \Rightarrow \mathbb{Z}_{\Delta^{\ell\ell'}Y|S=1}$$

$$\sqrt{\frac{n}{\omega_0}}\left(\Delta^{\ell,\ell'}\hat{Q}^*_{Y_j|S=1}(\tau)-\Delta^{\ell,\ell'}\hat{Q}_{Y|S=1}(\tau)\right) \Rightarrow \mathbb{Z}_{Q|S=1,\ell\ell'}(\tau)$$

On top of providing uniform confidence bands for the functionals of interest and the intermediate functions, the weighted bootstrap can be used to carry out uniform inference using, *e.g.*, a Kolmogorov-Smirnov test to any of the components of the decomposition of the unconditional quantiles. For example, one could test the null hypothesis that one of the components equals a specific value, $\Delta^{\ell,\ell'} Q_{Y|S=1}(\tau)$. The test statistic would be given by

$$KS_n = \sup_{\tau \in \mathcal{T}} \sqrt{n} \hat{\Sigma}_{Q|S=1,\ell\ell'}(\tau)^{-1/2} \left| \Delta^{\ell,\ell'} \left(\hat{Q}_{Y|S=1}(\tau) - Q_{Y|S=1}(\tau) \right) \right|$$

where $\hat{\Sigma}_{Q|S=1,\ell\ell'}(\tau)$ is an estimator of the asymptotic variance of $\Delta^{\ell,\ell'} \hat{Q}_{Y|S=1}(\tau)$, such as the one proposed in the weighted bootstrap algorithm. The critical value $c_{1-\alpha}$ would be $1-\alpha$ quantile of the distribution of the bootstrapped of the KS_n statistic. Similar uniform confidence bands can be constructed for other functionals of interest.

4.3 Ancillary Decompositions

The ancillary decompositions are also based on the vector $\upsilon_\ell(z, \tau, \eta, f)$. The counterfactual values of the mean propensity score and the mean value of the unobservables are given by

$$\hat{\mathbb{E}}\left[\pi^\ell | S=1\right] = \frac{1}{n_h} \sum_{i=1}^{n} \hat{\pi}_m(z_i) \mathbf{1}(d_i = h) \tag{19}$$

$$\hat{\mathbb{E}}\left[U^\ell | S=1\right] = \frac{1}{n_h} \sum_{i=1}^{n} \int_{\varepsilon}^{1-\varepsilon} u d\hat{G}_{l,x}(u, \hat{\pi}_m(z_i)) \mathbf{1}(d_i = h) \tag{20}$$

The asymptotic distribution of the components of the two ancillary distributions is established in the following theorem:

Theorem 4. *Let the estimator $\hat{\upsilon}_\ell(z, \tau, \eta)$ satisfy Condition 1. Under Assumptions 1-4, the following hold for all (ℓ, ℓ'):*

$$\sqrt{n}\Delta^{\ell,\ell'} \left(\hat{\mathbb{E}}[\pi | S=1] - \mathbb{E}[\pi | S=1] \right) \Rightarrow \mathbb{Z}_{\Delta^{\ell\ell'}\pi|S=1}$$

$$\sqrt{n}\Delta^{\ell,\ell'} \left(\hat{\mathbb{E}}[U | S=1] - \mathbb{E}[U | S=1] \right) \Rightarrow \mathbb{Z}_{\Delta^{\ell\ell'}U|S=1}$$

where $\mathbb{Z}_{\Delta^{\ell\ell'}\pi|S=1}$ and $\mathbb{Z}_{\Delta^{\ell\ell'}U|S=1}$ are zero-mean Gaussian random variables.

5 Particular Cases

Several particular cases are nested by the general model. I discuss how the different estimands change and which methods can be used to estimation them.

5.1 Additively separable unobserved term

An interesting case arises when $Y = \left(g_D(X) + \tilde{U}\right) S$, where $\tilde{U} = Q_{U|D,X}(U)$ denotes the conditional quantile function of the uniformly distributed random variable U. For example, if the additive error term $\tilde{U}$ is homoskedastic and normally distributed, then $Q_{U|D,X}(U) = \Phi^{-1}(U)$, where $\Phi(\cdot)$ is the standard normal cdf. Importantly, as long as $g_j(x)$ contains an intercept, $\mathbb{E}(\tilde{U}) = 0$. This case yields a substantial simplification to the expected outcome:

$$\begin{aligned}\mathbb{E}\left[Y^{hklm}|S=1\right] &= \int_{\mathcal{Z}}\int_0^1 \left[g_k(x) + Q_{U|k,x}(u)\right] dG_{l,x}(u, \pi_m(z))\, dF_Z^h(z) \\ &= \int_{\mathcal{X}} g_k(x)\, dF_X^h(x) + \int_{\mathcal{Z}}\int_0^1 Q_{U|k,x}(u)\, dG_{l,x}(u, \pi_m(z))\, dF_Z^h(z)\end{aligned}$$

where $F_X^h(x)$ is the cdf of X for group $h = 0, 1$ and $\mathcal{X}$ its support. Similarly,

$$\mathbb{E}\left[Y^{hklm}\right] = \int_{\mathcal{X}} g_k(x)\, \pi_m(z)\, dF_Z^h(z) + \int_{\mathcal{Z}}\int_0^1 Q_{U|k,x}(u)\, dC_{l,x}(u, \pi_m(z))\, dF_Z^h(z)$$

Both means can be split into two terms. One depends exclusively on the separable term $g_d(x)$ and the distribution of the covariates, without the instrument. The other depends on the copula, the propensity score and the distribution of the covariates including the instrument, but not on $g_d(x)$. Such a simplification follows from the linearity of the expectation operator, which does not hold for the unconditional quantiles. Note also that a change in the covariates results in a parallel shift of the conditional distribution function, without affecting its shape, limiting the amount of heterogeneity that the model can display.

The estimation of the SQF is simplified because it is comprised of the sum of a potentially nonlinear term that depends exclusively on the covariates, and the unobserved ability. Thus, it is possible to estimate $g_d(x)$ nonparametrically. For example, Das et al. (2003) proposes a two-step series estimator which, in the first step nonparametrically regresses the selection variable on the covariates and the instruments, and in the second step it nonparametrically regresses the outcome variable on the covariates and the correction term.

5.2 Linear model

A particular case of the previous one appears when $g_d(x,u) = x'\beta_d + \tilde{u}$, where again $\tilde{u} = Q_{u|d,x}(u)$. This assumption slightly simplifies the expression of the first term of the mean outcome relative to the additively separable unobserved term model. Specifically, $\int_{\mathcal{X}} g_k(x)\, dF_X^h(x) = \mathbb{E}_h(X)'\beta_k$, where $\mathbb{E}_h$ stands for the expectation with respect to F_X^h.

Another two-step estimator that can be used in this case is the one proposed by Newey (2009). This is similar to Das et al. (2003), with the main difference being the fact that the regression in the second step is linear in the covariates whilst keeping the power series on the correction term. Other alternative estimators based on stronger parametric assumptions are those proposed, *e.g.*, by Heckman (1976, 1979); Lee (1983).[15]

5.3 No self-selection

Usually, it is assumed that the ability of participants is unrelated to their participation decision. Mathematically, the copula is independent: $C_{d,x}(u,v) = uv$ and $G_{d,x}(u,v) = u$. This assumption implies that, conditional on participation, there are no differences in the distribution of ability between the members of both groups, so the selection component vanishes in all decompositions. Moreover, the outcome value of participants is comparable to the potential value of non-participants. Hence, absence of self-selection also implies that the participation component vanishes in the decompositions for participants (Equations 9 and 11). This can be seen in Equations 3 and 5, as the statistics of interest depend on the propensity score exclusively through the copula. However, the same simplification does not apply to the decompositions for the entire population, as both the mean and the unconditional quantiles directly depend on the propensity score.

[15] See Vella (1998) for a review of these and other estimators for data with sample selection.

The combination of this assumption with the additively linear unobserved term yields another meaningful simplification, as both the participation and selection components equal zero. To see this, note that

$$\int_0^1 Q_{U|k,x}(u)\, dG_{l,x}(u, \pi_m(z)) = \int_0^1 Q_{U|k,x}(u)\, du = 0$$

where the second equality follows by construction.

In this case the estimation of the copula is superfluous: because the distribution of the unobserved ability is the same for participants and non-participants, there is no need to account for sample selection in the estimation of the SQF. This opens the possibility of using methods that are consistent under exogeneity. For example, one can use quantile regression and then apply the decomposition proposed by Machado and Mata (2005), distributional regression and follow Chernozhukov et al. (2013), or one could use reweighting methods, such as DiNardo et al. (1996) or Firpo et al. (2018).[16] Additionally, if the linearity assumption is combined with either no self-selection or all participants, the resulting model can be estimated by OLS. Regardless, the decompositions for the entire population still depend on the propensity score.

5.4 All participants

A particular case of the previous one takes place when every considered individual is a participant. In this case, the propensity score equals one for all individuals, so $S_i = 1$ for all $i = 1, ..., N$. As a consequence, the participation component vanishes. Also, because $C_{d,x}(u, 1) = G_{d,x}(u, 1) = u$, the selection component also disappears. Moreover, because there are no non-participants, the decomposition for participants and non-participants are the same. Additionally, under linearity the decomposition becomes the Oaxaca-Blinder decomposition.

[16] A detailed comparison of the advantages and disadvantages of these methods is provided by Fortin et al. (2011). Furthermore, Leorato and Peracchi (2015) compares quantile and distributional regression.

6 Evolution of the Gender Wage Gap

I study the evolution of the gender gap between earnings distributions using the Current Population Survey (CPS) dataset. I extend the analysis in Maasoumi and Wang (2019) by decomposing several features of the distribution of actual earnings for employed workers and the entire population. In addition, I do the two ancillary decompositions regarding differences in participation and self-selection between men and women.

To preserve the comparability to Maasoumi and Wang (2019), the sample covers the 1976-2013 period, and the regressions are done on a year-by-year basis, abstracting from any dynamics. I restrict the analysis to individuals between 18 and 64 years old, who work for wages and salary, do not live in group quarters and worked at least for 20 weeks and 25 hours per week in the previous year.[17] Like them, I estimate the propensity score using the probit estimator. I use the same regressors they used, *i.e.*, a third degree polynomial of age, four levels of education, four regional dummies, marital status, an indicator for white race, and the interactions between age and the other listed covariates, plus another variable they did not use: the number of children.

The dependent variable is mean log hourly wages.[18] The specification for the QRS estimator uses the same set of variables (except for the interactions between age and the remaining covariates), the same instrument (number of children below 5 years old) and two parametric copulas: the Frank and the Gaussian copulas.[19] However, there are some slight differences along several dimensions: I include the regressor number of children, the quantile grid, the propensity score, and the objective function used to estimate the copula.[20]

[17] Note that this analysis excludes part-time workers, seasonal workers, self-employed, or the total number of hours worked (which affects the intensive margin). Despite this, the gaps considered in this paper are relevant for a large fraction of the working age population, and the elements analyzed could also be relevant for the analysis of a more comprehensive gender gap that accounts for the aforementioned factors. Such analysis is beyond the scope of this paper.

[18] As in Maasoumi and Wang (2019), it is computed as the logarithm of the total wage and salary income divided by the number of week and hours worked during the previous year and adjusted for inflation using the 1999 consumer price index adjustment factor.

[19] A discussion of the validity of the exclusion restriction is provided in Maasoumi and Wang (2019). They acknowledge that the instrument may be stronger for women than for men, as well as for earlier years than for the latter. In addition, it may be stronger for married workers than for those who are not.

[20] For precision, the quantile grid I used for the estimation is $(0.01, 0.02, ..., 0.99)$, while the one used

Additionally, I also allow for more flexible specifications that separately estimate the main equation according to race (white vs non-white), level of education (college vs less than college) and marital status (married vs non-married), which are reported in Section 6.5, and the estimates using the same specification as in Maasoumi and Wang (2019), which are reported in Appendix E.[21]

6.1 Evolution of Participation

To analyze the evolution of the gender earnings gap, I begin by analyzing the evolution of labor market participation for both genders. Table 1 reports the average estimated propensity score by gender for the entire period. There has been a marked catch-up between female and male participation: in 1976, female participation was roughly one third, steadily increasing to over one half in the early 2000s, to fall slightly in the aftermath of the financial crisis. Meanwhile, male participation has been more stable: it has been equal to around two thirds until the financial crisis, falling to about 60% afterwards.

Consequently, the gender participation gap has more than halved during the period, from 33% to 14% (Table 2). Its decomposition shows that almost the entirety of the gap is explained by the coefficients components. In words, there has been a structural increase in female participation into employment for women unrelated to gender differences in covariates. On the other hand, the endowments component has been either statistically not significant or slightly negative for most of the period. Indeed, only between 1987 and 1989 was this component positive and significant. Hence, the catch up in college education rates for female workers has not contributed to an increased participation in the labor force relative to men.

by Maasoumi and Wang (2019) was $(0.3, 0.4, ..., 0.7)$; I use the propensity score as the instrument $\varphi(u, z) = \hat{\pi}(z)$ as suggested in Arellano and Bonhomme (2017), whereas Maasoumi and Wang (2019) use $\varphi(u, z) = \sqrt{u(1-u)}\hat{\pi}(z)$, which puts less weight on values that are further away from the median; the objective function equals Equation 18, while the objective function used by Maasoumi and Wang (2019) is $\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{J}\left(\varphi(\tau_j, z_i)\left(\mathbf{1}\left(y_i \leq x_i'\hat{\beta}(\tau_j)\right) - G\left(\tau_j, \hat{\pi}(z_i); \hat{\theta}\right)\right)\right)$; the implemented quantile regression estimates from Stata in Maasoumi and Wang (2019) were in some cases numerically slightly worse than the ones in Matlab, *i.e.*, the value of the check function was smaller for the latter. See Appendix E for further details.

[21] For completeness, I also report the results of the decompositions of potential outcomes and the estimates of the general entropy measures considered by Maasoumi and Wang (2019) in Appendix E.

Table 1: Average propensity to work by gender

Year	Male	Female	Year	Male	Female
1976	68.0	35.0	1995	66.7	48.0
1977	68.6	36.1	1996	67.2	49.1
1978	68.6	37.5	1997	67.1	49.7
1979	69.8	39.2	1998	67.7	50.3
1980	69.1	40.7	1999	68.7	51.1
1981	67.6	40.4	2000	68.8	52.0
1982	64.8	38.7	2001	69.2	52.5
1983	61.6	37.9	2002	68.3	51.5
1984	61.4	39.0	2003	66.3	50.3
1985	63.1	40.8	2004	65.2	49.5
1986	64.4	41.7	2005	65.3	49.3
1987	64.7	42.6	2006	65.7	50.0
1988	66.6	46.0	2007	66.3	50.4
1989	67.0	46.8	2008	65.8	50.9
1990	68.8	47.7	2009	64.0	49.8
1991	67.8	47.6	2010	59.8	47.7
1992	66.7	47.8	2011	59.2	46.8
1993	65.6	47.4	2012	60.1	46.7
1994	65.5	47.2	2013	60.8	46.9

Notes: average estimated propensity score by year and gender; coefficients scaled by 100.

Table 2: Participation decomposition

Year	Total	EC	CC	Year	Total	EC	CC
1976	33.0***	-0.6***	33.7***	1995	18.7***	-0.4**	19.1***
1977	32.5***	-0.7***	33.2***	1996	18.0***	-0.2	18.3***
1978	31.1***	-0.6***	31.7***	1997	17.4***	-0.2	17.6***
1979	30.7***	-0.5***	31.2***	1998	17.4***	-0.3	17.7***
1980	28.4***	-0.4**	28.8***	1999	17.6***	-0.1	17.7***
1981	27.2***	-0.1	27.3***	2000	16.8***	0.0	16.8***
1982	26.1***	-0.1	26.2***	2001	16.7***	0.0	16.8***
1983	23.7***	0.0	23.8***	2002	16.8***	-0.2**	17.0***
1984	22.5***	0.0	22.5***	2003	15.9***	-0.2	16.1***
1985	22.3***	0.0	22.3***	2004	15.7***	-0.4***	16.1***
1986	22.6***	0.1	22.5***	2005	16.1***	-0.3**	16.3***
1987	22.1***	0.3**	21.8***	2006	15.7***	-0.2	15.9***
1988	20.6***	0.3*	20.3***	2007	15.8***	-0.4***	16.2***
1989	20.3***	0.3**	20.0***	2008	14.9***	-0.4***	15.4***
1990	21.1***	0.1	21.0***	2009	14.2***	-0.7***	15.0***
1991	20.2***	-0.3**	20.5***	2010	12.1***	-0.8***	12.9***
1992	18.9***	-0.2	19.0***	2011	12.4***	-0.8***	13.3***
1993	18.2***	-0.2	18.4***	2012	13.4***	-0.8***	14.2***
1994	18.3***	-0.2	18.5***	2013	13.9***	-1.0***	14.8***

Notes: Total, EC and CC respectively denote total difference, endowments component and coefficients component; coefficients scaled by 100; *, ** and ** respectively denote statistical significance at the 90%, 95% and 99% confidence level.

6.2 Evolution of Self-Selection

The second feature of interest is the evolution of differences in self selection, which depends mainly on the estimated copula. Because the values for different copulas are not directly comparable, I report the Kendall's τ correlation coefficients.[22] Table 3 reports the baseline estimates for each year and gender, both with the Frank and the Gaussian copulas. These coefficients indicate that the amount of selection into employment has steadily increased for female workers: until the early 80s, there used to be negative selection that turned positive afterwards.[23] On the other hand, the amount of selection for male workers has fluctuated more over time, being either above or below that of females depending on the year. The results are very similar with both copulas, which suggests that the choice of the parametric copula is of secondary importance. To assess the sensitivity of these results to the model used, I report in Table 28 in Appendix F a comparison of the correlation coefficients with those of the Heckman 2-stage estimator (Heckman, 1979). The findings show that the copula estimated by both models are very similar throughout the entire period.

A more informative way to understand these estimates is to compare the average value of the unobservable u across genders and periods. This is shown in Table 4 and Figure 1. For the entire period considered, the average value of u for full-time employed females steadily increased from slightly above 40 to around 60. The average value for employed males has slightly increased over time. However, it displayed much more fluctuation over time, attaining its maximum value of about 57 in 2010. In other words, along with the increase in participation, there has been an increase in the amount of self-selection into employment for women. Hence, potential earnings of non-employed women are lower than actual earnings of those employed, given the same observed characteristics.

The average selection difference between male and female workers has consequently become negative, decreasing by about 11.3 percentage points, although it also reflects the oscillation of the estimates for males. The participation components experienced a decrease

[22] Unlike to the more common Spearman's ρ correlation coefficient, Kendall's τ is invariant to the distribution of the marginals.

[23] Recall that a negative (positive) coefficient implies positive (negative) selection into employment.

Table 3: Kendall's τ correlation coefficients

	Frank copula		Gaussian copula	
Year	Male	Female	Male	Female
1976	0.26***	0.14***	0.26***	0.13***
1977	0.15*	0.09**	0.13	0.08**
1978	0.19**	-0.01	0.23***	-0.01
1979	0.23***	-0.02	0.21*	-0.02
1980	0.22***	0.00	0.24***	0.01
1981	0.19**	0.01	0.19**	0.01
1982	0.13*	-0.03	0.13	-0.04
1983	0.02	-0.10**	-0.05	-0.09**
1984	-0.10	0.02	-0.15*	0.02
1985	0.06	-0.05	0.10	-0.05
1986	0.08	-0.12***	0.08	-0.10***
1987	0.08	-0.19***	0.12*	-0.18***
1988	-0.07	-0.11***	-0.09	-0.10***
1989	-0.12	-0.10**	-0.17*	-0.10*
1990	-0.14*	-0.22***	-0.09	-0.22***
1991	-0.08	-0.11**	-0.09	-0.10**
1992	-0.09	-0.12***	-0.11	-0.12***
1993	-0.30***	-0.13***	-0.31***	-0.14***
1994	-0.27***	-0.17***	-0.24***	-0.17***
1995	-0.28***	-0.21***	-0.34***	-0.22***
1996	0.03	-0.22***	0.04	-0.21***
1997	-0.03	-0.20***	-0.01	-0.21***
1998	-0.30***	-0.21***	-0.27***	-0.21***
1999	0.04	-0.23***	-0.06	-0.22***
2000	0.05	-0.14**	0.07	-0.15**
2001	0.12**	-0.11**	0.13*	-0.13***
2002	0.03	-0.13**	0.03	-0.14***
2003	0.03	-0.24***	0.04	-0.25***
2004	0.07	-0.22***	0.04	-0.22***
2005	0.08	-0.22***	0.07	-0.22***
2006	0.26***	-0.17***	0.30***	-0.17**
2007	0.06	-0.28***	0.08	-0.30***
2008	0.11	-0.31***	0.10	-0.33***
2009	-0.12	-0.26***	-0.15	-0.26***
2010	-0.29***	-0.27***	-0.34***	-0.28***
2011	-0.04	-0.33***	-0.10	-0.33***
2012	-0.29***	-0.29***	-0.33***	-0.29***
2013	0.05	-0.25***	0.01	-0.26***

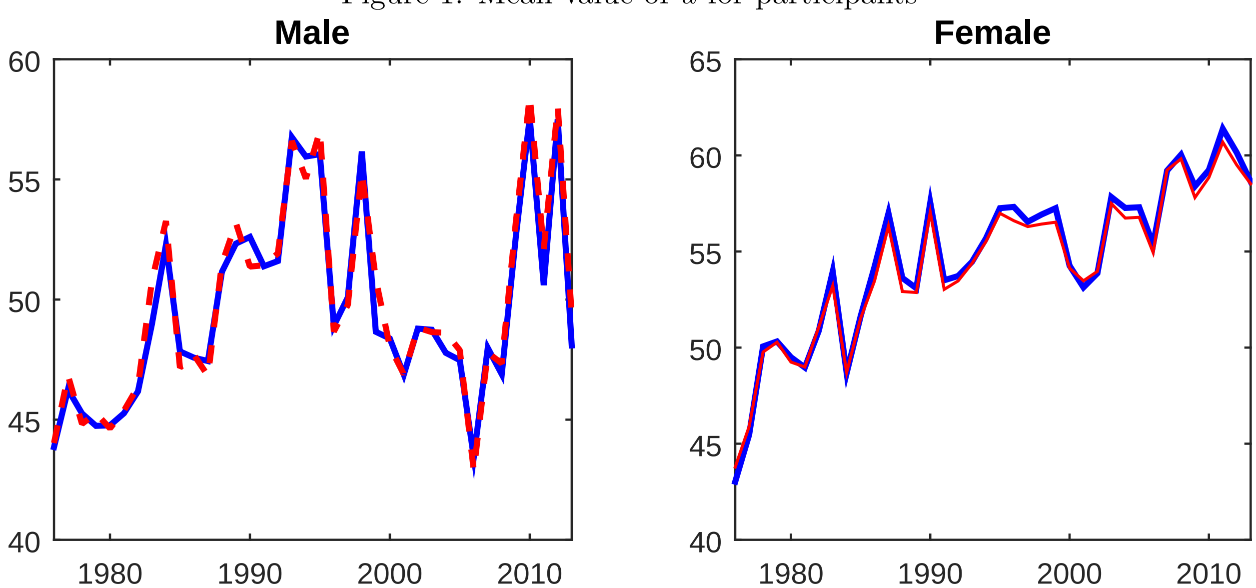
Notes: Kendall's τ correlation coefficients of the copula estimates by year and gender; *, ** and ** respectively denote statistical significance at the 90%, 95% and 99% confidence level.

Table 4: Mean value of u for participants

	Frank copula		Gaussian copula			Frank copula		Gaussian copula	
Year	Male	Female	Male	Female	Year	Male	Female	Male	Female
1976	43.9	43.0	44.0	43.8	1995	56.0	57.3	57.0	57.0
1977	46.3	45.5	46.8	45.9	1996	48.9	57.3	48.7	56.6
1978	45.3	50.1	44.8	49.8	1997	50.1	56.5	49.8	56.3
1979	44.7	50.3	45.3	50.3	1998	56.2	56.9	55.3	56.4
1980	44.8	49.5	44.6	49.2	1999	48.7	57.3	50.8	56.5
1981	45.3	49.0	45.4	49.0	2000	48.4	54.3	48.0	54.1
1982	46.2	50.9	46.4	51.1	2001	46.9	53.1	46.9	53.5
1983	49.0	54.0	50.8	53.2	2002	48.8	53.9	48.8	54.0
1984	52.3	48.7	53.3	48.7	2003	48.7	57.8	48.6	57.5
1985	47.8	51.6	47.2	51.5	2004	47.8	57.3	48.6	56.7
1986	47.6	54.2	47.7	53.5	2005	47.5	57.3	47.9	56.8
1987	47.4	57.0	46.8	56.3	2006	43.4	55.4	42.9	54.9
1988	51.2	53.6	51.5	52.9	2007	48.0	59.2	47.8	59.2
1989	52.3	53.1	53.2	52.9	2008	46.9	60.0	47.3	59.8
1990	52.6	57.5	51.4	57.1	2009	52.6	58.4	53.2	57.8
1991	51.4	53.5	51.4	53.0	2010	57.6	59.2	58.5	58.8
1992	51.6	53.7	51.9	53.5	2011	50.6	61.4	52.1	60.7
1993	56.8	54.5	56.6	54.4	2012	57.5	60.1	58.0	59.5
1994	56.0	55.7	55.0	55.5	2013	48.1	58.6	49.3	58.5

Notes: coefficients scaled by 100.

Figure 1: Mean value of u for participants



Notes: the solid blue line denotes the estimate with the Frank copula for all participants; the dashed red line denotes the estimate with the Gaussian copula for all participants; coefficients scaled by 100.

of about 6 percentage points, roughly half of the difference. Hence, even if the amount of self-selection (copula) had been the same for both genders, the increase of female employment rates contributed to the increase in average unobserved ability, setting a gap relative to male workers. On the other hand, the selection component is slightly less precisely estimated and it displays an unstable evolution. This is a direct consequence of the oscillating behavior of the male copula estimates. Regardless, the long term trend also points at an increase of the gap in favor of employed women. Lastly, because the baseline estimates are homogeneous with respect to the covariates, the endowments component is negligible.[24]

Table 5: Self-selection decomposition (Frank copula)

Year	Total	EC	SC	PC	Year	Total	EC	SC	PC
1976	0.8	-0.1***	-2.4	3.4***	1995	-1.2	0.1**	1.5	-2.8***
1977	0.8	-0.1*	-1.2	2.1**	1996	-8.4***	0.0	-5.6***	-2.8***
1978	-4.8**	-0.1**	-4.4**	-0.3	1997	-6.5***	0.0*	-4.0**	-2.4***
1979	-5.6**	-0.1***	-5.1***	-0.4	1998	-0.8	0.1	1.8	-2.6***
1980	-4.7*	-0.1***	-4.7***	0.0	1999	-8.6***	0.0	-5.8***	-2.8***
1981	-3.7*	0.0	-3.9**	0.2	2000	-5.9**	0.0	-4.2**	-1.7**
1982	-4.7**	0.0	-4.1*	-0.6	2001	-6.3***	0.0	-5.0***	-1.3**
1983	-5.0**	0.0	-3.3*	-1.7**	2002	-5.1**	0.0	-3.5*	-1.5**
1984	3.6	0.0	3.3	0.3	2003	-9.1***	0.0	-6.4***	-2.7***
1985	-3.8	0.0	-3.0**	-0.8	2004	-9.5***	0.0	-7.0***	-2.5***
1986	-6.6***	0.0	-4.8***	-1.8***	2005	-9.8***	0.0	-7.3***	-2.5***
1987	-9.5***	0.0	-6.8	-2.8***	2006	-11.9***	0.0	-10.0***	-1.9***
1988	-2.5	0.0	-0.9	-1.5***	2007	-11.2***	0.0	-8.1***	-3.2***
1989	-0.8	0.0	0.6	-1.3*	2008	-13.1***	0.0	-9.8***	-3.3***
1990	-4.9***	0.0	-1.7	-3.2***	2009	-5.8**	0.1	-3.2	-2.6***
1991	-2.1	0.0	-0.6	-1.6**	2010	-1.7	0.2***	0.6	-2.4***
1992	-2.1	0.0	-0.6	-1.5***	2011	-10.8***	0.0	-7.9***	-2.9***
1993	2.3	0.0	4.0*	-1.7***	2012	-2.6	0.2***	0.0	-2.8***
1994	0.3	0.0	2.4	-2.2***	2013	-10.5***	0.0	-8.0***	-2.5***

Notes: Total, EC, SC and PC respectively denote total difference, endowments component, selection component and participation component; coefficients scaled by 100; *, ** and ** respectively denote statistical significance at the 90%, 95% and 99% confidence level.

[24]Note that it is not exactly zero because there is variation in the copula because of differences in the propensity score, deriving themselves from differences in the distribution of Z between genders.

6.3 Evolution of Labor Earnings

Next consider the distributions of actual earnings for participants and the entire population by gender. Specifically, I report their means and the value of their 10th, 25th, 50th, 75th and 90th percentiles in Tables 6-7.[25] The actual earnings distribution for participants shows a small decrease in mean earnings for male workers and a slightly larger increase for mean female earnings.

This catch-up, however, masks an increase in the inequality within each gender, as interquantile ranges (IQR) increased both for male and female distributions: the 90-10 IQR increased from 132 to 170 percentage points for men, and from 112 to 158 for women; the 75-25 IQR similarly increased from 69 to 89 and 60 to 82 percentage points for male and female workers, respectively. Still, the evolution has been quite heterogeneous across the distributions of earnings. For male workers earnings followed a long term decrease for percentiles below the 75th, and a steady increase for those at the top of the distribution; for female workers there has been a gain for those above the 25th percentile, more pronounced at the top. Despite this catch-up, there is still a gap in favor of men at all quantiles.

By construction, earnings at any given quantile are smaller on the distribution for the full population than on that for participants, resulting in a smaller mean for both genders. Nonetheless, including non-participants increases the earnings gap. Following the increase in female labor participation, this distribution has steadily increased for females, reducing the gap relative to males by a bigger fraction than for the distribution of participants. On the other hand, mean earnings for males have oscillated across time following the changes in participation and average earnings for participants. Note that the fall in the male participation rate in the last years of the sample has been starker than that of females, prompting a decrease in mean earnings for both genders, along with a decrease of the gap.

[25] The estimates of the decompositions of mean earnings with the Heckman 2-stage estimator are very similar to those found with the QRS estimator. These are shown in Figures 14-15 in Appendix F.

Table 6: Actual earnings distributions for participants by gender (Frank copula)

	Male						Female					
Year	Mean	P10	P25	P50	P75	P90	Mean	P10	P25	P50	P75	P90
1976	2.72	2.03	2.40	2.77	3.09	3.36	2.29	1.74	2.00	2.30	2.60	2.86
1977	2.73	2.03	2.40	2.78	3.10	3.37	2.32	1.77	2.03	2.33	2.62	2.88
1978	2.74	2.03	2.40	2.78	3.12	3.38	2.31	1.76	2.02	2.33	2.63	2.89
1979	2.75	2.05	2.40	2.79	3.13	3.39	2.33	1.78	2.04	2.33	2.63	2.90
1980	2.73	2.03	2.40	2.78	3.12	3.38	2.32	1.78	2.04	2.32	2.62	2.89
1981	2.68	1.97	2.34	2.73	3.07	3.34	2.29	1.74	2.00	2.29	2.59	2.87
1982	2.67	1.95	2.31	2.71	3.06	3.34	2.27	1.72	1.98	2.28	2.59	2.86
1983	2.65	1.91	2.28	2.70	3.07	3.36	2.28	1.70	1.98	2.29	2.60	2.88
1984	2.65	1.90	2.28	2.70	3.07	3.37	2.30	1.70	1.99	2.31	2.63	2.91
1985	2.66	1.88	2.28	2.71	3.08	3.38	2.31	1.69	1.98	2.32	2.65	2.94
1986	2.66	1.88	2.28	2.71	3.09	3.39	2.32	1.69	1.99	2.33	2.67	2.97
1987	2.68	1.89	2.29	2.73	3.11	3.42	2.34	1.70	2.01	2.36	2.71	3.02
1988	2.68	1.89	2.29	2.72	3.11	3.42	2.35	1.69	2.02	2.38	2.72	3.03
1989	2.68	1.90	2.30	2.73	3.11	3.42	2.36	1.69	2.02	2.38	2.73	3.04
1990	2.67	1.89	2.29	2.71	3.10	3.43	2.37	1.69	2.03	2.39	2.74	3.06
1991	2.64	1.86	2.25	2.67	3.07	3.40	2.36	1.69	2.01	2.38	2.74	3.05
1992	2.63	1.84	2.23	2.67	3.06	3.39	2.36	1.69	2.00	2.38	2.74	3.06
1993	2.62	1.83	2.22	2.66	3.06	3.40	2.37	1.69	2.01	2.39	2.75	3.07
1994	2.60	1.80	2.19	2.64	3.04	3.39	2.36	1.66	2.00	2.38	2.75	3.08
1995	2.60	1.80	2.20	2.64	3.05	3.41	2.36	1.65	1.99	2.38	2.76	3.10
1996	2.61	1.80	2.19	2.63	3.04	3.39	2.36	1.65	1.99	2.37	2.76	3.10
1997	2.61	1.82	2.20	2.63	3.04	3.39	2.37	1.66	2.00	2.38	2.76	3.11
1998	2.63	1.84	2.23	2.66	3.06	3.43	2.39	1.70	2.02	2.40	2.78	3.11
1999	2.67	1.87	2.26	2.68	3.08	3.45	2.41	1.71	2.04	2.43	2.81	3.15
2000	2.68	1.87	2.26	2.69	3.10	3.48	2.43	1.71	2.05	2.44	2.83	3.17
2001	2.70	1.90	2.27	2.70	3.11	3.51	2.45	1.74	2.08	2.47	2.85	3.19
2002	2.70	1.90	2.28	2.70	3.11	3.52	2.47	1.75	2.09	2.48	2.86	3.22
2003	2.70	1.90	2.28	2.70	3.12	3.52	2.48	1.77	2.10	2.49	2.88	3.24
2004	2.69	1.88	2.26	2.69	3.12	3.51	2.49	1.76	2.11	2.50	2.88	3.25
2005	2.68	1.88	2.25	2.68	3.11	3.51	2.47	1.75	2.10	2.49	2.88	3.25
2006	2.68	1.88	2.24	2.67	3.10	3.50	2.46	1.72	2.08	2.47	2.86	3.23
2007	2.67	1.87	2.24	2.67	3.10	3.51	2.47	1.73	2.08	2.48	2.89	3.27
2008	2.68	1.88	2.25	2.67	3.10	3.50	2.48	1.75	2.10	2.50	2.89	3.27
2009	2.66	1.85	2.23	2.66	3.10	3.52	2.46	1.73	2.08	2.47	2.88	3.25
2010	2.68	1.87	2.25	2.68	3.11	3.53	2.48	1.73	2.09	2.49	2.90	3.28
2011	2.66	1.85	2.22	2.67	3.10	3.51	2.47	1.72	2.07	2.48	2.89	3.29
2012	2.64	1.82	2.20	2.65	3.10	3.52	2.46	1.71	2.06	2.47	2.87	3.27
2013	2.64	1.80	2.19	2.64	3.09	3.50	2.46	1.70	2.06	2.47	2.88	3.28

Table 7: Actual earnings distributions for the full population by gender (Frank copula)

	Male						Female					
Year	Mean	P10	P25	P50	P75	P90	Mean	P10	P25	P50	P75	P90
1976	1.88	0.00	0.00	2.50	2.97	3.28	0.81	0.00	0.00	0.00	2.09	2.60
1977	1.90	0.00	0.00	2.51	2.99	3.29	0.84	0.00	0.00	0.00	2.13	2.62
1978	1.91	0.00	0.00	2.50	3.00	3.31	0.87	0.00	0.00	0.00	2.16	2.64
1979	1.94	0.00	0.00	2.52	3.01	3.32	0.92	0.00	0.00	0.00	2.20	2.65
1980	1.91	0.00	0.00	2.50	3.00	3.30	0.95	0.00	0.00	0.00	2.23	2.66
1981	1.84	0.00	0.00	2.42	2.94	3.26	0.93	0.00	0.00	0.00	2.19	2.63
1982	1.76	0.00	0.00	2.34	2.91	3.25	0.89	0.00	0.00	0.00	2.15	2.61
1983	1.67	0.00	0.00	2.24	2.88	3.25	0.87	0.00	0.00	0.00	2.14	2.63
1984	1.66	0.00	0.00	2.23	2.89	3.26	0.91	0.00	0.00	0.00	2.19	2.67
1985	1.70	0.00	0.00	2.27	2.90	3.27	0.95	0.00	0.00	0.00	2.23	2.71
1986	1.74	0.00	0.00	2.30	2.92	3.28	0.98	0.00	0.00	0.00	2.26	2.74
1987	1.76	0.00	0.00	2.32	2.94	3.31	1.01	0.00	0.00	0.00	2.30	2.79
1988	1.81	0.00	0.00	2.35	2.95	3.32	1.10	0.00	0.00	0.00	2.37	2.82
1989	1.82	0.00	0.00	2.36	2.95	3.32	1.12	0.00	0.00	0.00	2.39	2.84
1990	1.85	0.00	0.00	2.38	2.95	3.32	1.14	0.00	0.00	0.00	2.41	2.86
1991	1.81	0.00	0.00	2.33	2.91	3.29	1.14	0.00	0.00	0.00	2.40	2.85
1992	1.77	0.00	0.00	2.29	2.89	3.28	1.15	0.00	0.00	0.00	2.40	2.86
1993	1.74	0.00	0.00	2.26	2.88	3.27	1.15	0.00	0.00	0.00	2.42	2.87
1994	1.72	0.00	0.00	2.23	2.86	3.27	1.14	0.00	0.00	0.00	2.41	2.87
1995	1.76	0.00	0.00	2.26	2.87	3.28	1.15	0.00	0.00	0.00	2.41	2.89
1996	1.78	0.00	0.00	2.27	2.87	3.27	1.18	0.00	0.00	0.00	2.42	2.89
1997	1.78	0.00	0.00	2.28	2.87	3.27	1.20	0.00	0.00	0.00	2.44	2.90
1998	1.80	0.00	0.00	2.31	2.89	3.30	1.22	0.00	0.00	1.16	2.46	2.92
1999	1.86	0.00	0.00	2.36	2.93	3.33	1.25	0.00	0.00	1.38	2.50	2.95
2000	1.87	0.00	0.00	2.37	2.94	3.36	1.28	0.00	0.00	1.51	2.53	2.98
2001	1.89	0.00	0.00	2.38	2.95	3.38	1.31	0.00	0.00	1.58	2.55	3.00
2002	1.87	0.00	0.00	2.37	2.95	3.39	1.29	0.00	0.00	1.49	2.55	3.01
2003	1.82	0.00	0.00	2.33	2.94	3.39	1.27	0.00	0.00	1.24	2.55	3.02
2004	1.78	0.00	0.00	2.30	2.93	3.37	1.25	0.00	0.00	0.00	2.55	3.02
2005	1.78	0.00	0.00	2.29	2.92	3.37	1.24	0.00	0.00	0.00	2.54	3.02
2006	1.79	0.00	0.00	2.28	2.91	3.37	1.25	0.00	0.00	0.85	2.54	3.01
2007	1.80	0.00	0.00	2.29	2.92	3.38	1.27	0.00	0.00	1.22	2.55	3.04
2008	1.79	0.00	0.00	2.29	2.92	3.37	1.29	0.00	0.00	1.38	2.57	3.05
2009	1.74	0.00	0.00	2.25	2.90	3.37	1.25	0.00	0.00	0.00	2.54	3.03
2010	1.64	0.00	0.00	2.15	2.89	3.37	1.21	0.00	0.00	0.00	2.54	3.05
2011	1.62	0.00	0.00	2.11	2.87	3.36	1.19	0.00	0.00	0.00	2.53	3.04
2012	1.63	0.00	0.00	2.13	2.87	3.36	1.18	0.00	0.00	0.00	2.51	3.02
2013	1.65	0.00	0.00	2.13	2.86	3.35	1.19	0.00	0.00	0.00	2.51	3.03

6.4 Main Decompositions

Tables 8-9 report the decompositions of the mean earnings gap for the two populations considered. The mean gap for participants has more than halved during the period: from over 40% gender gap for workers, it was equal to 18%. Out of the four components, the largest one in every period has been the coefficients component. Its size displays some yearly variation driven by the fluctuation of the slope and copula parameters for male workers. Analogously, the selection component also displays an erratic behavior, which is again a consequence of the estimates of the copula for males. The sign of this component is often negative, and when it is positive it is not significant at the 95% confidence level. Moreover, it displays a slightly downward trend, thus increasingly helping in the reduction of the mean gap (by 14 percentage points in the last year of analysis). This contrasts with the coefficients component, which does not exhibit a clear trend.[26]

The dynamics of the remaining two components has been more stable: they were initially positive, and they eventually became negative, therefore reducing the mean gender gap. Moreover, the magnitude of these two components has been more modest than that of the other two: the endowments components changed from 1 to -4 percentage points, whereas the participation component experienced a more pronounced fall (from 4 to -4 percentage points).

The gap for the entire population has followed a similar trend, steadily decreasing to less than a half of the gap in 1976. Specifically, from a gap of over 100%, it fell to less than 50%. However, its magnitude has always been larger, owing to the gender participation gap. Indeed, the participation component constitutes the lion share of the gap, and its reduction has been responsible for the majority of the reduction of the gap: from an initial 81 percentage points, it fell to 33. The remaining three components display a similar behavior

[26]The erratic behavior of these two components could be linked to the strength of the instrument, which is weaker for male workers. To see this, note that the autocorrelation of the total gap is 0.99. Out of the four components, it is also large for the endowments (0.99) and participation components (0.79), whereas for the coefficients and selection components are significantly smaller: 0.47 and 0.37, respectively. However, the autocorrelation of the sum of these two components is equal to 0.93. Therefore, it is plausible that the lack of strength of the instrument may be responsible for the large variations in the size of these two components even in consecutive years.

Table 8: Mean decomposition, actual earnings for participants (Frank copula)

Year	Total	EC	CC	SC	PC
1976	0.43***	0.01***	0.40***	-0.03	0.04***
1977	0.41***	0.01**	0.39***	-0.02	0.03**
1978	0.42***	0.01***	0.47***	-0.06**	0.00
1979	0.42***	0.01***	0.48***	-0.07***	-0.01
1980	0.41***	0.01**	0.47***	-0.06***	0.00
1981	0.40***	0.00	0.44***	-0.05**	0.00
1982	0.40***	0.01**	0.46***	-0.05*	-0.01
1983	0.38***	0.01**	0.44***	-0.04*	-0.02**
1984	0.35***	0.01**	0.30***	0.04	0.00
1985	0.35***	0.01***	0.40***	-0.04	-0.01
1986	0.34***	0.01***	0.43***	-0.07**	-0.03***
1987	0.33***	0.01**	0.47***	-0.10***	-0.04***
1988	0.32***	0.01***	0.35***	-0.01	-0.02***
1989	0.32***	0.01***	0.32***	0.01	-0.02*
1990	0.30***	0.01***	0.37***	-0.03	-0.05***
1991	0.28***	0.00	0.31***	-0.01	-0.02**
1992	0.26***	0.00	0.29***	-0.01	-0.02***
1993	0.25***	0.00	0.21***	0.06*	-0.03***
1994	0.23***	0.00	0.23***	0.04	-0.03***
1995	0.24***	0.00	0.27***	0.02	-0.05***
1996	0.25***	0.00	0.39***	-0.09***	-0.05***
1997	0.25***	-0.01**	0.36***	-0.06**	-0.04***
1998	0.24***	-0.01***	0.27***	0.03	-0.04***
1999	0.26***	-0.01**	0.40***	-0.09***	-0.05***
2000	0.25***	-0.01**	0.35***	-0.07**	-0.03**
2001	0.24***	-0.01***	0.35***	-0.08***	-0.02**
2002	0.23***	-0.01***	0.32***	-0.06*	-0.02**
2003	0.22***	-0.02***	0.38***	-0.10***	-0.04***
2004	0.20***	-0.02***	0.38***	-0.12***	-0.04***
2005	0.21***	-0.02***	0.40***	-0.12***	-0.04***
2006	0.22***	-0.03***	0.44***	-0.16***	-0.03**
2007	0.20***	-0.03***	0.43***	-0.14***	-0.06***
2008	0.19***	-0.03***	0.45***	-0.17***	-0.06***
2009	0.20***	-0.04***	0.33***	-0.05	-0.04***
2010	0.19***	-0.04***	0.26***	0.01	-0.04***
2011	0.19***	-0.04***	0.42***	-0.14***	-0.05***
2012	0.19***	-0.03***	0.27***	0.00	-0.05***
2013	0.18***	-0.04***	0.40***	-0.14***	-0.04***

Notes: Total, EC, CC, SC and PC respectively denote total difference, endowments component, coefficients component, selection component and participation component; *, ** and ** respectively denote statistical significance at the 90%, 95% and 99% confidence level.

to the one found in the decomposition of actual earnings for participants, although their size is slightly scaled down.

The same decompositions are performed for the unconditional distributions. I present the estimates for a number of years (1976, 1984, 1992, 2000, 2007, 2013) in Figures 2-3. Additionally, I report the estimates for several quantiles in Tables 29-38 in Appendix F.

Figure 2 shows the evolution of the gap for the entire quantile process. Several changes have taken place. First, the gap increases monotonically with the quantiles of the distribution in every year, with the exception of the extreme top quantiles. Second, there has been a generalized reduction at all quantiles and, the decrease in the gap in absolute value has also been larger for higher quantiles. Namely, the gap at the 10th percentile fell from 29% to 10%, whereas the gap at the 90th percentile decreased from 49% to 22%. As it was the case for the mean decomposition, the coefficients component has been the largest one for almost every quantile and every year. In contrast, the selection component has been relatively flat across quantiles, although it again displayed great variation across years, switching sign several times.

Regarding the participation and endowments components, their behavior is quite similar: their magnitude is smaller than that of the other two components, they initially were positive for the majority of the distribution and had a mild upward slope, and they have progressively flattened out, becoming negative and therefore reducing the gender gap. Moreover, their magnitude is similar to that of the decomposition of the mean.

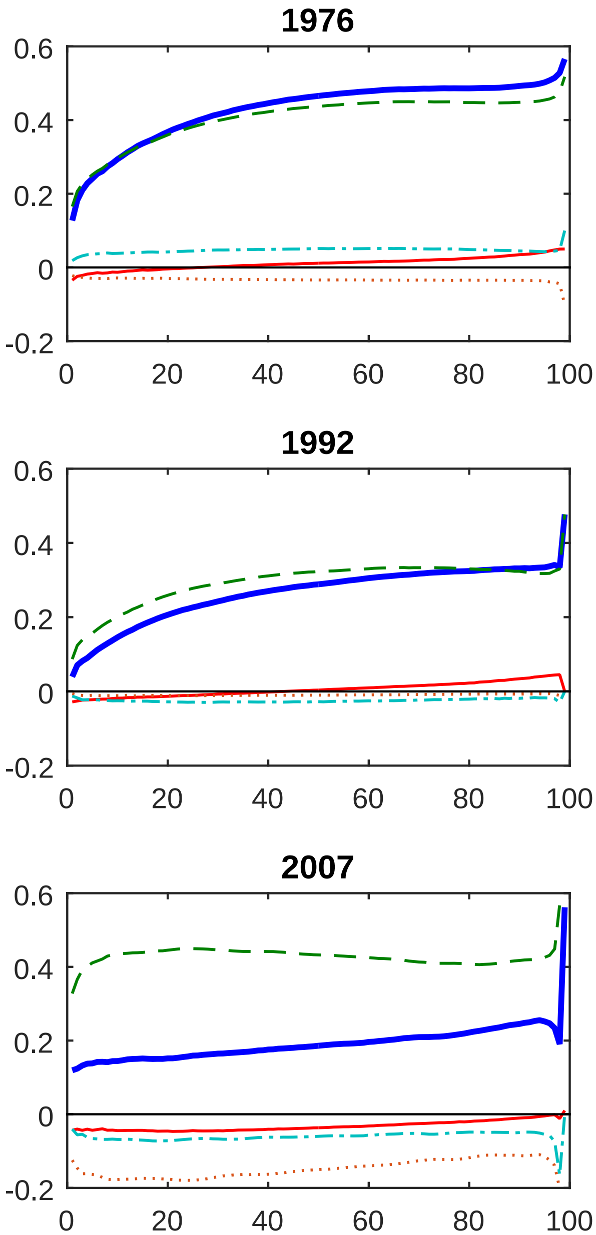
The distributional gap for the entire population has an unconventional shape, as it displays a thick spike for a large part of the distribution. Its width equals the difference in the participation rates between men and women, and its height equals earnings of male workers from the left tail of their gender distribution. Therefore, the width of this spike has progressively diminished over time with the reduction of the participation gap. However, it still remains the main factor of difference between the two distributions. Second, since the fraction of non-participants is positive for both men and women, the lower tail of the gap equals zero, as workers of both genders on that tail do not have any labor earnings.

Table 9: Mean decomposition, actual earnings for the full population (Frank copula)

Year	Total	EC	CC	SC	PC
1976	1.07***	0.00	0.29***	-0.02	0.81***
1977	1.06***	-0.01	0.28***	-0.01	0.79***
1978	1.03***	0.00	0.34***	-0.03**	0.73***
1979	1.03***	0.00	0.35***	-0.04***	0.72***
1980	0.96***	0.00	0.33***	-0.04**	0.67***
1981	0.91***	0.01	0.31***	-0.03**	0.63***
1982	0.87***	0.01	0.30***	-0.03	0.59***
1983	0.79***	0.01*	0.28***	-0.02	0.53***
1984	0.75***	0.01**	0.20***	0.02	0.52***
1985	0.75***	0.01**	0.26***	-0.02	0.50***
1986	0.76***	0.01**	0.28***	-0.04**	0.50***
1987	0.75***	0.02**	0.31***	-0.06***	0.48***
1988	0.71***	0.02**	0.24***	-0.01	0.46***
1989	0.69***	0.02***	0.22***	0.01	0.45***
1990	0.71***	0.01**	0.26***	-0.02	0.46***
1991	0.67***	0.00	0.21***	-0.01	0.46***
1992	0.62***	0.00	0.20***	-0.01	0.43***
1993	0.59***	0.00	0.15***	0.03*	0.41***
1994	0.59***	0.00	0.16***	0.02	0.41***
1995	0.61***	-0.01	0.18***	0.01	0.41***
1996	0.60***	-0.01	0.26***	-0.05***	0.40***
1997	0.58***	-0.01	0.24***	-0.04**	0.39***
1998	0.58***	-0.01**	0.19***	0.02	0.39***
1999	0.61***	0.00	0.28***	-0.05***	0.39***
2000	0.59***	0.00	0.24***	-0.04**	0.39***
2001	0.58***	-0.01*	0.24***	-0.05***	0.39***
2002	0.58***	-0.01***	0.23***	-0.03*	0.40***
2003	0.55***	-0.01***	0.26***	-0.06***	0.37***
2004	0.53***	-0.02***	0.25***	-0.06***	0.37***
2005	0.54***	-0.02***	0.26***	-0.07***	0.37***
2006	0.53***	-0.02***	0.28***	-0.09***	0.36***
2007	0.53***	-0.03***	0.29***	-0.08***	0.36***
2008	0.51***	-0.03***	0.30***	-0.10***	0.34***
2009	0.49***	-0.04***	0.22***	-0.03	0.34***
2010	0.43***	-0.04***	0.17***	0.00	0.30***
2011	0.43***	-0.04***	0.25***	-0.07***	0.29***
2012	0.45***	-0.04***	0.18***	0.00	0.31***
2013	0.46***	-0.05***	0.25***	-0.07***	0.33***

Notes: Total, EC, CC, SC and PC respectively denote total difference, endowments component, coefficients component, selection component and participation component; *, ** and ** respectively denote statistical significance at the 90%, 95% and 99% confidence level.

Figure 2: Unconditional quantiles decompositions, actual earnings for participants (Frank copula)



Notes: the solid thick blue line denotes the total gap between male and female workers; the solid thin red line denotes the endowments component; the dashed thin green line denotes the coefficients component; the dotted thin orange line denotes the selection component; the dashed-dotted thin cyan line denotes the participation component.

Additionally, the participation component has a decreasing shape after the end of the spike, reflecting a shifting in the distribution between male and female worker. To see this, denote by τ_f the quantile at which women earnings becomes positive, *i.e.*, $\tau_f = 1 - \mathbb{E}[\pi_f(Z)]$. Men above τ_f represent those above a certain level of earnings, which would be equivalent to comparing the distribution of earnings of employed women at a given quantile to the distribution of employed men of a higher quantile.

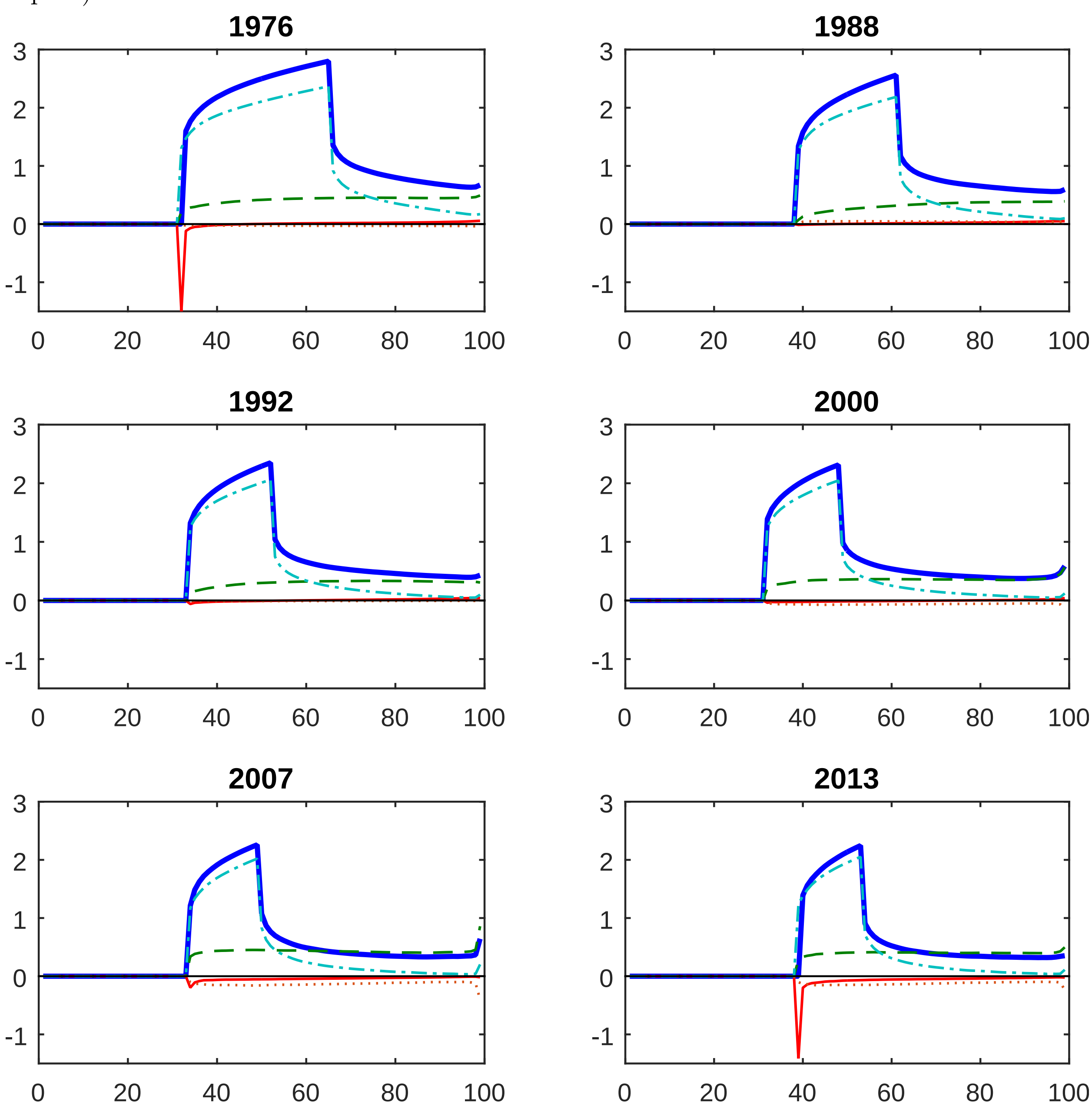
The other three components have a similar shape to the one found in the decomposition for participants, although their size is somewhat smaller. Consequently, the coefficients component is the second largest one, remaining an important determinant of the gap. Finally, the endowments component and selection components are flatter, relatively small, but they have become negative over time reducing the gap at all quantiles.

These results are related to the findings in Olivetti and Petrongolo (2008). In their cross-country comparison, they found that countries with the highest participation gaps tended to have lower wage gaps for participants. However, their estimated gaps accounting for self-selection showed that gaps increased much more in countries with larger participation gaps. This is captured by the participation component in the decomposition of the wage gap: for a given level of self-selection, narrowing the participation gap would make the participation component shrink to zero. However, this does not eliminate the impact of self-selection on unobservables, as differences in the amount of self-selection could be due to differences in the copula. Indeed, the estimates in this paper showed how, as participation gaps decreased over the considered period, the intensity of self-selection increased much more for women. Therefore, a more comprehensive cross-country analysis could account for this factor to explain how differences in self-selection intensity have determined different wage gaps in different countries.

6.5 Heterogeneous copulas

One potentially strong assumption regards the fact that the baseline copulas used in the estimation are homogeneous across the covariates. This limits any selection differences across

Figure 3: Unconditional quantiles decompositions, actual earnings for participants (Gaussian copula)



Notes: the solid thick blue line denotes the total gap between male and female workers; the solid thin red line denotes the endowments component; the dashed thin green line denotes the coefficients component; the dotted thin orange line denotes the selection component; the dashed-dotted thin cyan line denotes the participation component.

some of these characteristics to the propensity score channel. To address this issue, I repeat the estimation separately for three different categories: race (white vs non-white), education level (college graduates vs. less than college) and marital status (married vs unmarried).

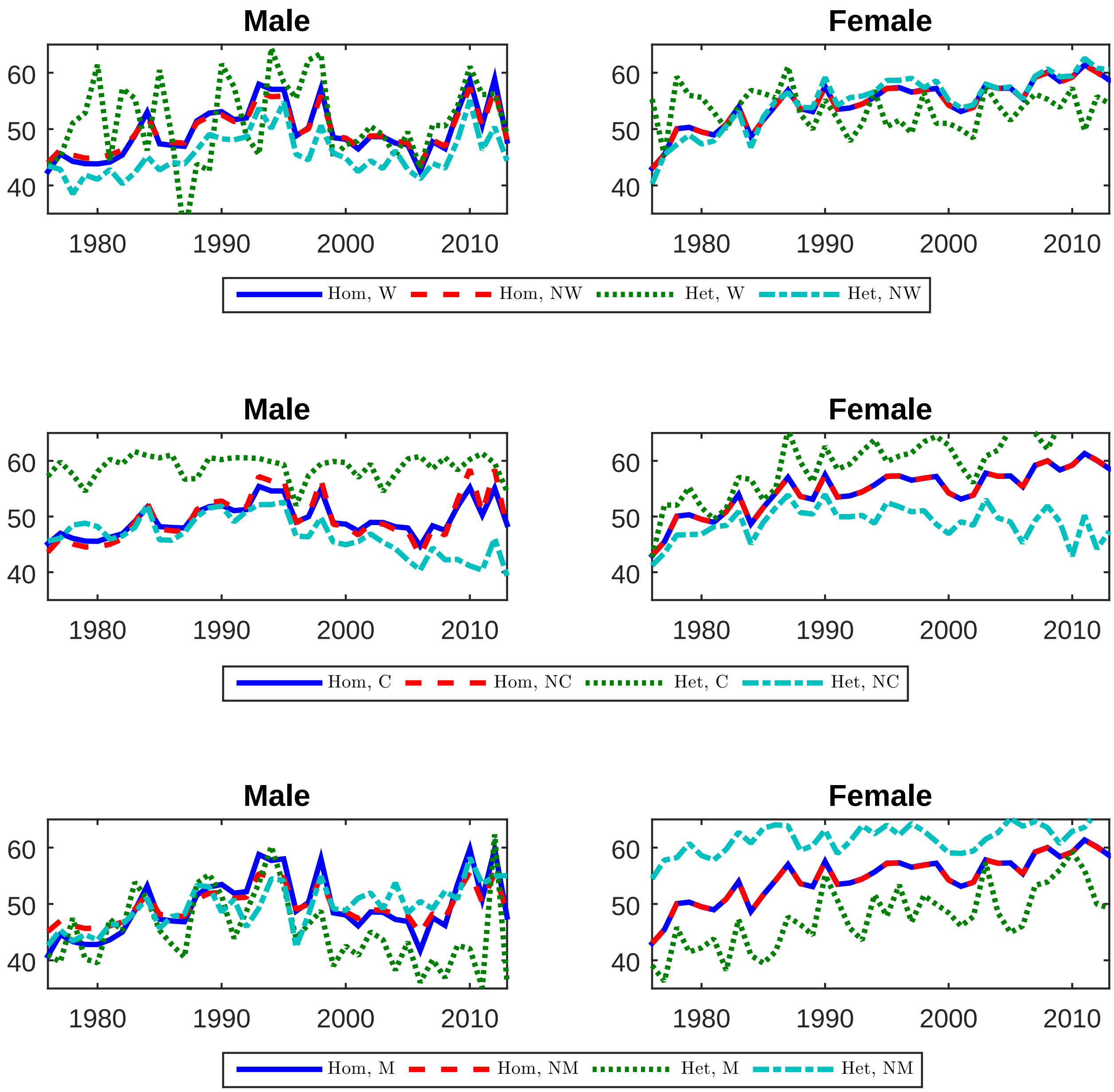
Some of the estimates are slightly sensitive to the heterogeneous copulas. The mean value of unobserved ability (Figure 4, top panel) is most similar when one obtains the estimates by race. However, it has been in general larger for white males relative to the baseline estimates, although the estimates are more volatile. In contrast, the estimates for non-white males show a smaller value than the baseline for almost the entire period. In contrast, the estimates for white females show a more stable evolution of the average value of unobserved ability, whereas for non-white female workers, the evolution has been positive, very closely to the baseline estimates.

A more evident difference relative to the baseline case appears in the estimates split by education level (Figure 4, central panel), showing a great divide in the average level of unobserved ability between those with a college degree and those without. For the first group, this level has been higher and relatively stable, whereas for the latter it has been lower and decreasing steadily since the mid-nineties. The same divide can be observed for female workers although the average level of unobserved ability steadily increased for college-educated women, and has remained stable for lower-educated female workers.

Finally, if one allows the copula to be different for married and unmarried workers (Figure 4, bottom panel), we find the largest differences for women: in particular, unmarried female workers tend to have a much higher level of unobserved ability, although the gap with married women has decreased in the last two decades. In contrast, the estimates for men are much similar to the baseline ones. The main difference corresponds to the period beginning in the late nineties, in which the average level of unobserved ability for married male workers is smaller.

Despite these differences in the amount of self-selection, the mean earnings gap for participants remains largely unaltered in these specifications (Figure 5). However, the decompositions do vary slightly. The most noticeable differences relative to the baseline

Figure 4: Mean value of u for participants



Notes: Hom, Het, W, NW, C, NC, M and NM respectively denote homogeneous copula, heterogeneous copula, white, non-white, college graduates, less than college, married and unmarried.

estimates arise in the coefficients and selection components. In particular, the coefficients components is generally larger for the estimates with heterogeneous copulas by race and marital status, and smaller for the estimates that are heterogeneous by education level. On the other hand, the coefficients and selection components almost cancel each other out entirely. This reinforces the hypothesis that the instrument is weak for male workers. The only exception is the model with heterogeneous copula by marital status, for which the the participation components is more negative during the entire period, *i.e.*, in favor of female workers.

7 Conclusion

In this paper I have introduced a new way to decompose differences in outcomes between two groups when there is self-selection into participation. In particular, rather than considering the decomposition of potential outcomes for the entire population, I decompose differences in actual outcomes for both participants and the entire population. These differences are decomposed into four components: endowments, coefficients, participation and selection. Moreover, I propose to provide two additional ancillary decompositions regarding differences in participation and self-selection.

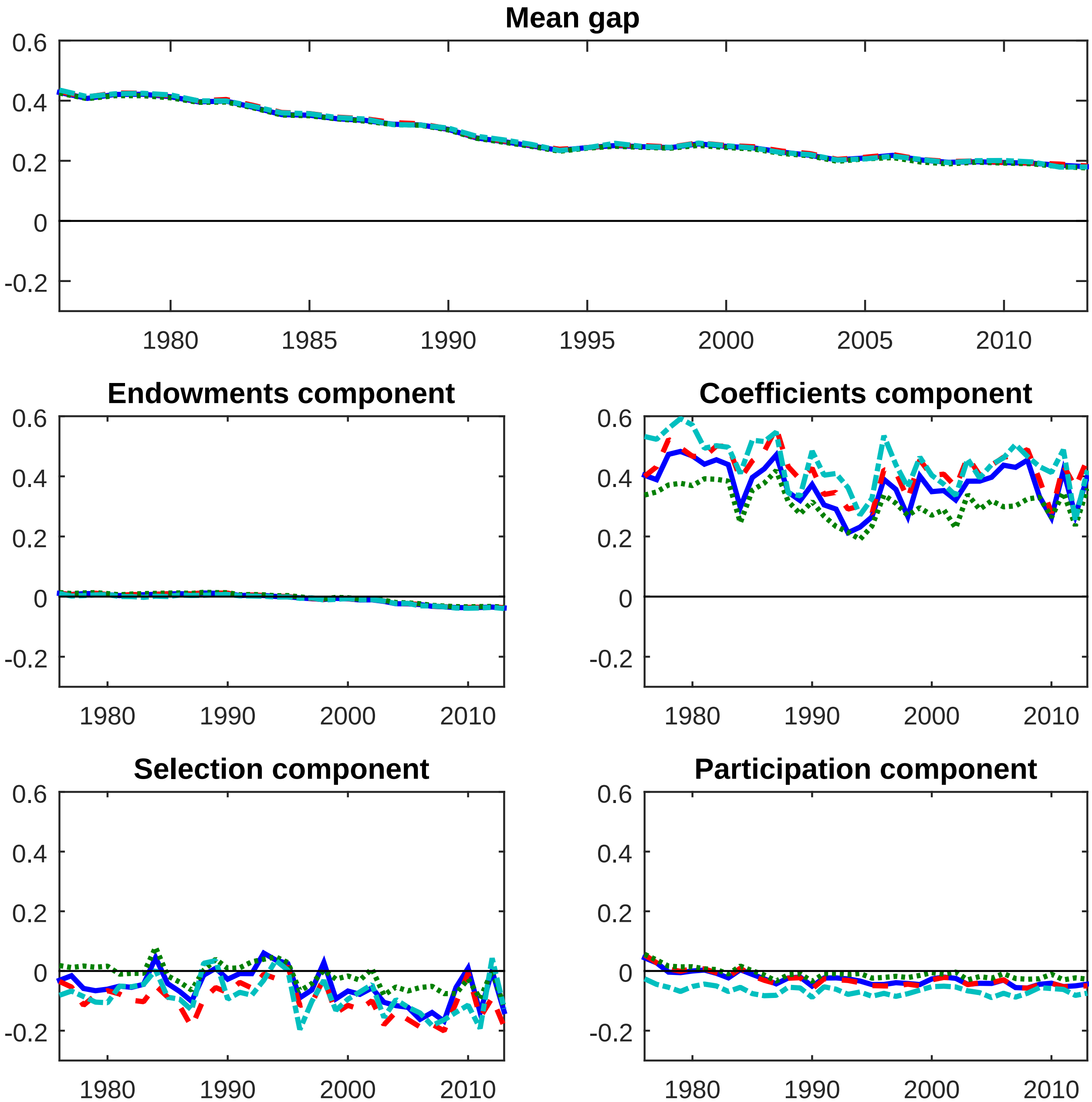
I apply this methodology to analyze the labor earnings gap between males and females. I find that increases in female labor market participation and improvements in self-selection that led to an increase in unobserved ability for females have been responsible for a large share of the fall of the gap. Moreover, considering the gap for participants greatly underestimates the gap for the entire population, due to the still existing gender participation gap.

References

Arellano, M. and S. Bonhomme (2017). Quantile selection models. *Econometrica 85*(1), 1–28.

Barsky, R., J. Bound, K. K. Charles, and J. P. Lupton (2002). Accounting for the black–white wealth gap: a nonparametric approach. *Journal of the American statistical Association 97*(459), 663–673.

Figure 5: Mean decomposition, actual earnings for participants with heterogeneous copulas



Notes: the solid blue line indicates the gap with the homogeneous copulas; the dashed red line indicates the gap with the heterogeneous copulas by race; the dotted green line indicates the gap with the heterogeneous copulas by education level; the dashed-dotted cyan line indicates the gap with the heterogeneous copulas by marital status.

Bauer, T., S. Göhlmann, and M. Sinning (2007). Gender differences in smoking behavior. *Health Economics 16*(9), 895–909.

Björklund, A. and R. Moffitt (1987). The estimation of wage gains and welfare gains in self-selection models. *The Review of Economics and Statistics 69*(1), 42–49.

Blinder, A. S. (1973). Wage discrimination: reduced form and structural estimates. *Journal of Human resources 8*(4), 436–455.

Borooah, V. K. and S. Iyer (2005). The decomposition of inter-group differences in a logit model: Extending the oaxaca-blinder approach with an application to school enrolment in india. *Journal of Economic and Social Measurement 30*(4), 279–293.

Bustamante, A. V., H. Fang, J. A. Rizzo, and A. N. Ortega (2009). Heterogeneity in health insurance coverage among us latino adults. *Journal of General Internal Medicine 24*(3), 561.

Card, D. (1996). The effect of unions on the structure of wages: A longitudinal analysis. *Econometrica: Journal of the Econometric Society 694*(4), 957–979.

Chernozhukov, V., I. Fernández-Val, and A. Galichon (2010). Quantile and probability curves without crossing. *Econometrica 78*(3), 1093–1125.

Chernozhukov, V., I. Fernández-Val, and S. Luo (2019). Distribution regression with sample selection, with an application to wage decompositions in the uk. *arXiv preprint*.

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

Chernozhukov, V., I. Fernández-Val, and B. Melly (2022). Fast algorithms for the quantile regression process. *Empirical economics 62*, 7–33.

Chernozhukov, V. and C. Hansen (2006). Instrumental quantile regression inference for structural and treatment effect models. *Journal of Econometrics 132*(2), 491–525.

Cukrowska-Torzewska, E. and A. Lovasz (2016). Are children driving the gender wage gap? comparative evidence from poland and hungary. *Economics of Transition 24*(2), 259–297.

Das, M., W. K. Newey, and F. Vella (2003). Nonparametric estimation of sample selection models. *The Review of Economic Studies 70*(1), 33–58.

Depalo, D. and S. Pereda-Fernández (2020). Consistent estimates of the public/private wage gap. *Empirical Economics 58*(6), 2937–2947.

DiNardo, J., N. Fortin, and T. Lemieux (1996). Labor market institutions and the distribution of wages, 1973-1992: A semiparametric approach. *Econometrica 64*(5), 1001–1044.

Fernández-Val, I., F. Peracchi, A. van Vuuren, F. Vella, et al. (2022). Selection and the us distribution of real female hourly wage. Technical report, arXiv. org.

Firpo, S. P., N. M. Fortin, and T. Lemieux (2018). Decomposing wage distributions using recentered influence function regressions. *Econometrics 6*(2), 28.

Fortin, N., T. Lemieux, and S. Firpo (2011). Decomposition methods in economics. In *Handbook of labor economics*, Volume 4, pp. 1–102. Elsevier.

Han, S. and E. J. Vytlacil (2017). Identification in a generalization of bivariate probit models with dummy endogenous regressors. *Journal of Econometrics 199*, 63–73.

Heckman, J. J. (1976). The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. *Annals of Economic and Social Measurement 5*(4), 475–492.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica: Journal of the econometric society 47*(1), 153–161.

Heckman, J. J. and E. Vytlacil (2005). Structural equations, treatment effects, and econometric policy evaluation. *Econometrica 73*(3), 669–738.

Heckman, J. J. and E. J. Vytlacil (2007). Econometric evaluation of social programs, part ii: Using the marginal treatment effect to organize alternative econometric estimators to evaluate social programs, and to forecast their effects in new environments. *Handbook of econometrics 6*, 4875–5143.

Huber, M. (2015). Causal pitfalls in the decomposition of wage gaps. *Journal of Business & Economic Statistics 33*(2), 179–191.

Huber, M., A. Solovyeva, et al. (2020). On the sensitivity of wage gap decompositions. *Journal of Labor Research 41*, 1–33.

Kato, K. (2011). A note on moment convergence of bootstrap m-estimators. *Statistics & Risk Modeling 28*(1), 51–61.

Koenker, R. and G. Bassett (1978). Regression quantiles. *Econometrica: journal of the Econometric Society 46*, 33–50.

Krieg, J. M. and P. Storer (2006). How much do students matter? applying the oaxaca decomposition to explain determinants of adequate yearly progress. *Contemporary Economic Policy 24*(4), 563–581.

Kunze, A. (2008). Gender wage gap studies: consistency and decomposition. *Empirical Economics 35*(1), 63–76.

Lee, L.-F. (1983). Generalized econometric models with selectivity. *Econometrica: Journal of the Econometric Society 51*(2), 507–512.

Lee, S. (2003). Efficient semiparametric estimation of a partially linear quantile regression model. *Econometric theory 19*(1), 1–31.

Leorato, S. and F. Peracchi (2015). Comparing distribution and quantile regression. Technical report, EIEF Working Paper.

Ma, S. and M. R. Kosorok (2005). Robust semiparametric m-estimation and the weighted bootstrap. *Journal of Multivariate Analysis 96*(1), 190–217.

Maasoumi, E. and L. Wang (2017). What can we learn about the racial gap in the presence of sample selection? *Journal of Econometrics 199*(2), 117–130.

Maasoumi, E. and L. Wang (2019). The gender gap between earnings distributions. *Journal of Political Economy 127*(5), 000–000.

Machado, J. A. and J. Mata (2005). Counterfactual decomposition of changes in wage distributions using quantile regression. *Journal of applied Econometrics 20*(4), 445–465.

McEwan, P. J. and J. H. Marshall (2004). Why does academic achievement vary across countries? evidence from cuba and mexico. *Education Economics 12*(3), 205–217.

Melly, B. (2006). Estimation of counterfactual distributions using quantile regression. *Review of Labor Economics 68*(4), 543–572.

Mogstad, M., A. Torgovitsky, and C. R. Walters (2021). The causal interpretation of two-stage least squares with multiple instrumental variables. *American Economic Review 111*(11), 3663–98.

Mora, R. (2008). A nonparametric decomposition of the mexican american average wage gap. *Journal of Applied Econometrics 23*(4), 463–485.

Neuman, S. and R. L. Oaxaca (2003). Gender versus ethnic wage differentials among professionals: Evidence from israel. *Annales d'Economie et de Statistique* (71-72), 267–292.

Neuman, S. and R. L. Oaxaca (2004). Wage decompositions with selectivity-corrected wage equations: A methodological note. *The Journal of Economic Inequality 2*(1), 3–10.

Newey, W. K. (2009). Two-step series estimation of sample selection models. *The Econometrics Journal 12*(s1), S217–S229.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International economic review 14*(3), 693–709.

Olivetti, C. and B. Petrongolo (2008). Unequal pay or unequal employment? a cross-country analysis of gender gaps. *Journal of Labor Economics 26*(4), 621–654.

Pereda-Fernández, S. (2022). Identification and estimation of triangular models with a binary treatment. *Journal of Econometrics*, 1–39.

Portnoy, S. and R. Koenker (1997). The gaussian hare and the laplacian tortoise: computability of squared-error versus absolute-error estimators. *Statistical Science 12*(4), 279–300.

Rothenberg, T. J. (1971). Identification in parametric models. *Econometrica: Journal of the Econometric Society 39*(3), 577–591.

Sohn, R. (2008). The gender math gap: Is it growing. Technical report, V mimeo, SUNY Albany.

van der Vaart, A. W. and J. A. Wellner (1996). *Weak Convergence and Empirical Processes With Applications to Statistics*. Springer.

Vella, F. (1998). Estimating models with sample selection bias: a survey. *Journal of Human Resources 33*(1), 127–169.

A Mathematical proofs

A.1 Proof of Lemma 1

For $d = 0, 1$, let $V = \tilde{F}_{V|d,X}\left(\tilde{V}|d, X\right)$. By definition, $V \sim U(0,1)$. Moreover, $\tilde{V} < \tilde{\pi}_d(Z) \Leftrightarrow \tilde{F}_{V|d,X}\left(\tilde{V}|d, X\right) < \tilde{F}_{V|d,X}\left(\tilde{\pi}_d(Z)|d, X\right) \equiv \pi_d(Z)$. Hence, $\tilde{F}_{V|S,X} = F_{V|S,X}$.

Similarly, let $U = \tilde{F}_{U|d,X}\left(\tilde{U}|d, X\right)$, which is also uniformly distributed on the unit interval. It follows that $Y = \tilde{g}_d\left(X, \tilde{U}\right) S = \tilde{g}_d\left(X, \tilde{F}^{-1}_{U|d,X}(U|d, X)\right) S \equiv g_d(X, U) S$. The joint distribution of $\left(\tilde{U}, \tilde{V}\right)$ can be written as

$$\begin{aligned}
\mathbb{P}\left(\tilde{U} \leq \tau, \tilde{V} \leq \tilde{\pi}_d(z)|Z = z\right) &= \mathbb{P}\left(\tilde{F}^{-1}_{U|d,x}(U|d, x) \leq \tau, \tilde{F}^{-1}_{V|d,x}(V|d, x) \leq \tilde{\pi}_d(z)|Z = z\right) \\
&= \mathbb{P}\left(U \leq \tilde{F}_{U|d,x}(\tau|x), V \leq \pi_d(z)|Z = z\right) \\
&= C_{d,x}\left(\tilde{F}_{U|d,x}(\tau|d, x), \pi_d(z)\right)
\end{aligned}$$

where the first equality follows by the invertibility of $\tilde{U}$ and $\tilde{V}$, the second one by the first result of the lemma, and the third one by definition of the copula.

Define $\tilde{G}_{d,x}(\tau, \pi_d(z)) \equiv \mathbb{P}\left(\tilde{U} \leq \tau|S = 1, Z = z\right)$. It can be expressed as

$$\begin{aligned}
\tilde{G}_{d,x}(\tau, \tilde{\pi}_d(z)) &\equiv \frac{\mathbb{P}\left(\tilde{U} \leq \tau, \tilde{V} \leq \tilde{\pi}_d(z)|Z = z\right)}{\mathbb{P}\left(\tilde{V} \leq \tilde{\pi}_d(z)|Z = z\right)} \\
&= \frac{\mathbb{P}\left(\tilde{F}^{-1}_{U|d,x}(U_1|d, x) \leq \tau, \tilde{F}^{-1}_{V|d,x}(V|d, x) \leq \tilde{\pi}_d(z)|Z = z\right)}{\mathbb{P}\left(\tilde{F}^{-1}_{V|d,x}(V|d, x) \leq \tilde{\pi}_d(z)\middle| Z = z\right)} \\
&= \frac{C_{d,x}\left(\tilde{F}_{U|d,x}(\tau|d, x), \pi_d(z)\right)}{\pi_d(z)} = G_{d,x}\left(\tilde{F}_{U|d,x}(\tau|x), \pi_d(z)\right) \qquad (21)
\end{aligned}$$

Then, the distribution of Y, conditional on $S=1$ and $Z=z$ equals

$$\begin{aligned}\mathbb{P}\left(Y\leq y|S=1,Z=z\right)&=\int\mathbf{1}\left(\tilde{g}_{d}\left(x,\tilde{u}\right)\leq y\right)d\tilde{G}_{d,x}\left(\tilde{u},\tilde{\pi}_{d}\left(z\right)\right)\\&=\int\mathbf{1}\left(\tilde{g}_{d}\left(x,\tilde{u}\right)\leq y\right)dG_{d,x}\left(\tilde{F}_{U|d,x}\left(\tilde{u}|x\right),\pi_{d}\left(z\right)\right)\\&=\int\mathbf{1}\left(\tilde{g}_{d}\left(x,\tilde{F}_{U|d,x}^{-1}\left(u\right)\right)\leq y\right)dG_{d,x}\left(u,\pi_{d}\left(z\right)\right)\\&=\int\mathbf{1}\left(g_{d}\left(x,u\right)\leq y\right)dG_{d,x}\left(u,\pi_{d}\left(z\right)\right)\end{aligned}$$

where the second equality follows by Equation 21, the third one by the invertibility of $\tilde{U}$, and the fourth one by the definition of $\tilde{g}_d$, completing the proof.

A.2 Proof of Theorem 1

By the functional delta method,

$$\begin{aligned}&\sqrt{n}\left(\hat{\mathbb{E}}\left[Y^{\ell}|S=1\right]-\mathbb{E}\left[Y^{\ell}|S=1\right]\right)\\&\quad=\sqrt{n}\int_{\mathcal{Z}}\int_{\varepsilon}^{1-\varepsilon}\left(\hat{g}_{k}\left(x,\tau\right)-g_{k}\left(x,\tau\right)\right)dG_{l,x}\left(\tau|\pi_{m}\left(z\right)\right)sdF_{Z}^{h}\left(z\right)\\&\quad+\sqrt{n}\int_{\mathcal{Z}}\int_{\varepsilon}^{1-\varepsilon}g_{k}\left(x,\tau\right)d\left(\hat{G}_{l,x}\left(\tau|\pi_{m}\left(z\right)\right)-G_{l,x}\left(\tau|\pi_{m}\left(z\right)\right)\right)sdF_{Z}^{h}\left(z\right)\\&\quad+\sqrt{n}\int_{\mathcal{Z}}\int_{\varepsilon}^{1-\varepsilon}g_{k}\left(x,\tau\right)d\left(\nabla_{\pi}G_{l,x}\left(\tau|\pi_{m}\left(z\right)\right)\right)\left(\hat{\pi}_{m}\left(z\right)-\pi_{m}\left(z\right)\right)sdF_{Z}^{h}\left(z\right)\\&\quad+\sqrt{n}\int_{\mathcal{Z}}\int_{\varepsilon}^{1-\varepsilon}g_{k}\left(x,\tau\right)dG_{l,x}\left(\tau|\pi_{m}\left(z\right)\right)sd\left(\hat{F}_{Z}^{h}\left(z\right)-F_{Z}^{h}\left(z\right)\right)+o_{P}\left(1\right)\\&\quad\Rightarrow\int_{\mathcal{Z}}\int_{\varepsilon}^{1-\varepsilon}\mathbb{Z}_{g_{k}}\left(\tau,x\right)dG_{l,x}\left(\tau|\pi_{m}\left(z\right)\right)dF_{Z}^{h}\left(z\right)\\&\quad+\int_{\mathcal{Z}}\int_{\varepsilon}^{1-\varepsilon}g_{k}\left(x,\tau\right)\frac{1}{\pi_{m}\left(z\right)}\mathbb{Z}_{c_{l}}\left(\tau,\pi_{m}\left(z\right)\right)d\tau dF_{Z}^{h}\left(z\right)\\&\quad+\int_{\mathcal{Z}}\int_{\varepsilon}^{1-\varepsilon}g_{k}\left(x,\tau\right)d\left(\nabla_{\pi}G_{l,x}\left(\tau|\pi_{m}\left(z\right)\right)\right)\mathbb{Z}_{\pi_{m}}\left(z\right)dF_{Z}^{h}\left(z\right)\\&\quad+\sqrt{p_{h}}\mathbb{Z}_{Z_{h}}\left(\int_{\varepsilon}^{1-\varepsilon}g_{k}\left(\cdot,\tau\right)dG_{l,x}\left(\tau|\pi\left(\cdot\right)\right)\right)\equiv\mathbb{Z}_{Y^{\ell}|S=1}\end{aligned}$$

where $\nabla_\pi G_{d,x}(\tau|\pi)$ is the Hadamard derivative of $G_{d,x}(\tau|\pi)$ with respect to π. Apply the functional delta method once more to obtain

$$\sqrt{n}\Delta^{\ell,\ell'}\left(\hat{\mathbb{E}}[Y|S=1]-\mathbb{E}[Y|S=1]\right)\Rightarrow\mathbb{Z}_{Y^\ell|S=1}-\mathbb{Z}_{Y^{\ell'}|S=1}\equiv\mathbb{Z}_{\Delta^{\ell\ell'}Y|S=1}$$

To show the asymptotic distribution of $\Delta^{\ell,\ell'}\hat{Q}_{Y|S=1}(\tau)$, I firstly show some intermediate steps. First, note that we can write

$$\begin{aligned}
&\hat{F}^\ell_{Y|S=1}(y|z)=\frac{1}{n_h}\sum_{i=1}^{n}\left[\varepsilon+\int_\varepsilon^{1-\varepsilon}\mathbf{1}\left(\hat{g}_k(x_i,u)\le y\right)d\hat{G}_{l,x}\left(u,\hat{\pi}_m(z_i)\right)\right]\mathbf{1}(d_i=h)\\
&=\int_{\mathcal{Z}}\varepsilon+\int_\varepsilon^{1-\varepsilon}\mathbf{1}\left(x'\hat{\beta}_k(u)\le y\right)\frac{\hat{c}_{l,x}\left(u,\hat{\pi}_m(z)\right)}{\hat{\pi}_m(z)}du\,d\hat{F}^h_Z(z)\equiv\int_{\mathcal{Z}}\hat{F}^\ell_{Y|Z,S=1}(y|z)\,d\hat{F}_{Z_m}(z)
\end{aligned}$$

Next, consider the joint asymptotic distribution of $\left(\hat{F}^\ell_{Y|Z,S=1}(y|z),\int_{\mathcal{Z}}f\,d\hat{F}_{Z_m}(z)\right)$. The following expansion holds:

$$\begin{aligned}
&\sqrt{n}\left(\hat{F}^\ell_{Y|Z,S=1}(y|z)-F^\ell_{Y|Z,S=1}(y|z)\right)\\
&\quad=\sqrt{n}\int_\varepsilon^{1-\varepsilon}\left(\mathbf{1}\left(x'\hat{\beta}_k(\tau)\le y\right)-\mathbf{1}\left(x'\beta_k(\tau)\le y\right)\right)\frac{c_{l,x}\left(\tau,\pi_m(z)\right)}{\pi_m(z)}d\tau\\
&\quad+\sqrt{n}\int_\varepsilon^{1-\varepsilon}\mathbf{1}\left(x'\beta_k(\tau)\le y\right)\frac{1}{\pi_m(z)}\left(\hat{c}_{l,x}\left(\tau,\pi_m(z)\right)-c_{l,x}\left(\tau,\pi_m(z)\right)\right)d\tau\\
&\quad+\sqrt{n}\int_\varepsilon^{1-\varepsilon}\mathbf{1}\left(x'\beta_k(\tau)\le y\right)\frac{\nabla_\pi c_{l,x}\left(u,\pi_m(z)\right)\pi_m(z)-c_{l,x}\left(\tau,\pi_m(z)\right)}{\pi_m(z)^2}\left(\hat{\pi}_m(z)-\pi_m(z)\right)d\tau+o_P(1)
\end{aligned}$$

where $\nabla_\pi c_{d,x}(u,\pi)$ denotes the Hadamard derivative of $c_{d,x}(u,\pi)$ with respect to its second argument. Define $\tilde{u}_d(x,y)\equiv\{\min u: x'\beta_d(u)\le y\}$. By Lemma 3, the mapping $\nu:\mathbb{D}_\nu\subset\ell^\infty(\mathcal{U})^{d_x}\to\ell^\infty(\mathcal{YX}\Pi)$, defined as $b\to\varepsilon+\int_\varepsilon^{1-\varepsilon}\mathbf{1}(x'b(u)\le y)\,dG_{d,x}(u,\pi)$ is Hadamard differentiable at $b(\cdot)=\beta_d(\cdot)$ tangentially to $C(\mathcal{U})^{d_x}$, with derivative

$$D_h(y|x)=-f_{Y|Z,S=1}(y|z)\frac{1}{\pi_d(z)}c_{d,x}\left(C_{d,x}\left(\tilde{u}_d(x,y),\pi_d(z)\right),\pi_d(z)\right)x'h\left(F_{Y|Z,S=1}(y|z)\,|z\right)$$

Hence, it follows that

$$
\begin{aligned}
&\sqrt{n}\left(\hat{F}^{\ell}_{Y|Z,S=1}\left(y|z\right)-F^{\ell}_{Y|Z,S=1}\left(y|z\right)\right)\Rightarrow \\
&\quad -f_{Y|Z,S=1}\left(y|z\right)\frac{1}{\pi_m\left(z\right)}c_{l,x}\left(C_{l,x}\left(\tilde{u}_k\left(x,y\right),\pi_m\left(z\right)\right),\pi_m\left(z\right)\right)x'\mathbb{Z}_{\beta_k}\left(F_{Y|Z,S=1}\left(y|z\right)\right) \\
&\quad +\int_{\varepsilon}^{1-\varepsilon}\mathbf{1}\left(x'\beta_k\left(\tau\right)\leq y\right)\frac{1}{\pi_m\left(z\right)}\mathbb{Z}_{c_l}\left(\tau,\pi_m\left(z\right)\right)d\tau \\
&\quad +\int_{\varepsilon}^{1-\varepsilon}\mathbf{1}\left(x'\beta_k\left(\tau\right)\leq y\right)\frac{\pi_m\left(z\right)\nabla_{\pi}c_{l,x}\left(u,\pi_m\left(z\right)\right)-c_{l,x}\left(\tau,\pi_m\left(z\right)\right)}{\pi_m\left(z\right)^2}\mathbb{Z}_{\pi_m}\left(z\right)d\tau \\
&\quad \equiv\mathbb{Z}_{F^{\ell}_{Y}|Z,S=1}\left(y,z\right)
\end{aligned}
$$

in $\ell^{\infty}\left(\mathcal{YZ}\right)$. Therefore,

$$
\sqrt{n}\begin{pmatrix}\hat{F}^{\ell}_{Y|Z,S=1}\left(y|z\right)-F^{\ell}_{Y|Z,S=1}\left(y|z\right)\\ \int_{\mathcal{Z}}fd\left(\hat{F}^{h}_{Z}\left(z\right)-F^{h}_{Z}\left(z\right)\right)\end{pmatrix}\Rightarrow\begin{pmatrix}\mathbb{Z}_{F^{\ell}_{Y}|X,S=1}\left(y,z\right)\\ \sqrt{p_h}\mathbb{Z}_{Z_h}\left(f\right)\end{pmatrix}
$$

in $\ell^{\infty}\left(\mathcal{YZF}\right)$.

The next step is to show the asymptotic distribution of the estimator of the unconditional distribution $\hat{F}^{\ell}_{Y|S=1}\left(y\right)$. By the functional delta method,

$$
\begin{aligned}
\sqrt{n}\left(\hat{F}^{\ell}_{Y|S=1}\left(y\right)-F^{\ell}_{Y|S=1}\left(y\right)\right)&=\sqrt{n}\int_{\mathcal{Z}}\left(\hat{F}^{\ell}_{Y|Z,S=1}\left(y|z\right)-F^{\ell}_{Y|Z,S=1}\left(y|z\right)\right)dF^{h}_{Z}\left(z\right) \\
&\quad+\sqrt{n}\int_{\mathcal{Z}}F^{\ell}_{Y|Z,S=1}\left(y|z\right)d\left(\hat{F}^{h}_{Z}\left(z\right)-F^{h}_{Z}\left(z\right)\right) \\
&\Rightarrow\int_{\mathcal{Z}}\mathbb{Z}_{F^{\ell}_{Y}|Z,S=1}\left(y|z\right)dF^{h}_{Z}\left(z\right)+\sqrt{p_h}\mathbb{Z}_{Z_h}\left(F^{\ell}_{Y|Z,S=1}\left(y|z\right)\right) \\
&\equiv\mathbb{Z}_{F^{\ell}_{Y|S=1}}\left(y\right)
\end{aligned}
$$

uniformly in $\ell^{\infty}\left(\mathcal{Y}\right)$.

The next step is to show the asymptotic distribution of $\hat{Q}^{\ell}_{Y|S=1}\left(\tau\right)$. By the functional delta method,

$$\begin{aligned}\sqrt{n}\left(\hat{Q}^{\ell}_{Y|S=1}(\tau)-Q^{\ell}_{Y|S=1}(\tau)\right) &= -\frac{\sqrt{n}\left(\hat{F}^{\ell}_{Y|S=1}\left(Q^{\ell}_{Y|S=1}(\tau)\right)-F^{\ell}_{Y|S=1}\left(Q^{\ell}_{Y|S=1}(\tau)\right)\right)}{f^{\ell}_{Y|S=1}\left(Q^{\ell}_{Y|S=1}(\tau)\right)}+o_P(1)\\ &\Rightarrow -\frac{\mathbb{Z}_{F^{\ell}_{Y|S=1}}\left(Q^{\ell}_{Y|S=1}(\tau)\right)}{f^{\ell}_{Y|S=1}\left(Q^{\ell}_{Y|S=1}(\tau)\right)}\\ &\equiv \mathbb{Z}_{Q^{\ell}_{Y|S=1}}(\tau)\end{aligned}$$

jointly in $\ell \in \mathcal{D}^4$, where I have used the Hadamard differentiability of the quantile operator (Chernozhukov et al., 2010). $\tau \to Q^{\ell}_{Y|S=1}(\tau)$ is a.s. uniformly continuous by Assumption 4, and together with the a.s. uniform continuity of $\mathbb{Z}_{F^{\ell}_{Y|S=1}}(y)$, it follows that $\mathbb{Z}_{Q^{\ell}_{Y|S=1}}(\tau)$ is a.s. uniformly continuous with respect to τ.

Finally, note that

$$\Delta^{\ell,\ell'}\left(\hat{Q}_{Y|S=1}(\tau)-Q_{Y|S=1}(\tau)\right)=\hat{Q}^{\ell}_{Y|S=1}(\tau)-Q^{\ell}_{Y|S=1}(\tau)-\left(\hat{Q}^{\ell'}_{Y|S=1}(\tau)-Q^{\ell'}_{Y|S=1}(\tau)\right)$$

Therefore, $\sqrt{n}\Delta^{\ell,\ell'}\left(\hat{Q}_{Y|S=1}(\tau)-Q_{Y|S=1}(\tau)\right)\Rightarrow \mathbb{Z}_{Q^{\ell}_{Y|S=1}}(\tau)-\mathbb{Z}_{Q^{\ell'}_{Y|S=1}}(\tau)\equiv \mathbb{Z}_{Q|S=1,\ell\ell'}(\tau)$, finishing the proof.

A.3 Proof of Theorem 2

Let $W \equiv (Y, S, D, Z)$. Moreover, define

$$r_d(W,\beta,\theta,\gamma,\tau)\equiv\begin{bmatrix}\mathbf{1}(D=d)\,SX\zeta_{G_{d,x}(\tau,\pi(Z;\gamma),\theta_d)}(Y-X'\beta_d)\\ \int_0^1\mathbf{1}(D=d)\,S\varphi(u,Z)\,\zeta_{G_{d,x}(\tau,\pi(Z;\gamma),\theta_d)}(Y-X'\beta_d)\,du\\ s_d(S,Z;\gamma)\end{bmatrix}$$

$$q_d(W,\beta,\theta,\gamma,\tau) \equiv \begin{bmatrix} \mathbf{1}(D=d)\, SX\rho_{G_{d,x}(\tau,\pi(Z;\gamma),\theta_d)}(Y-X'\beta_d) \\ \mathbf{1}(D=d)\, S\int_0^1 \varphi(u,Z)\,\rho_{G_{d,x}(\tau,\pi(Z;\gamma),\theta_d)}(Y-X'\beta_d)\,du \\ s_d(S,Z;\gamma) \end{bmatrix}$$

$f \mapsto \mathbb{E}_n[f(W)] \equiv \frac{1}{n}\sum_{i=1}^n f(W)$, $f \mapsto \mathbb{G}_n[f(W)] \equiv \frac{1}{\sqrt{n}}\sum_{i=1}^n f(W) - \mathbb{E}(f(W))$, $Q_{d,n}(\beta,\theta,\gamma,\tau) \equiv \mathbb{E}_n[q_d(W,\beta,\theta,\gamma,\tau)]$, and $Q_d(\beta,\theta,\gamma,\tau) \equiv \mathbb{E}[q_d(W,\beta,\theta,\gamma,\tau)]$, where $\rho_\tau(u) \equiv (\tau - \mathbf{1}(u<0))\,u$, $\zeta_\tau(u) \equiv (\mathbf{1}(u<0)-\tau)$, $\epsilon_d(\tau) \equiv Y - X'\beta_d(\tau)$, and $\hat{\epsilon}_d(\tau) \equiv Y - X'\hat{\beta}_d(\tau)$.

First I show the consistency of the estimator. By Assumptions 3 to 6, $Q_d(\beta,\theta,\gamma,\tau)$ is continuous over $\mathcal{B}\times\Theta\times\Gamma\times\mathcal{T}$. By Lemma 6, $\sup_{(\beta,\theta,\gamma,\tau)\in\mathcal{B}\times\Theta\times\Gamma\times\mathcal{T}} \|Q_{d,n}(\beta,\theta,\gamma,\tau) - Q_d(\beta,\theta,\gamma,\tau)\| \xrightarrow{P} 0$, uniformly in $\mathcal{D}$. Thus, by Lemma 5, $\sup_{\tau\in\mathcal{T}} \left\|\hat{\vartheta}_d(\tau) - \vartheta_d(\tau)\right\| \xrightarrow{P} 0$, uniformly in $\mathcal{D}$.

Next, I show its asymptotic distribution. By Theorem 3 in Koenker and Bassett (1978), it is possible to show that

$$O\left(\frac{1}{\sqrt{n}}\right) = \sqrt{n}\mathbb{E}_n\left[\mathbf{1}(D=d)\, SX\zeta_{G_{d,x}\left(\tau,\pi(Z;\hat{\gamma}),\hat{\theta}_d\right)}(\hat{\epsilon}_d(\tau))\right]$$

By Lemma 6 and Assumption 8, the following expansion holds in $\ell^\infty(\mathcal{T})$:

$$\begin{aligned} O\left(\frac{1}{\sqrt{n}}\right) &= \mathbb{G}_n\left[\mathbf{1}(D=d)\, SX\zeta_{G_{d,x}\left(\tau,\pi(Z;\hat{\gamma});\hat{\theta}_d\right)}(\hat{\epsilon}_d(\tau))\right] + \sqrt{n}\mathbb{E}\left[\mathbf{1}(D=d)\, SX\zeta_{G_{d,x}\left(\tau,\pi(Z;\hat{\gamma});\hat{\theta}_d\right)}(\hat{\epsilon}_d(\tau))\right] \\ &= \mathbb{G}_n\left[\mathbf{1}(D=d)\, SX\zeta_{G_{d,x}(\tau,\pi(Z;\gamma);\theta_d)}(\epsilon_d(\tau))\right] + o_P(1) \\ &+ \sqrt{n}\mathbb{E}\left[\mathbf{1}(D=d)\, SX\zeta_{G_{d,x}\left(\tau,\pi(Z;\hat{\gamma});\hat{\theta}_d\right)}(\hat{\epsilon}_d(\tau))\right] \\ &= \mathbb{G}_n\left[\mathbf{1}(D=d)\, SX\zeta_{G_{d,x}(\tau,\pi(Z;\gamma);\theta_d)}(\epsilon_d(\tau))\right] + J_{\beta d}(\tau)\sqrt{n}\left(\hat{\beta}_d(\tau) - \beta_d(\tau)\right) \\ &- J_{\gamma d}(\tau)\sqrt{n}(\hat{\gamma}-\gamma) - J_{\theta d}(\tau)\sqrt{n}\left(\hat{\theta}_d - \theta_d\right) + o_P(1) \end{aligned}$$

where

$$J_{\beta d}(\tau) \equiv \frac{\partial\mathbb{E}\left[\mathbf{1}(D=d)\, SX\zeta_{G_{d,x}(\tau,\pi(Z;\gamma);\theta_d)}(\epsilon_d(\tau))\right]}{\partial\beta_d}$$

$$J_{\gamma d}(\tau) \equiv -\frac{\partial \mathbb{E}\left[\mathbf{1}(D=d) SX\zeta_{G_{d,x}(\tau,\pi(Z;\gamma);\theta_d)}(\epsilon_d(\tau))\right]}{\partial \gamma}$$

$$J_{\theta d}(\tau) \equiv -\frac{\partial \mathbb{E}\left[\mathbf{1}(D=d) SX\zeta_{G_{d,x}(\tau,\pi(Z;\gamma);\theta_d)}(\epsilon_d(\tau))\right]}{\partial \theta_d}$$

Rearranging and solving for $\sqrt{n}\left(\hat{\beta}_d(\tau)-\beta_d(\tau)\right)$,

$$\begin{aligned}\sqrt{n}\left(\hat{\beta}_d(\tau)-\beta_d(\tau)\right) = &-J_{\beta d}(\tau)^{-1}\Big\{\mathbb{G}_n\left[\mathbf{1}(D=d) SX\zeta_{G_{d,x}(\tau,\pi(Z;\gamma);\theta_d)}(\epsilon_d(\tau))\right] \\ &-J_{\gamma d}(\tau)\sqrt{n}(\hat{\gamma}-\gamma) - J_{\theta d}(\tau)\sqrt{n}\left(\hat{\theta}_d-\theta_d\right)\Big\} + o_P(1) \end{aligned} \tag{22}$$

in $\ell^\infty(\mathcal{T})$.

Using Theorem 3 in Koenker and Bassett (1978) again, it is possible to show that

$$O\left(\frac{1}{\sqrt{n}}\right) = \sqrt{n}\mathbb{E}_n\left[\int_\varepsilon^{1-\varepsilon}\mathbf{1}(D=d) S\varphi(u,Z)\,\zeta_{G_{d,x}\left(u,\pi(Z;\hat{\gamma});\hat{\theta}_d\right)}(\hat{\epsilon}_d(u))\,du\right]$$

By Lemma 6 and Assumption 8, the following expansion holds:

$$\begin{aligned} O\left(\frac{1}{\sqrt{n}}\right) &= \mathbb{G}_n\left[\int_\varepsilon^{1-\varepsilon}\mathbf{1}(D=d) S\varphi(u,Z)\,\zeta_{G_{d,x}\left(u,\pi(Z;\hat{\gamma});\hat{\theta}_d\right)}(\hat{\epsilon}_d(u))\,du\right] \\ &+\sqrt{n}\int_\varepsilon^{1-\varepsilon}\mathbb{E}\left[\mathbf{1}(D=d) S\varphi(u,Z)\,\zeta_{G_{d,x}\left(u,\pi(Z;\hat{\gamma});\hat{\theta}_d\right)}(\hat{\epsilon}_d(u))\right]du \\ &= \mathbb{G}_n\left[\int_\varepsilon^{1-\varepsilon}\mathbf{1}(D=d) S\varphi(u,Z)\,\zeta_{G_{d,x}(u,\pi(Z;\gamma);\theta_d)}(\epsilon_d(u))\,du\right] + o_P(1) \\ &+\sqrt{n}\int_\varepsilon^{1-\varepsilon}\mathbb{E}\left[\mathbf{1}(D=d) S\varphi(u,Z)\,\zeta_{G_{d,x}\left(u,\pi(Z;\hat{\gamma});\hat{\theta}_d\right)}(\hat{\epsilon}_d(u))\right]du \\ &= \mathbb{G}_n\left[\int_\varepsilon^{1-\varepsilon}\mathbf{1}(D=d) S\varphi(u,Z)\,\zeta_{G_{d,x}(u,\pi(Z;\gamma);\theta_d)}(\epsilon_d(u))\,du\right] \\ &+\sqrt{n}\int_\varepsilon^{1-\varepsilon}\tilde{J}_{\beta d}(u)\left(\hat{\beta}_d(u)-\beta_d(u)\right)du - \sqrt{n}\int_\varepsilon^{1-\varepsilon}\tilde{J}_{\theta d}(u)\,du\left(\hat{\theta}_d-\theta_d\right) \\ &-\sqrt{n}\int_\varepsilon^{1-\varepsilon}\tilde{J}_{\gamma d}(u)\,du\,(\hat{\gamma}-\gamma) + o_P(1) \end{aligned}$$

where

$$\tilde{J}_{\beta d}(\tau) \equiv \frac{\partial \mathbb{E}\left[\mathbf{1}(D=d)\, S\varphi(\tau, Z)\, \zeta_{G_{d,x}(\tau,\pi(Z;\gamma);\theta_d)}\left(\epsilon_d(\tau)\right)\right]}{\partial \beta_d}$$

$$\tilde{J}_{\gamma d}(\tau) \equiv -\frac{\partial \mathbb{E}\left[\mathbf{1}(D=d)\, S\varphi(\tau, Z)\, \zeta_{G_{d,x}(\tau,\pi(Z;\gamma);\theta_d)}\left(\epsilon_d(\tau)\right)\right]}{\partial \gamma}$$

$$\tilde{J}_{\theta d}(\tau) \equiv -\frac{\partial \mathbb{E}\left[\mathbf{1}(D=d)\, S\varphi(\tau, Z)\, \zeta_{G_{d,x}(\tau,\pi(Z;\gamma);\theta_d)}\left(\epsilon_d(\tau)\right)\right]}{\partial \theta_d}$$

Rearranging and solving for $\sqrt{n}\left(\hat{\theta}_d - \theta_d\right)$,

$$\begin{aligned}\sqrt{n}\left(\hat{\theta}_d - \theta_d\right) = \left[\int_{\varepsilon}^{1-\varepsilon} \tilde{J}_{\theta d}(u)\, du\right]^{-1} \Bigg\{ \mathbb{G}_n \left[\int_{\varepsilon}^{1-\varepsilon} \mathbf{1}(D=d)\, S\varphi(u, Z)\, \zeta_{G_{d,x}(u,\pi(Z;\gamma);\theta_d)}\left(\epsilon_d(u)\right) du\right] \\ + \sqrt{n} \int_{\varepsilon}^{1-\varepsilon} \tilde{J}_{\beta d}(u) \left(\hat{\beta}_d(u) - \beta_d(u)\right) du - \sqrt{n} \int_{\varepsilon}^{1-\varepsilon} \tilde{J}_{\gamma d}(u)\, du\, (\hat{\gamma} - \gamma) \Bigg\} + o_P(1)\end{aligned} \tag{23}$$

Now define

$$A_d(\tau) \equiv \hat{\vartheta}_d(\tau) - \vartheta_d(\tau)$$

$$C_d(\tau) \equiv \begin{bmatrix} -J_{\beta d}(\tau)^{-1} & 0 & 0 \\ 0 & \left[\int_{\varepsilon}^{1-\varepsilon} \tilde{J}_{\theta d}(u)\, du\right]^{-1} & 0 \\ 0 & 0 & -B_d^{-1} \end{bmatrix}$$

$$D_d(\tau) \equiv \begin{bmatrix} 0 & 0 & 0 \\ \left[\int_{\varepsilon}^{1-\varepsilon} \tilde{J}_{\theta d}(u)\, du\right]^{-1} \tilde{J}_{\beta d}(\tau) & 0 & -\left[\int_{\varepsilon}^{1-\varepsilon} \tilde{J}_{\theta d}(u)\, du\right]^{-1} \tilde{J}_{\gamma d}(\tau) \\ 0 & 0 & 0 \end{bmatrix}$$

$$F_d(\tau) \equiv \begin{bmatrix} 0 & J_{\beta d}(\tau)^{-1} J_{\theta d}(\tau) & J_{\beta d}(\tau)^{-1} J_{\gamma d}(\tau) \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\psi_d(\tau) \equiv r_d(W, \beta(\tau), \theta, \gamma, \tau)$$

Combining Equations 22 to 23 yields

$$A_d(\tau) = F_d(\tau) A_d(\tau) + \int_{\varepsilon}^{1-\varepsilon} D_d(u) A_d(u)\, du + C_d(\tau) \frac{1}{\sqrt{n}} \mathbb{G}_n \psi_d(\tau) + o_P\left(\frac{1}{\sqrt{n}}\right) \tag{24}$$

in $\ell^\infty(\mathcal{T})$. Equation 24 is a particular case of a Fredholm integral equation of the second kind. The solution to this type of equations is a Liouville-Neumann series. By Lemma 4, the solution to this equation is given by:

$$\begin{aligned} \sqrt{n} A_d(\tau) = {} & F_d^I(\tau) \left(I - \int_{\varepsilon}^{1-\varepsilon} D_d(u) F_d^I(u)\, du\right)^{-1} \int_{\varepsilon}^{1-\varepsilon} D_d(u) F_d^I(u) C_d(u) \mathbb{G}_n \psi_d(u)\, du \\ & + F_d^I(\tau) C_d(\tau) \mathbb{G}_n \psi_d(\tau) + o_P(1) \end{aligned} \tag{25}$$

in $\ell^\infty(\mathcal{T})$, where $F_d^I(\tau) \equiv (I - F_d(\tau))^{-1} = I + F_d(\tau)$. Using the Functional Delta Method and Lemmas 2 and 5, it follows that $\sqrt{n}\left(\hat{\vartheta}_d(\tau) - \vartheta_d(\tau)\right) \Rightarrow \sqrt{p_d} H_d(\tau) \mathbb{Z}_{R_d}(\tau) \equiv \mathbb{Z}_{\vartheta_d}(\tau)$, a zero-mean Gaussian process with covariance $\Sigma_{\vartheta_d}(\tau, \tau') = \sqrt{p_d p_{d'}} H_d(\tau) \Sigma_{R_d}(\tau, \tau') H_d(\tau)$.

A.4 Proof of Corollary 1

The first step is to show the asymptotic distribution of the first three components of $\left(\hat{g}_k(x,\tau), \hat{c}_{l,x}(\tau|\pi), \hat{\pi}_m(z), \int_{\mathcal{Z}} f d\hat{F}_Z^h\right)$. By Theorem 2 and the functional delta method,

$$\sqrt{n}\left(\left(\hat{g}_k(x,\tau), \hat{c}_{l,x}(\tau,\eta), \hat{\pi}_m(z)\right) - \left(g_k(x,\tau), c_{l,x}(\tau,\eta), \pi_m(z)\right)\right) \Rightarrow J_{\upsilon_\ell}(z,\tau,\eta)\,\mathbb{Z}_{\vartheta_\ell}(\tau)$$

where

$$J_{\upsilon_\ell}(z,\tau,\eta) = \begin{pmatrix} x' & 0 & 0 \\ 0 & \nabla_{\theta_l} c_{l,x}(\tau,\eta;\theta_l) & 0 \\ 0 & 0 & \nabla_{\gamma_m}\pi_m(z;\gamma_m) \end{pmatrix}$$

and I have also used the fact that the mapping $b \to x'b(u)$ is linear and therefore Hadamard differentiable at $b(\cdot) = \beta_d(\cdot)$, where the derivative equals $D_h(\tau,x) = x'h(\tau,x)$. By Lemma E.4 in Chernozhukov et al. (2013), $\sqrt{n_h}\int_{\mathcal{Z}} f d\left(\hat{F}_Z^h - F_Z^h\right) \Rightarrow \mathbb{Z}_{Z_h}(f)$.[27] Taking these together yields $\sqrt{n}\left(\hat{\upsilon}_\ell(z,\tau,\eta,f) - \upsilon_\ell(z,\tau,\eta,f)\right) \Rightarrow \mathbb{Z}_{\upsilon_\ell}(z,\tau,\eta,f)$, where

$$\mathbb{Z}_{\upsilon_\ell}(z,\tau,\eta,f) \equiv \begin{pmatrix} \mathbb{Z}_{g_k}(\tau,x) \\ \mathbb{Z}_{c_l}(\tau,\pi_m(z)) \\ \mathbb{Z}_{\pi_m}(z) \\ \sqrt{p_h}\mathbb{Z}_{Z_h}(f) \end{pmatrix} = \begin{pmatrix} J_{\upsilon_\ell}(z,\tau,\eta)\,\mathbb{Z}_{\vartheta_\ell}(\tau) \\ \sqrt{p_h}\mathbb{Z}_{Z_h}(f) \end{pmatrix}$$

A.5 Proof of Theorem 3

First, I show the distribution of $\vartheta_d^*(\tau)$. Using the same arguments used in Theorem 2 and Assumption 10, it follows that

$$\sqrt{n}\left(\hat{\vartheta}_d^*(\tau) - \vartheta_d(\tau)\right) = F_d^I(\tau)\left(I - \int_\varepsilon^{1-\varepsilon} D_d(u)\,F_d^I(u)\,du\right)^{-1}\int_\varepsilon^{1-\varepsilon} D_d(u)\,F_d^I(u)\,C_d(u)\,\mathbb{G}_n^*\psi_d(u)\,du$$
$$+ F_d^I(\tau)\,C_d(\tau)\,\mathbb{G}_n^*\psi_d(\tau) + o_P(1) \quad (26)$$

[27] Recall that $\hat{F}_Z^h(z) = \frac{1}{n_h}\sum_{i=1}^n \mathbf{1}(z_i \le z)\,\mathbf{1}(d_i = h)$, so that $\sqrt{n}\int_{\mathcal{Z}} f d\left(\hat{F}_Z^h - F_Z^h\right) \Rightarrow \sqrt{p_h}\mathbb{Z}_{Z_h}(f)$.

where $f \mapsto \mathbb{G}_n^* [f(W)] \equiv \frac{1}{\sqrt{n}} \sum_{i=1}^n w_i f(W) - \mathbb{E}(f(W))$. Therefore, $\sqrt{n}\left(\hat{\vartheta}_d^* - \vartheta_d\right) \Rightarrow \mathbb{Z}_{\vartheta_d}^* (\tau) \equiv \sqrt{\omega_0}\mathbb{Z}_{\vartheta_d}(\tau)$, a zero-mean Gaussian process with covariance $\omega_0 \Sigma_{\vartheta_d}(\tau, \tau')$.

Now substract Equation 25 from Equation 26 to get

$$
\begin{aligned}
&\sqrt{n}\left(\hat{\vartheta}_d^*(\tau) - \hat{\vartheta}_d(\tau)\right) \\
&= F_d^I(\tau)\left(I - \int_\varepsilon^{1-\varepsilon} D_d(u) F_d^I(u)\, du\right)^{-1} \int_\varepsilon^{1-\varepsilon} D_d(u) F_d^I(u) C_d(u) \frac{1}{\sqrt{n}} \sum_i^n (w_i - 1)\psi_d(u)\, du \\
&+ F_d^I(\tau) C_d(\tau) \frac{1}{\sqrt{n}} \sum_i^n (w_i - 1)\psi_d(\tau) + o_P(1)
\end{aligned} \tag{27}
$$

By Assumption 10, it follows that $\sqrt{\frac{n}{\omega_0}}\left(\hat{\vartheta}_d^*(\tau) - \hat{\vartheta}_d(\tau)\right) \Rightarrow \mathbb{Z}_{\vartheta_d}(\tau)$. By the functional delta method, Theorem 1, and the previous result, it is straightforward to show that:

$$
\begin{aligned}
\sqrt{\frac{n}{\omega_0}}\left(\Delta^{\ell,\ell'}\hat{\mathbb{E}}^*[Y|S=1] - \Delta^{\ell,\ell'}\mathbb{E}[Y|S=1]\right) &\Rightarrow \mathbb{Z}_{\Delta^{\ell\ell'}Y|S=1} \\
\sqrt{\frac{n}{\omega_0}}\left(\Delta^{\ell,\ell'}\hat{\mathbb{E}}^*[Y|S=1] - \Delta^{\ell,\ell'}\hat{\mathbb{E}}[Y|S=1]\right) &\Rightarrow \mathbb{Z}_{\Delta^{\ell\ell'}Y|S=1} \\
\sqrt{\frac{n}{\omega_0}}\left(\Delta^{\ell,\ell'}\hat{Q}_{Y|S=1}^*(\tau) - \Delta^{\ell,\ell'}Q_{Y|S=1}(\tau)\right) &\Rightarrow \mathbb{Z}_{Q|S=1,\ell\ell'}(\tau) \\
\sqrt{\frac{n}{\omega_0}}\left(\Delta^{\ell,\ell'}\hat{Q}_{Y|S=1}^*(\tau) - \Delta^{\ell,\ell'}\hat{Q}_{Y|S=1}(\tau)\right) &\Rightarrow \mathbb{Z}_{Q|S=1,\ell\ell'}(\tau)
\end{aligned}
$$

A.6 Proof of Theorem 4

By the functional delta method,

$$
\begin{aligned}
\sqrt{n}\left(\hat{\mathbb{E}}\left[\pi^\ell|S=1\right] - \mathbb{E}\left[\pi^\ell|S=1\right]\right) &= \sqrt{n}\int_{\mathcal{Z}} \left(\hat{\pi}_m(z) - \pi_m(z)\right) dF_Z^h(z) + o_P(1) \\
&\Rightarrow \int_{\mathcal{Z}} \mathbb{Z}_{\pi_m}(z)\, dF_Z^h(z)
\end{aligned}
$$

Using the functional delta method once more, it follows that

$$
\sqrt{n}\Delta^{\ell,\ell'}\left(\hat{\mathbb{E}}[\pi|S=1] - \mathbb{E}[\pi|S=1]\right) \Rightarrow \mathbb{Z}_{\pi^\ell|S=1} - \mathbb{Z}_{\pi^{\ell'}|S=1} \equiv \mathbb{Z}_{\Delta^{\ell\ell'}\pi|S=1}
$$

finishing the proof of the first part of the Theorem.

For the second part, again by the functional delta method,

$$\begin{aligned}
&\sqrt{n}\left(\hat{\mathbb{E}}\left[U^{\ell}|S=1\right]-\mathbb{E}\left[U^{\ell}|S=1\right]\right)\\
&\quad=\sqrt{n}\int_{\mathcal{Z}}\int_{\varepsilon}^{1-\varepsilon}\tau d\left(\hat{G}_{l,x}\left(\tau|\pi_{m}\left(z\right)\right)-G_{l,x}\left(\tau|\pi_{m}\left(z\right)\right)\right)dF_{Z}^{h}\left(z\right)\\
&\quad+\sqrt{n}\int_{\mathcal{Z}}\int_{\varepsilon}^{1-\varepsilon}\tau d\left(\nabla_{\pi}G_{l,x}\left(\tau|\pi_{m}\left(z\right)\right)\right)\left(\hat{\pi}_{m}\left(z\right)-\pi_{m}\left(z\right)\right)dF_{Z}^{h}\left(z\right)\\
&\quad+\sqrt{n}\int_{\mathcal{Z}}\int_{\varepsilon}^{1-\varepsilon}\tau dG_{l,x}\left(\tau|\pi_{m}\left(z\right)\right)d\left(\hat{F}_{Z}^{h}\left(z\right)-F_{Z}^{h}\left(z\right)\right)+o_{P}\left(1\right)\\
&\quad\Rightarrow\int_{\mathcal{Z}}\int_{\varepsilon}^{1-\varepsilon}\tau\frac{1}{\pi_{m}\left(z\right)}\mathbb{Z}_{c_{l}}\left(\tau,\pi_{m}\left(z\right)\right)d\tau dF_{Z}^{h}\left(z\right)\\
&\quad+\int_{\mathcal{Z}}\int_{\varepsilon}^{1-\varepsilon}\tau d\left(\nabla_{\pi}G_{l,x}\left(\tau|\pi_{m}\left(z\right)\right)\right)\mathbb{Z}_{\pi_{m}}\left(z\right)dF_{Z}^{h}\left(z\right)\\
&\quad+\sqrt{p_{h}}\mathbb{Z}_{Z_{h}}\left(\int_{\varepsilon}^{1-\varepsilon}\tau dG_{l,x}\left(\tau|\pi\left(\cdot\right)\right)\right)\equiv\mathbb{Z}_{U^{\ell}|S=1}
\end{aligned}$$

Applying the functional delta method again,

$$\sqrt{n}\Delta^{\ell,\ell'}\left(\hat{\mathbb{E}}\left[U|S=1\right]-\mathbb{E}\left[U|S=1\right]\right)\Rightarrow\mathbb{Z}_{U^{\ell}|S=1}-\mathbb{Z}_{U^{\ell'}|S=1}\equiv\mathbb{Z}_{\Delta^{\ell\ell'}U|S=1}$$

which finishes the proof.

B Auxiliary Lemmas

B.1 Hadamard Derivatives

Lemma 2. *Define* $\kappa\left(u\right)\equiv\int_{0}^{1}\lambda\left(u\right)\nu\left(u\right)du$ *and* $\kappa\left(u;h_{t}\right)\equiv\int_{0}^{1}\lambda\left(u\right)\left(\nu\left(u\right)+th_{t}\left(u\right)\right)du$*. As* $t\rightarrow0$*,*

$$D_{h_{t}}\left(t\right)=\frac{1}{t}\left(\kappa\left(u;h_{t}\right)-\kappa\left(u\right)\right)\rightarrow D_{h}$$

where $D_h \equiv \int_0^1 \lambda(u) h(u) du$. *The convergence holds uniformly in any compact subset of* $\mathcal{T}$ *for any* $h_t : \|h_t - h\|_\infty \to 0$, *where* $h_t \in \ell^\infty(\mathcal{T})$ *and* $h \in C(\mathcal{T})$.

Proof.

$$
\begin{aligned}
D_{h_t}(h_t) &= \frac{\int_0^1 \lambda(u)(\nu(u) + th_t(u)) du - \int_0^1 \lambda(u)\nu(u) du}{t} \\
&= \frac{1}{t}\int_0^1 \lambda(u) th_t(u) du \to D_h
\end{aligned}
$$

□

Lemma 3. *Define* $F_{Y|Z,S=1}(y|z, h_t) \equiv \int_0^1 \mathbf{1}(g_d(x,u) + th_t(u|z) \leq y) \frac{1}{\pi(z)} c_{d,x}(u, \pi(z)) du$. *Under Assumption 6, as* $t \to 0$,

$$
D_{h_t}(y|z, h_t) = \frac{1}{t}\left[F_{Y|Z,S=1}(y|z, h_t) - F_{Y|Z,S=1}(y|z)\right] \to D_h(y|z)
$$

where $D_h(y|z) \equiv -f_{Y|Z,S=1}(y|z) \frac{1}{\pi_d(z)} c_{d,x}\left(C_{d,x}\left(g_d^{-1}(x,y), \pi_d(z)\right), \pi_d(z)\right) h\left(F_{Y|Z,S=1}(y|z)|z\right)$. *The convergence holds uniformly in any compact subset of* $\mathcal{UZ}$ *for any* $h_t : \|h_t - h\|_\infty \to 0$, $h_t \in \ell^\infty(\mathcal{UZ})$ *and* $h \in C(\mathcal{UZ})$.

Proof. This proof is partly based on the proof to Proposition 2 in Chernozhukov et al. (2010).

For any $\delta > 0 \exists \epsilon > 0$: for $u \in B_\epsilon\left(F_{Y|Z,S=1}(y|z)\right)$ and for small enough $t \geq 0$

$$
\mathbf{1}(g_d(x,u) + th_t(u|z) \leq y) \leq \mathbf{1}\left(g_d(x,u) + t\left(h\left(F_{Y|Z,S=1}(y|z)|z\right) - \delta\right) \leq y\right)
$$

whereas $\forall u \notin B_\epsilon\left(F_{Y|Z,S=1}(y|z)\right)$

$$
\mathbf{1}(g_d(x,u) + th_t(u|z) \leq y) \leq \mathbf{1}(g_d(x,u) \leq y)
$$

Therefore, for small enough $t \geq 0$

$$\frac{1}{t}\left[\int_0^1 \mathbf{1}\left(g_d\left(x,u\right)+th_t\left(u|z\right)\leq y\right)\frac{c_{d,x}\left(u,\pi_d\left(z\right)\right)}{\pi_d\left(z\right)}du-\int_0^1 \mathbf{1}\left(g_d\left(x,u\right)\leq y\right)\frac{c_{d,x}\left(u,\pi_d\left(z\right)\right)}{\pi_d\left(z\right)}du\right] \tag{28}$$

$$\leq \frac{1}{t}\int_{B_\epsilon\left(F_{Y|Z,S=1}(y|z)\right)}\left[\mathbf{1}\left(g_d\left(x,u\right)+th_t\left(u|z\right)\leq y\right)-\mathbf{1}\left(g_d\left(x,u\right)\leq y\right)\right]\frac{c_{d,x}\left(u,\pi_d\left(z\right)\right)}{\pi_d\left(z\right)}du \tag{29}$$

which by the change of variable $\tilde{y}=g_d\left(x,C_{d,x}^{-1}\left(u,\pi_d\left(z\right)\right)\right)$, where $C_{d,x}^{-1}\left(u,\pi_d\left(z\right)\right)$ denotes the inverse of $C_{d,x}\left(u,\pi_d\left(z\right)\right)$ with respect to its first argument, is equal to

$$\frac{1}{t}\int_{J\cap\left[y,y-t\left(h\left(F_{Y|Z,S=1}(y|z)|z\right)-\delta\right)\right]}\frac{c_{d,x}\left(C_{d,x}\left(g_d^{-1}\left(x,\tilde{y}\right),\pi_d\left(z\right)\right),\pi_d\left(z\right)\right)}{\left|\nabla_u g_d\left(x,C_{d,x}\left(g_d^{-1}\left(x,\tilde{y}\right),\pi_d\left(z\right),\right)\right)\right|\pi_d\left(z\right)}d\tilde{y}$$

where $g_d^{-1}\left(x,y\right)$ denotes the inverse of $g_d\left(x,u\right)$ with respect to its second argument and J is the image of $B_\epsilon\left(F_{Y|Z,S=1}\left(y|z\right)\right)$ under $u\mapsto g_d\left(x,C_{d,x}^{-1}\left(u,\pi_d\left(z\right)\right)\right)$. The change of variables is possible because $g_d\left(x,C_{d,x}^{-1}\left(u,\pi_d\left(z\right)\right)\right)$ is a bijection between $B_\epsilon\left(F_{Y|Z,S=1}\left(y|z\right)\right)$ and J.

Fix $\epsilon>0$ for $t\to 0$. Then, we have that $J\cap\left[y,y-t\left(h\left(F_{Y|Z,S=1}\left(y|z\right)|z\right)-\delta\right)\right]=\left[y,y-t\left(h\left(F_{Y|Z,S=1}\left(y|z\right)|z\right)-\delta\right)\right]$ and

$$\begin{aligned}&\frac{c_{d,x}\left(C_{d,x}\left(g_d^{-1}\left(x,\tilde{y}\right),\pi_d\left(z\right)\right),\pi_d\left(z\right)\right)}{\left|\nabla_u g_d\left(x,C_{d,x}\left(g_d^{-1}\left(x,\tilde{y}\right),\pi_d\left(z\right)\right)\right)\right|\pi_d\left(z\right)}\\&\qquad\to\frac{c_{d,x}\left(F_{Y|Z,S=1}\left(y|z\right),\pi_d\left(z\right)\right)}{\left|\nabla_u g_d\left(x,F_{Y|Z,S=1}\left(y|z\right)\right)\right|\pi_d\left(z\right)}\end{aligned}$$

as $C_{d,x}\left(g_d^{-1}\left(x,\tilde{y}\right),\pi_d\left(z\right)\right)\to F_{Y|Z,S=1}\left(y|z\right)$. Thus, Equation 29 is no greater than

$$\frac{\left[-h\left(F_{Y|Z,S=1}\left(y|z\right)|z\right)-\delta\right]c_{d,x}\left(F_{Y|Z,S=1}\left(y|z\right),\pi_d\left(z\right)\right)}{\left|\nabla_u g_d\left(x,F_{Y|Z,S=1}\left(y|z\right)\right)\right|\pi_d\left(z\right)}+o\left(1\right)$$

By a similar argument,

$$\frac{\left[-h\left(F_{Y|Z,S=1}\left(y|z\right)|z\right)+\delta\right]c_{d,x}\left(F_{Y|Z,S=1}\left(y|z\right),\pi_{d}\left(z\right)\right)}{\left|\nabla_{u}g_{d}\left(x,F_{Y|Z,S=1}\left(y|z\right)\right)\right|\pi_{d}\left(z\right)})+o\left(1\right)$$

bounds Equation 28 from below. Since $\delta>0$ can be made arbitrarily small, the desired result follows.

To show that the result holds uniformly in $(y,z)\in K$, a compact subset of $\mathcal{YZ}$, we use Lemma B.4 in Chernozhukov et al. (2013). Take a sequence (y_t,z_t) in K that converges to $(y,z)\in K$. Then, the preceding argument applies to this sequence, since the function $\frac{c_{d,x}\left(F_{Y|Z,S=1}(y|z),\pi_d(z)\right)}{\left|\nabla_u g_d\left(x,F_{Y|Z,S=1}(y|z)\right)\right|\pi_d(z)}$ is uniformly continuous on K. This result follows by the assumed continuity of $h\left(u|x\right)$, $F_{Y|Z,S=1}\left(y|z\right)$, $\left|\nabla_{u}g_{d}\left(x,F_{Y|Z,S=1}\left(y|z\right)\right)\right|$ in both their arguments, of $c_{d,x}\left(u,\pi_{d}\left(z\right)\right)$, $C_{d,x}\left(u,\pi_{d}\left(z\right)\right)$ and $g_{d}\left(x,u\right)$ in u, as well as the compactness of K.

□

B.2 Solution to the Fredholm Integral Equation

Lemma 4. *Let $L\left(\tau\right)=M_{1}\left(\tau\right)L\left(\tau\right)+M_{2}\left(\tau\right)+\int_{0}^{1}M_{3}\left(u\right)L\left(u\right)du$ be a Fredholm integral equation of the second kind. Moreover, define $\tilde{M}_{2}\left(\tau\right)\equiv\left(I-M_{1}\left(\tau\right)\right)^{-1}M_{2}\left(\tau\right)$ and $\tilde{M}_{3}\left(\tau\right)\equiv M_{3}\left(\tau\right)\left(I-M_{1}\left(\tau\right)\right)^{-1}$. Let*

(i) $I-M_{1}\left(\tau\right)$ is invertible $\forall\tau\in[0,1]$

(ii) $\lim_{n\to\infty}\left[\int_{0}^{1}\tilde{M}_{3}\left(u\right)du\right]^{n}=0$

Under (i)-(ii), the solution to this equation is given by

$$L\left(\tau\right)=\tilde{M}_{2}\left(\tau\right)+\left(I-M_{1}\left(\tau\right)\right)^{-1}\left(I-\int_{0}^{1}\tilde{M}_{3}\left(u\right)du\right)^{-1}\int_{0}^{1}\tilde{M}_{3}\left(u\right)M_{2}\left(u\right)du$$

Proof.

$$\begin{aligned}
L(\tau) &= M_1(\tau) L(\tau) + M_2(\tau) + \int_0^1 M_3(u) L(u) du \\
&= \tilde{M}_2(\tau) + (I - M_1(\tau))^{-1} \int_0^1 M_3(u) L(u) du \\
&= \tilde{M}_2(\tau) + (I - M_1(\tau))^{-1} \sum_{n=0}^{\infty} \left[\int_0^1 \tilde{M}_3(u) du\right]^n \int_0^1 \tilde{M}_3(u) M_2(u) du \\
&+ \lim_{n\to\infty} (I - M_1(\tau))^{-1} \left[\int_0^1 \tilde{M}_3(u) du\right]^n \int_0^1 M_3(u) L(u) du \\
&= \tilde{M}_2(\tau) + (I - M_1(\tau))^{-1} \left(I - \int_0^1 \tilde{M}_3(u) du\right)^{-1} \int_0^1 \tilde{M}_3(u) M_2(u) du
\end{aligned}$$

where the second equality follows by (i), the third one by iteratively substituting $L(u)$ inside the integral, and the fourth one by (ii) and the following result: define $S \equiv \sum_{n=0}^{\infty} C^n$, and A, B and C be square matrices. Then, $ASB - ACSB = A(I - C)SB = AB$. If $I - C$ is invertible, then $S = (I - C)^{-1}$. Premultiply both sides of the equation by A and postmultiply them by B to obtain the desired result. □

B.3 Argmax Process

Lemma 5. *(Chernozhukov and Hansen, 2006) Suppose that uniformly in π in a compact set Π and for a compact set K (i) $Z_n(\pi)$ is s.t. $Q_n(Z_n(\pi)|\pi) \geq \sup_{z\in K} Q_n(z|\pi) - \epsilon_n$, $\epsilon \searrow 0$; $Z_n(\pi) \in K$ wp $\to 1$, (ii) $Z_\infty(\pi) \equiv \arg\sup_{z\in K} Q_\infty(z|\pi)$ is a uniquely defined continuous process in $\ell^\infty(\Pi)$, (iii) $Q_n(\tau|\tau) \xrightarrow{p} Q_\infty(\tau|\tau)$ in $\ell^\infty(K \times \Pi)$, where $Q_\infty(\tau|\tau)$ is continuous. Then $Z_n(\tau) = Z_\infty(\tau) + o_P(1)$ in $\ell^\infty(\Pi)$*

Proof. See Chernozhukov and Hansen (2006). □

B.4 Stochastic Expansion

Lemma 6. *Under Assumptions 1-8, the following statements hold uniformly over $d \in \mathcal{D}$:*

1. $\sup_{(\beta,\theta,\gamma,\tau)\in\mathcal{B}\times\Theta\times\Gamma\times\mathcal{T}} |\mathbb{E}_n[q_d(W,\beta,\theta,\gamma,\tau)] - \mathbb{E}[q_d(W,\beta,\theta,\gamma,\tau)]| = o_P(1)$

2. $\mathbb{G}_n r_d(W, \beta(\tau), \theta, \gamma, \tau) \Rightarrow \mathbb{Z}_{R_d}(\tau)$ *in* $\ell^\infty(\mathcal{T})$*, where* $\mathbb{Z}_{R_d}(\tau)$ *is a zero-mean Gaussian process with covariance* $\Sigma_{R_d}(\tau, \tau')$ *defined below in the proof. Moreover, for any* $\hat{\vartheta}_d(\tau)$ *such that* $\sup_{(\tau,d)\in\mathcal{T}\times\mathcal{D}} \left\|\hat{\vartheta}_d(\tau) - \vartheta_d(\tau)\right\| = o_P(1)$*, the following holds:*

$$\sup_{\tau\in\mathcal{T}} \left\|\mathbb{G}_n r_d\left(W, \hat{\beta}(\tau), \hat{\theta}, \hat{\gamma}, \tau\right) - \mathbb{G}_n r_d(W, \beta(\tau), \theta, \gamma, \tau)\right\| = o_P(1)$$

Proof. Let $\mathcal{F}$ be the class of uniformly smooth functions in z with the uniform smoothness order $\omega > \frac{\dim(s,z)}{2}$ and $\|f(\tau', z) - f(\tau, z)\| < \overline{K}(\tau - \tau')^a$ for $\overline{K} > 0$, $a > 0$, $\forall (z, \tau, \tau') \forall f \in \mathcal{F}$. The bracketing number of $\mathcal{F}$, by Corollary 2.7.4 in van der Vaart and Wellner (1996) satisfies

$$\log N_{[\cdot]}(\epsilon, \mathcal{F}, L_2(P)) = O\left(\epsilon^{-\frac{\dim(z)}{\omega}}\right) = O\left(\epsilon^{-2-\delta}\right)$$

for some $\delta < 0$. Therefore, $\mathcal{F}$ is Donsker with a constant envelope. By Corollary 2.7.4, the bracketing number of

$$\mathcal{D}_d \equiv \{\beta_d \mapsto X'\beta_d, \beta_d \in \mathcal{B}_d\}$$

satisfies

$$\log N_{[\cdot]}(\epsilon, \mathcal{D}_d, L_2(P)) = O\left(\epsilon^{-\frac{\dim(s,x)}{\omega}}\right) = O\left(\epsilon^{-2-\delta'}\right)$$

for some $\delta' < 0$ and $d = 0, 1$. Since the indicator function is bounded and monotone, and the density functions $f_{Y|d,Z}(y|z)$ are bounded by Assumption 4, the bracketing number of

$$\mathcal{E}_d \equiv \{\beta_d \mapsto \mathbf{1}(Y < X'\beta_d), \beta_d \in \mathcal{B}_d\}$$

satisfies

$$\log N_{[\cdot]}(\epsilon, \mathcal{E}_d, L_2(P)) = O\left(\epsilon^{-2-\delta'}\right)$$

Since $\mathcal{E}_d$ has a constant envelope, it is Donsker. Now consider the function $G_{d,x}$. By Assumptions 6 and 9, the mean value theorem can be applied to show

$$\left\|G_{d,x}\left(\tau,\pi\left(z,\gamma\right);\theta_d\right)-G_{d,x}\left(\tau',\pi\left(z,\gamma\right);\theta_d\right)\right\|=\left\|\tau-\tau'\right\|\left\|\frac{\partial}{\partial\tau}G_{d,x}\left(\tau'',\pi\left(z,\gamma\right);\theta_d\right)\right\|$$

for some τ'' between τ and τ'. By Assumptions 6 and 9, the second term is bounded $\forall z,\tau'',d$, so it follows that $G_{d,x}\in\mathcal{F}$.[28] Let $\mathcal{T}\equiv\{\tau\mapsto\tau\}$ and define

$$\mathcal{H}_d\equiv\{h=(\beta,\theta,\gamma,\tau)\mapsto r_d\left(W,\beta,\theta,\gamma,\tau\right),(\beta,\theta,\gamma)\in\mathcal{B}\times\Theta\times\Gamma\}$$

The first subvector of $\mathcal{H}_d$ is $\mathcal{E}_d\times\mathcal{F}-\mathcal{T}\times\mathcal{F}$, the second subvector is $\mathcal{E}_d\times\mathcal{F}-\mathcal{T}\times\mathcal{F}$, and the third subvector is $\mathcal{F}$. Since $\mathcal{H}_d$ is Lipschitz over $(\mathcal{T},\mathcal{F},\mathcal{E}_d)$, it follows that it is Donsker by Theorem 2.10.6 in van der Vaart and Wellner (1996). Define

$$h_d\equiv(\beta,\theta,\gamma,\tau)\mapsto\mathbb{G}_n r_d\left(W,\beta,\theta,\gamma,\tau\right)$$

h_d is Donsker in $\ell^\infty\left(\mathcal{H}\right)$. Consider the process

$$\tau\mapsto\mathbb{G}_n r_d\left(W,\beta,\theta,\gamma,\tau\right)$$

By the uniform Hölder continuity of $\tau\mapsto(\tau,\beta\left(\tau\right))$ in τ with respect to the supremum norm, it is also Donsker in $\ell^\infty\left(\mathcal{T}\right)$ for all $d=0,1$. This, together with Assumption 2 implies

$$\mathbb{G}_n r_d\left(W,\beta\left(\tau\right),\theta,\gamma,\tau\right)\Rightarrow\sqrt{p_d}\mathbb{Z}_{R_d}\left(\tau\right)$$

[28]To see this, notice that both $\frac{\partial}{\partial\tau}C_{d,x}\left(\tau,\pi\right)\in[0,1]$ and $\pi_d\left(\tau\right)\in[0,1]$. Hence, it suffices to show that $\lim_{\pi\to1}\frac{\partial}{\partial\tau}G_{d,x}\left(\tau,\pi\right)=\lim_{\pi\to1}C_{d,x}\left(\tau,\pi\right)<\infty$, where I have used L'Hôpital rule. Since the derivative is bounded by Assumption 6, the result follows.

with covariate function

$$\Sigma_{R_d}(\tau,\tau') \equiv \mathbb{E}\left[\mathbb{Z}_{R_d}(\tau)\,\mathbb{Z}_{R_{d'}}(\tau')'\right] = \begin{bmatrix} \Sigma^{11}_{R_{dd'}}(\tau,\tau') & \Sigma^{12}_{R_{d'}}(\tau')' & 0 \\ \Sigma^{12}_{R_d}(\tau) & \Sigma^{22}_{R_{dd'}} & 0 \\ 0 & 0 & \Sigma^{33}_{R_{dd'}} \end{bmatrix}$$

where

$$\Sigma^{11}_{R_{dd'}}(\tau,\tau') = \mathbb{E}\left[S\left(G_{d,X,\tau} \wedge G_{d',X,\tau'} - G_{d,X,\tau}G_{d,X,\tau'}\right)XX'\right]$$

$$\Sigma^{12}_{R_d}(\tau) = \mathbb{E}\left[S\int_0^1 X\varphi(u,Z)'\left[G_{d,X,\tau} \wedge G_{d',X,u} - G_{d,X,\tau}G_{d',X,u}\right]du\right]$$

$$\Sigma^{22}_{R_{dd'}} = \mathbb{E}\left[\int_0^1\int_0^1 \varphi(u,Z)\,\varphi(v,Z)'\left[G_{d,X,u} \wedge G_{d',X,v} - G_{d,X,u}G_{d',X,v}\right]dvdu\right]$$

$$\Sigma^{33}_{R_{dd'}} = \mathbb{E}\left[s_d(S,Z;\gamma)\,s_{d'}(S,Z;\gamma)'\right]$$

where $\wedge$ denotes the minimum between two variables, and $G_{d,X,\tau} \equiv G_{d,X}(\tau,\pi(Z);\theta_d)$. Define ξ as the $L_2(P)$ pseudometric on $\mathcal{H}_d$:

$$\xi\left(\tilde{h}_d,h_d\right) \equiv \sqrt{\mathbb{E}\left\|r_d\left(W,\tilde{\beta},\tilde{\theta},\tilde{\gamma},\tilde{\tau}\right) - r_d(W,\beta,\theta,\gamma,\tau)\right\|^2}$$

Define $\delta_n \equiv \sup_{\tau\in\mathcal{T}} \xi\left(\tilde{h}_d(\tau),h_d(\tau)\right)\Big|_{\tilde{h}_d(\tau)=\hat{h}_d(\tau)}$. Since $\hat{\vartheta}_d(\tau) \xrightarrow{p} \vartheta_d(\tau)$ uniformly in τ for all $d=0,1$, by Assumption 4, $\delta_n \xrightarrow{p} 0$. Therefore, as $\delta_n \xrightarrow{p} 0$,

$$\begin{aligned} &\sup_{\tau\in\mathcal{T}}\left\|\mathbb{G}_n r_d\left(W,\hat{\beta},\hat{\theta},\hat{\gamma},\tau\right) - \mathbb{G}_n r_d(W,\beta,\theta,\gamma,\tau)\right\| \\ &\leq \sup_{\substack{\xi\left(\tilde{h}_d,h_d\right)\leq\delta_n \\ \tilde{h}_d,h_d\in\mathcal{H}_d}}\left\|\mathbb{G}_n r_d\left(W,\hat{\beta},\hat{\theta},\hat{\gamma},\tau\right) - \mathbb{G}_n r_d(W,\beta,\theta,\gamma,\tau)\right\| = o_P(1) \end{aligned}$$

by stochastic equicontinuity of $h_d \mapsto \mathbb{G}_n r_d(W,\beta,\theta,\gamma,\tau)$. Because the number of total groups is finite, this proves claim 2. To prove claim 1, define

$$\mathcal{A}_d \equiv \{(\beta,\theta,\gamma,\tau) \mapsto q_d(W,\beta,\theta,\gamma,\tau)\}$$

By Assumption 1, $\mathcal{A}_d$ is bounded, and it is also uniformly Lipschitz over $\mathcal{B}\times\Theta\times\Gamma\times\mathcal{T}$, so by Theorem 2.10.6 in van der Vaart and Wellner (1996), $\mathcal{A}_d$ is Donsker. Hence, the following ULLN holds:

$$\sup_{h_d \in \mathcal{H}_d} |\mathbb{E}_n q_d(W,\beta,\theta,\gamma,\tau) - \mathbb{E} q_d(W,\beta,\theta,\gamma,\tau)| \xrightarrow{p} 0$$

which gives

$$\sup_{(\beta,\theta,\gamma,\tau)\in\mathcal{B}\times\Theta\times\Gamma\times\mathcal{T}} |\mathbb{E}_n q_d(W,\beta,\theta,\gamma,\tau) - \mathbb{E} q_d(W,\beta,\theta,\gamma,\tau)| \xrightarrow{p} 0$$

Using the fact that the number of total groups is finite, it implies claim 1. □

C Identification under a Parametric Assumption

Consider the following assumptions:

Assumption 11. *(U,V) is jointly statistically independent of Z_1 given $X=x$.*

Assumption 12. *The bivariate distribution of (U,V) given $X=x$ is absolutely continuous with respect to the Lebesgue measure, with standard uniform marginals and rectangular support. Denote its cumulative distribution function as $C_{d,x}(u,v)$.*

Assumption 13. *The conditional cdf $F_{Y^*|X}(y|x)$ and its inverse are strictly increasing. In addition, $C_{d,x}(u,v)$ is strictly increasing in u.*

Assumption 14. *$\pi_d(Z) \equiv \mathbb{P}(S=1|D=d,Z) > 0$ with probability 1.*

Assumptions 11-14 are Assumptions A1-A4 from Arellano and Bonhomme (2017). They are not enough to identify the SQF and the copula, but they can be used to put some restrictions on the conditional copula, as they show in Lemma 1. Specifically, their Equation 6 can be written using this paper's notation as:

$$F_{Y|D=d,Z,S=1}\left(F^{-1}_{Y|D=d,Z,S=1}\left(\tau|z'\right)|z\right) = G_{d,x}\left(G^{-1}_{d,x}\left(\tau,\pi_d\left(z'\right)\right),\pi_d\left(z\right)\right) \tag{30}$$

where $F_{Y|D=d,Z,S=1}(\tau|z)$ is the cdf of Y, conditional on $D = d$, $Z = z$, and $S = 1$. Point identification is achieved when they assume either enough variation in the instrument such that one can use an identification at infinity argument, or that the copula is real analytic and the instrument displays some continuous variation. A feasible alternative would be to impose the following parametric assumption:

Assumption 15. *The copula $C_{d,x}(\tau,\pi)$ is known up to a scalar parameter $\theta_{d,x} \in \Theta_{d,x}$, for $d = 0,1$ and $\forall x \in \mathcal{X}$. $C_{d,x} : (0,1)^2 \to (0,1)$ is uniformly continuous and twice continuously differentiable in its arguments and in $\theta_{d,x}$ a.e. Moreover, for any $\theta_1 < \theta_2$, $C_{d,x}(\tau,\pi;\theta_2)$ is strictly more stochastically increasing in joint distribution than $C_{d,x}(\tau,\pi;\theta_1)$.*

The identification under this assumption is established by the following proposition:

Proposition 1. *Let Assumptions 11 to 15 hold, and $x \in \mathcal{X}$. Then, for $d = 0,1$, the functions $(\tau,\pi) \to G_{d,x}(\tau,\pi)$ and $\tau \to g_d(x,\tau)$ are globally identified.*

Proof. The proof is split in parts. First, I show the local identification of $\theta_{d,x}$, followed by its global identification, and I conclude by showing the identification of the SQF.

Define the functions $M_{d,x}(\tau,\theta_{d,x}) \equiv G_{d,x}\left(G^{-1}_{d,x}\left(\tau,\pi_d\left(z'\right);\theta_{d,x}\right),\pi_d\left(z\right);\theta_{d,x}\right)$ and $\phi_{d,x}(\tau) \equiv F_{Y|D=d.Z,S=1}\left(F^{-1}_{Y|D=d,Z,S=1}\left(\tau|z'\right)|z\right)$. By Equation 30, $M_{d,x}(\tau,\theta_{d,x}) = \phi_{d,x}(\tau)$, $\forall x \in \mathcal{X}, d = 0,1$. Taking the derivative with respect to the copula parameter for a generic value of θ, and dropping the (d,x) subscript from the functions M and ϕ and d from the propensity score for notational simplicity, yields

$$\begin{aligned}\nabla_\theta M(\tau,\theta) &= \nabla_\theta G\left(G^{-1}\left(\tau,\pi\left(z'\right);\theta\right),\pi\left(z\right);\theta\right)\\ &\quad - \nabla_u G\left(G^{-1}\left(\tau,\pi\left(z'\right);\theta\right),\pi\left(z\right);\theta\right)\frac{\nabla_\theta G\left(G^{-1}\left(\tau,\pi\left(z'\right);\theta\right),\pi\left(z'\right);\theta\right)}{\nabla_u G\left(G^{-1}\left(\tau,\pi\left(z'\right);\theta\right),\pi\left(z'\right);\theta\right)}\end{aligned} \tag{31}$$

Because $M(\tau,\theta)$ holds for any $\tau\in(0,1)$, there is an continuum of moments that pin down the parameter θ. Instead, consider a finite number of values of τ, given by $\{\tau_1,...,\tau_T\}$. Local identification holds when the matrix that collects the Jacobian for all values in this set is of full rank, as required by Theorem 6 in Rothenberg (1971). Because it is a scalar parameter, full rank is attained if $\nabla_\theta M(\tau,\theta)\neq 0$ for any of the values of τ, *i.e.*,

$$\frac{\nabla_\theta G\left(G^{-1}\left(\tau,\pi\left(z'\right);\theta\right),\pi\left(z\right);\theta\right)}{\nabla_u G\left(G^{-1}\left(\tau,\pi\left(z'\right);\theta\right),\pi\left(z\right);\theta\right)}-\frac{\nabla_\theta G\left(G^{-1}\left(\tau,\pi\left(z'\right);\theta\right),\pi\left(z'\right);\theta\right)}{\nabla_u G\left(G^{-1}\left(\tau,\pi\left(z'\right);\theta\right),\pi\left(z'\right);\theta\right)}\neq 0 \tag{32}$$

Let $\tau'\equiv G^{-1}\left(\tau,\pi\left(z'\right);\theta\right)\Leftrightarrow\tau=G\left(\tau',\pi\left(z'\right);\theta\right)$. Then, Equation 32 can be rewritten as

$$\frac{\nabla_\theta G\left(\tau',\pi\left(z\right);\theta\right)}{\nabla_u G\left(\tau',\pi\left(z\right);\theta\right)}-\frac{\nabla_\theta G\left(\tau',\pi\left(z'\right);\theta\right)}{\nabla_u G\left(\tau',\pi\left(z'\right);\theta\right)}\neq 0 \tag{33}$$

By the definition of the conditional copula, $\nabla_\theta G(\tau,\pi;\theta)/\nabla_u G(\tau,\pi;\theta)=\nabla_\theta C(\tau,\pi;\theta)/\nabla_u C(\tau,\pi;\theta)$, so Equation 33 is equivalent to

$$\frac{\nabla_\theta C\left(\tau',\pi\left(z\right);\theta\right)}{\nabla_u C\left(\tau',\pi\left(z\right);\theta\right)}-\frac{\nabla_\theta C\left(\tau',\pi\left(z'\right);\theta\right)}{\nabla_u C\left(\tau',\pi\left(z'\right);\theta\right)}\neq 0 \tag{34}$$

This is equivalent to the first equation in Condition 4.7 in Han and Vytlacil (2017). By Lemma 4.1 in Han and Vytlacil (2017), under Assumption 15, if the copula $G_{d,x}(\tau,\pi)$ satisfies Assumption 6 in Han and Vytlacil (2017), then Equation 34 is strictly decreasing in the second argument of the copula. If $\pi(z)\neq\pi(z')$, *i.e.*, if the instrument does not come from a degenerate distribution, then the copula parameter $\theta_{d,x}$ is locally identifiable by Proposition 4.1 in Han and Vytlacil (2017).

For global identification, I begin by showing that it is possible to apply Lemma 4.2 in

Han and Vytlacil (2017) on a restricted parameter space, extending it subsequently to the entire parameter space. Note that Equation 31 can be rewritten as

$$\nabla_\theta M(\tau,\theta) = \nabla_u G(\tau',\pi(z);\theta)\left[\frac{\nabla_\theta G(\tau',\pi(z);\theta)}{\nabla_u G(\tau',\pi(z);\theta)} - \frac{\nabla_\theta G(\tau',\pi(z');\theta)}{\nabla_u G(\tau',\pi(z');\theta)}\right]$$

where $\tau' \equiv G^{-1}(\tau,\pi(z');\theta)$. By Lemma 4.1 in Han and Vytlacil (2017), the term in brackets is positive when $\pi(z) < \pi(z')$. Moreover, $\nabla_u G(\tau,\pi;\theta) = \frac{1}{\pi}\nabla_u C(\tau,\pi;\theta) > 0$. Therefore, the Jacobian $\nabla_\theta M(\tau,\theta)$ is positive semidefinite if $\pi(z) < \pi(z')$ and negative semidefinite if $\pi(z) > \pi(z')$. In addition, it has full rank for any θ as long as $\pi(z) \neq \pi(z')$.

Let $\Theta_c \subseteq \Theta$ be a bounded open space with half spaces $\Theta_{c_1} \equiv \{\theta \in \Theta_c : \pi(z) < \pi(z')\}$, and $\Theta_{c_2} \equiv \{\theta \in \Theta_c : \pi(z) > \pi(z')\}$, which are simply connected. Define $\phi_{c_1}(\tau) = M(\tau,\Theta_{c_1})$ and $\phi_{c_2}(\tau) = M(\tau,\Theta_{c_2})$, and let $M|_{\Theta_{c_1}} : \Theta_{c_1} \to \phi_{c_1}$ and $M|_{\Theta_{c_2}} : \Theta_{c_2} \to \phi_{c_2}$ be the function $M(\tau,\cdot)$ on its restricted domains.

Because $M|_{\Theta_{c_1}}(\tau,\cdot)$ and $M|_{\Theta_{c_2}}(\tau,\cdot)$ are continuous, the pre-image of a closed set under $M|_{\Theta_{c_1}}(\tau,\cdot)$ and $M|_{\Theta_{c_2}}(\tau,\cdot)$ is closed. Because Θ_{c_1} and Θ_{c_2} are bounded, the pre-image of a bounded set is bounded. Thus, $M|_{\Theta_{c_1}}(\tau,\cdot)$ and $M|_{\Theta_{c_2}}(\tau,\cdot)$ are proper.

Because Θ_{c_1} and Θ_{c_2} are simply connected, $M|_{\Theta_{c_1}}(\tau,\cdot)$ and $M|_{\Theta_{c_2}}(\tau,\cdot)$ are continuous on Θ_{c_1} and Θ_{c_2}, respectively, and the Jacobian $\nabla_\theta M(\tau,\cdot)$ is positive semidefinite and negative semidefinite on Θ_{c_1} and Θ_{c_2}, respectively, it follows that ϕ_{c_1} and ϕ_{c_2} are simply connected.

Also, $\nabla_\theta M(\tau,\cdot)$ has full rank over Θ_{c_1} and Θ_{c_2}. Thus, by Lemma 4.2 in Han and Vytlacil (2017), $\phi(\tau) = M(\tau,\theta)$ has a unique solution on Θ_{c_1} and Θ_{c_2}, respectively. Because there exist $M|^{-1}_{\Theta_{c_1}}(\tau,\cdot) \in \Theta_{c_1}$ for $\phi \in \phi_{c_1}$ and $M|^{-1}_{\Theta_{c_2}}(\tau,\cdot) \in \Theta_{c_2}$ for $\phi \in \phi_{c_2}$, θ is globally identified.

Now let $\Theta_1 \equiv \{\theta \in \Theta : \pi(z) < \pi(z')\}$ and $\Theta_2 \equiv \{\theta \in \Theta : \pi(z) > \pi(z')\}$ be two simply connected, possibly unbounded spaces. Θ_1 and Θ_2 can be represented as a countable union of bounded open simply connected sets. *E.g.*, $\Theta_j = \cup_{i=1}^{\infty}\Theta_{ji}$, where Θ_{ji} is a sequence of bounded open simply connected sets in Θ_j such that $\Theta_{j1} \subset \Theta_{j2} \subset ... \subset \Theta_j$ for $j = 1, 2$.

Let $\phi_{ji}(\tau) \equiv M(\tau,\Theta_{ji})$ for $i = 1, 2, ...$ and $j = 1, 2$. Then, $\phi_j(\tau) = M(\tau,\Theta_j) = M(\tau,\cup_{i=1}^{\infty}\Theta_{ji}) = \cup_{i=1}^{\infty}M(\tau,\Theta_{ij}) = \cup_{i=1}^{\infty}\phi_{ij}(\tau)$, and $\phi_{j1} \subset \phi_{j2} \subset ... \subset \phi_j$. Then, for any given $\phi \in \phi_j$, $\exists q : \phi \in \phi_{ji} \forall i \geq q$, so $M|_{\Theta_{c_j}}^{-1}(\tau, \phi) \in \Theta_{ji} \forall i \geq q$, and therefore $M^{-1}(\tau, \phi) = M|_{\cup_{i=q}^{\infty}\Theta_{ji}}^{-1}(\tau, \phi) \in \cup_{i=q}^{\infty}\Theta_{ji} = \Theta_j$. Because $M^{-1}(\tau, \phi)$ is the unique solution on Θ_j, it is the unique solution of the full system with $\tau = \{\tau_1, ..., \tau_T\}$. Thus, θ is globally identified in Θ_j.

To show the identification of the SQF, note that by Equation 30, $F_{Y|D=d,Z,S=1}(g_d(x, \tau)|z) = G_{d,x}(\tau, \pi(z); \theta_{d,x})$. Therefore, one can solve for g_d for $d = 0, 1$, and express it in terms of either observed or identified functions: $g_d(x, \tau) = F_{Y|D=d,Z,S=1}^{-1}(G_{d,x}(\tau, \pi(z); \theta_{d,x}))$. □

D Additional Theoretical Results

The estimators of Equations 4, 6 and 8 are given by

$$\begin{aligned}
\hat{\mathbb{E}}\left[Y^{\ell}\right] &= \frac{1}{\tilde{n}_h}\sum_{i=1}^{n}\int_{\varepsilon}^{1-\varepsilon}\hat{g}_k(x_i, u)\, d\hat{C}_{l,x}(u, \hat{\pi}_m(z_i))\, \mathbf{1}(d_i = h) \\
\hat{F}_Y^{\ell}(y) &= \frac{1}{\tilde{n}_h}\sum_{i=1}^{n}\left[\varepsilon + \int_{\varepsilon}^{1-\varepsilon}\mathbf{1}(\hat{g}_k(x_i, u) \leq y)\, d\hat{C}_{l,x}(u, \hat{\pi}_m(z_i)) + (1 - \pi_m(z_i))\right]\mathbf{1}(d_i = h) \\
\hat{Q}_Y^{\ell}(\tau) &= \inf\left\{y : \tau \leq \hat{F}_Y^{\ell}(y)\right\}
\end{aligned}$$

where $\tilde{n}_h = \sum_{i=1}^{n}\mathbf{1}(d_i = h)$. Under the same condition required for Theorem 1 to hold, these estimators are consistent and asymptotically Gaussian:

Theorem 5. *Let the estimator $\hat{v}_{\ell}(z, \tau, \eta)$ satisfy Condition 1. Under Assumptions 1-4, the following hold for all (ℓ, ℓ'):*

$$\sqrt{n}\Delta^{\ell,\ell'}\left(\hat{\mathbb{E}}[Y] - \mathbb{E}[Y]\right) \Rightarrow \mathbb{Z}_{\Delta^{\ell\ell'}Y}$$

where $\mathbb{Z}_{\Delta^{\ell\ell'}Y}$ is a zero-mean Gaussian random variable, and

$$\sqrt{n}\Delta^{\ell,\ell'}\left(\hat{Q}_Y(\tau) - Q_Y(\tau)\right) \Rightarrow \mathbb{Z}_{Q,\ell\ell'}(\tau)$$

where $\mathbb{Z}_{Q,\ell\ell'}(\tau)$ is a zero-mean Gaussian process.

Proof. Using the same argument used in Theorem 1, it is straightforward to show that

$$\sqrt{n}\Delta^{\ell,\ell'}\left(\hat{\mathbb{E}}\left[Y\right]-\mathbb{E}\left[Y\right]\right)\Rightarrow\mathbb{Z}_{Y^{\ell}}-\mathbb{Z}_{Y^{\ell'}}\equiv\mathbb{Z}_{\Delta^{\ell\ell'}Y}$$

where

$$\begin{aligned}\mathbb{Z}_{\Delta^{\ell\ell'}Y}=&\int_{\mathcal{Z}}\int_{\varepsilon}^{1-\varepsilon}\mathbb{Z}_{g_k}\left(\tau,x\right)dC_{l,x}\left(\tau|\pi_m\left(z\right)\right)dF_Z^h\left(z\right)\\&+\int_{\mathcal{Z}}\int_{\varepsilon}^{1-\varepsilon}g_k\left(x,\tau\right)\mathbb{Z}_{c_l}\left(\tau,\pi_m\left(z\right)\right)d\tau dF_Z^h\left(z\right)\\&+\int_{\mathcal{Z}}\int_{\varepsilon}^{1-\varepsilon}g_k\left(x,\tau\right)d\left(\nabla_{\pi}C_{l,x}\left(\tau|\pi_m\left(z\right)\right)\right)\mathbb{Z}_{\pi_m}\left(z\right)dF_Z^h\left(z\right)\\&+\sqrt{p_h}\mathbb{Z}_{Z_h}\left(\int_{\varepsilon}^{1-\varepsilon}g_k\left(\cdot,\tau\right)dC_{l,x}\left(\tau|\pi\left(\cdot\right)\right)\right)\end{aligned}$$

Similarly,

$$\sqrt{n}\Delta^{\ell,\ell'}\left(\hat{Q}_Y\left(\tau\right)-Q_Y\left(\tau\right)\right)\Rightarrow\mathbb{Z}_{Q_Y^{\ell}}\left(\tau\right)-\mathbb{Z}_{Q_Y^{\ell'}}\left(\tau\right)\equiv\mathbb{Z}_{Q,\ell\ell'}\left(\tau\right)$$

where

$$\begin{aligned}\mathbb{Z}_{Q_Y^{\ell}}\left(\tau\right)=&-\frac{\mathbb{Z}_{F_Y^{\ell}}\left(Q_Y^{\ell}\left(\tau\right)\right)}{f_Y^{\ell}\left(Q_Y^{\ell}\left(\tau\right)\right)}\\\mathbb{Z}_{F_Y^{\ell}}\left(y\right)=&\int_{\mathcal{Z}}\left(\hat{F}_{Y|Z}^{\ell}\left(y|z\right)-F_{Y|Z}^{\ell}\left(y|z\right)\right)dF_Z^h\left(z\right)+\int_{\mathcal{Z}}F_{Y|Z}^{\ell}\left(y|z\right)d\left(\hat{F}_Z^h\left(z\right)-F_Z^h\left(z\right)\right)\\\Rightarrow&\int_{\mathcal{Z}}\mathbb{Z}_{F_Y^{\ell}|X}\left(y|z\right)dF_Z^h\left(z\right)+\sqrt{p_h}\mathbb{Z}_{Z_h}\left(F_{Y|Z}^{\ell}\left(y|z\right)\right)\\\mathbb{Z}_{F_Y^{\ell}|Z}\left(y,z\right)=&-f_{Y|Z}\left(y|z\right)\frac{1}{\pi_m\left(z\right)}c_{l,x}\left(C_{l,x}\left(\tilde{u}_k\left(x,y\right),\pi_m\left(z\right)\right),\pi_m\left(z\right)\right)x'\mathbb{Z}_{\beta_k}\left(F_{Y|Z}\left(y|z\right)\right)\\&+\int_{\varepsilon}^{1-\varepsilon}\mathbf{1}\left(x'\beta_k\left(\tau\right)\leq y\right)\mathbb{Z}_{c_l}\left(\tau,\pi_m\left(z\right)\right)d\tau\\&+\left[\int_{\varepsilon}^{1-\varepsilon}\mathbf{1}\left(x'\beta_k\left(\tau\right)\leq y\right)\nabla_{\pi}c_{l,x}\left(u,\pi_m\left(z\right)\right)d\tau-1\right]\mathbb{Z}_{\pi_m}\left(z\right)\end{aligned}$$

and an analog version of Lemma 3 for the distribution of the entire population, *i.e.*, $F_{Y|Z}\left(y|z\right)$, is applied. □

E Additional Results

E.1 Potential Outcomes for the Entire Population

Table 10 reports the potential earnings for the full population. Because the selection into employment was negative in the early years of the sample and turned positive afterwards, this distribution tends to be above the distribution of actual earnings for participants at the beginning of the period, and below it towards the end. This is most evident for the distributions of females. The different interquantile range measures (10-90th percentiles, 25-75th percentiles) are of a similar magnitude to those found in Table 6.

The decomposition of mean potential earnings for the entire population is reported in Table 11. The gap displays an erratic behavior, driven by the coefficients component, that is more variable than for the two main decompositions. Hence, this gap is more sensitive to the QRS estimates, which may be less robust if the instrument used is weak for men. Regardless, the endowments component switches sign over time, thus helping reduce the gender gap.

As it was the case for the mean, the decomposition of the unconditional distributions of potential earnings for the entire population are quite similar to those found for actual earnings for participants (Figure 6; Tables 12-16). Regarding its components, the coefficients component is dominant. On the other hand, the endowments component has a slightly increasing shape and is much smaller in magnitude. In the early years it was negative for the left tail and positive for the majority of the distribution, and it has progressively become more negative throughout the entire distribution, now contributing the the reduction of the gap.

Table 10: Potential earnings distributions for the full population by gender (Frank copula)

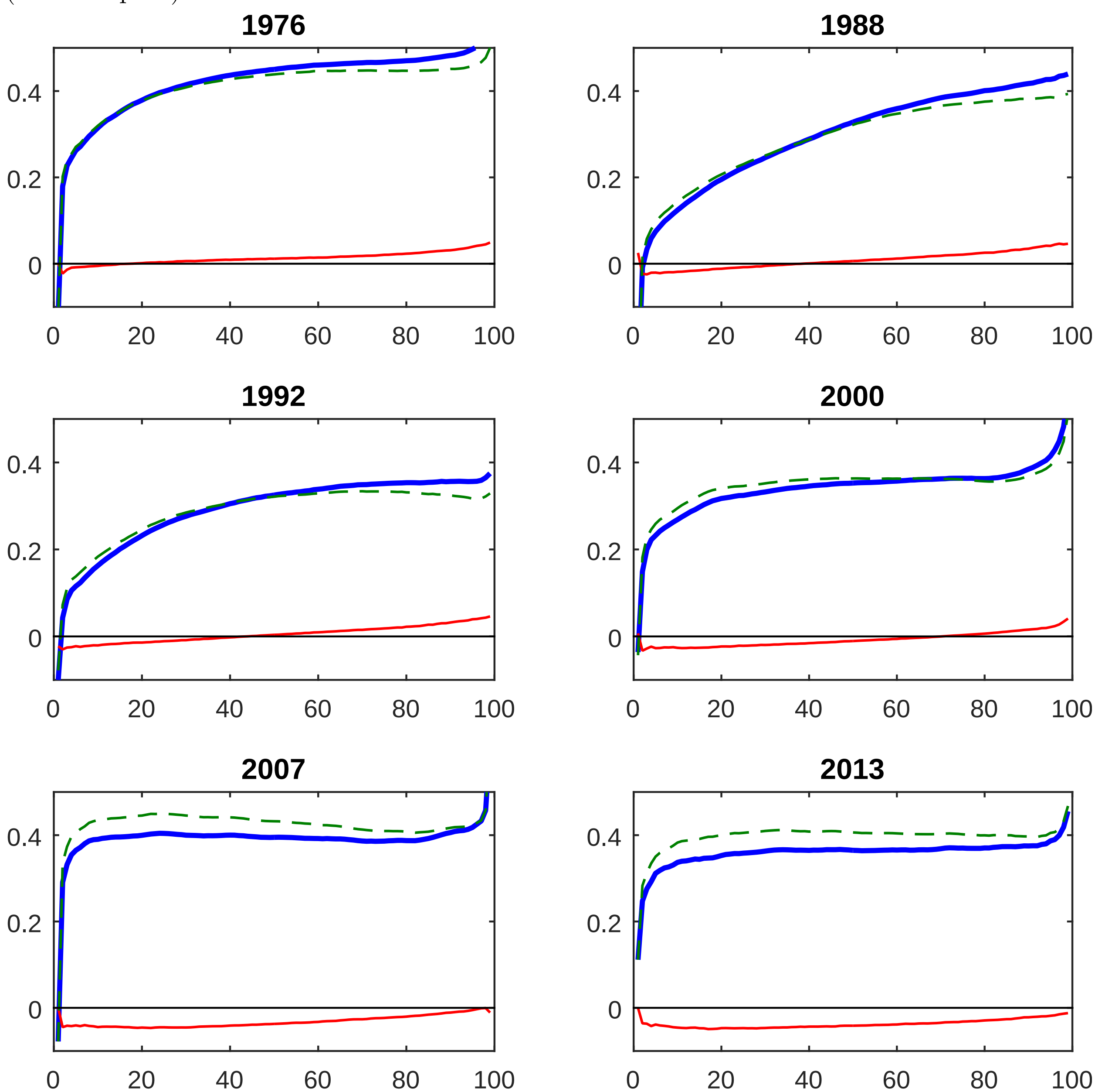
	Male						Female					
Year	Mean	P10	P25	P50	P75	P90	Mean	P10	P25	P50	P75	P90
1976	2.80	2.11	2.48	2.84	3.15	3.40	2.38	1.80	2.08	2.39	2.68	2.92
1977	2.77	2.06	2.44	2.81	3.13	3.38	2.37	1.80	2.07	2.37	2.67	2.91
1978	2.80	2.08	2.46	2.83	3.16	3.41	2.31	1.72	2.00	2.31	2.61	2.86
1979	2.82	2.10	2.47	2.85	3.18	3.43	2.31	1.74	2.02	2.31	2.61	2.87
1980	2.80	2.09	2.46	2.84	3.16	3.42	2.32	1.76	2.02	2.31	2.61	2.87
1981	2.75	2.02	2.40	2.78	3.11	3.37	2.29	1.73	2.00	2.29	2.58	2.85
1982	2.72	1.98	2.36	2.75	3.09	3.36	2.25	1.67	1.94	2.24	2.55	2.82
1983	2.66	1.89	2.27	2.69	3.06	3.34	2.22	1.61	1.89	2.21	2.53	2.80
1984	2.61	1.80	2.21	2.64	3.02	3.31	2.31	1.68	1.98	2.31	2.63	2.90
1985	2.68	1.88	2.29	2.72	3.09	3.37	2.28	1.63	1.93	2.27	2.61	2.89
1986	2.69	1.88	2.29	2.72	3.10	3.38	2.25	1.59	1.90	2.24	2.59	2.89
1987	2.71	1.90	2.31	2.74	3.12	3.42	2.23	1.53	1.86	2.22	2.59	2.90
1988	2.65	1.81	2.24	2.68	3.07	3.37	2.29	1.58	1.92	2.30	2.65	2.95
1989	2.63	1.80	2.22	2.66	3.05	3.36	2.31	1.60	1.93	2.30	2.66	2.97
1990	2.62	1.78	2.20	2.64	3.04	3.36	2.24	1.49	1.84	2.23	2.61	2.94
1991	2.61	1.78	2.19	2.63	3.03	3.35	2.30	1.59	1.92	2.30	2.66	2.97
1992	2.59	1.75	2.17	2.61	3.02	3.33	2.30	1.59	1.91	2.29	2.66	2.98
1993	2.49	1.60	2.03	2.51	2.94	3.28	2.29	1.57	1.90	2.29	2.67	2.98
1994	2.48	1.60	2.03	2.50	2.94	3.28	2.26	1.52	1.86	2.26	2.65	2.97
1995	2.49	1.59	2.02	2.50	2.94	3.29	2.23	1.47	1.81	2.22	2.63	2.97
1996	2.62	1.78	2.18	2.62	3.03	3.37	2.23	1.46	1.81	2.22	2.62	2.97
1997	2.60	1.77	2.17	2.61	3.01	3.35	2.26	1.50	1.83	2.24	2.64	2.99
1998	2.51	1.63	2.05	2.52	2.95	3.31	2.27	1.52	1.86	2.26	2.65	2.99
1999	2.68	1.85	2.26	2.68	3.08	3.43	2.29	1.52	1.86	2.27	2.68	3.02
2000	2.70	1.86	2.27	2.70	3.10	3.46	2.35	1.59	1.94	2.34	2.74	3.08
2001	2.74	1.92	2.30	2.73	3.13	3.52	2.40	1.64	1.99	2.39	2.77	3.11
2002	2.71	1.88	2.27	2.69	3.10	3.49	2.40	1.64	1.99	2.39	2.77	3.12
2003	2.71	1.88	2.27	2.70	3.11	3.49	2.34	1.56	1.91	2.33	2.74	3.09
2004	2.72	1.88	2.28	2.71	3.12	3.51	2.36	1.57	1.93	2.34	2.75	3.10
2005	2.72	1.88	2.27	2.70	3.12	3.51	2.34	1.55	1.91	2.33	2.75	3.10
2006	2.79	1.95	2.33	2.76	3.19	3.59	2.37	1.58	1.94	2.35	2.76	3.11
2007	2.70	1.87	2.25	2.68	3.11	3.50	2.30	1.48	1.84	2.29	2.72	3.10
2008	2.72	1.89	2.28	2.70	3.13	3.51	2.30	1.48	1.85	2.29	2.72	3.09
2009	2.61	1.73	2.15	2.59	3.03	3.43	2.31	1.49	1.87	2.30	2.72	3.09
2010	2.53	1.61	2.03	2.51	2.97	3.39	2.32	1.49	1.86	2.30	2.73	3.12
2011	2.64	1.79	2.18	2.63	3.07	3.46	2.26	1.39	1.79	2.24	2.70	3.08
2012	2.49	1.56	1.99	2.48	2.96	3.38	2.27	1.44	1.81	2.25	2.69	3.08
2013	2.66	1.79	2.20	2.64	3.09	3.48	2.30	1.46	1.84	2.28	2.72	3.11

Table 11: Mean decomposition, potential earnings for the full population (Frank copula)

Year	Total	EC	CC	Year	Total	EC	CC
1976	0.43	0.01	0.41	1995	0.25	0.00	0.25
1977	0.40	0.01	0.39	1996	0.39	0.00	0.39
1978	0.49	0.01	0.48	1997	0.35	-0.01	0.36
1979	0.50	0.01	0.49	1998	0.24	-0.01	0.25
1980	0.48	0.01	0.48	1999	0.40	-0.01	0.40
1981	0.45	0.00	0.45	2000	0.34	-0.01	0.35
1982	0.47	0.01	0.46	2001	0.35	-0.01	0.36
1983	0.45	0.01	0.44	2002	0.31	-0.01	0.32
1984	0.30	0.01	0.29	2003	0.37	-0.02	0.38
1985	0.41	0.01	0.40	2004	0.36	-0.02	0.39
1986	0.44	0.01	0.43	2005	0.37	-0.02	0.40
1987	0.48	0.01	0.47	2006	0.42	-0.03	0.45
1988	0.36	0.01	0.35	2007	0.40	-0.03	0.43
1989	0.33	0.01	0.32	2008	0.42	-0.03	0.45
1990	0.38	0.01	0.37	2009	0.29	-0.04	0.33
1991	0.31	0.00	0.30	2010	0.21	-0.04	0.25
1992	0.29	0.00	0.29	2011	0.38	-0.04	0.42
1993	0.20	0.00	0.19	2012	0.22	-0.04	0.26
1994	0.22	0.00	0.22	2013	0.36	-0.04	0.40

Notes: Total, EC and CC respectively denote total difference, endowments component and coefficients component.

Figure 6: Unconditional quantiles decompositions, potential earnings for the full population (Frank copula)



Notes: the solid thick blue line denotes the total gap between male and female workers; the solid thin red line denotes the endowments component; the dashed green line denotes the coefficients component.

Table 12: 10th percentile decomposition, potential earnings for the full population (Frank copula)

Year	Total	EC	CC	Year	Total	EC	CC
1976	0.31	-0.01	0.32	1995	0.12	-0.03	0.15
1977	0.26	-0.01	0.27	1996	0.32	-0.03	0.34
1978	0.35	-0.01	0.36	1997	0.27	-0.03	0.30
1979	0.36	0.00	0.36	1998	0.10	-0.03	0.13
1980	0.33	-0.01	0.34	1999	0.34	-0.03	0.36
1981	0.29	-0.01	0.31	2000	0.27	-0.03	0.29
1982	0.31	-0.01	0.32	2001	0.28	-0.03	0.31
1983	0.28	-0.02	0.30	2002	0.24	-0.03	0.27
1984	0.12	-0.02	0.14	2003	0.31	-0.04	0.35
1985	0.25	-0.02	0.27	2004	0.31	-0.04	0.35
1986	0.29	-0.01	0.31	2005	0.33	-0.04	0.37
1987	0.37	-0.02	0.38	2006	0.36	-0.04	0.41
1988	0.23	-0.02	0.25	2007	0.39	-0.04	0.43
1989	0.20	-0.01	0.21	2008	0.42	-0.05	0.46
1990	0.30	-0.01	0.31	2009	0.24	-0.05	0.29
1991	0.19	-0.02	0.21	2010	0.12	-0.05	0.17
1992	0.16	-0.02	0.18	2011	0.40	-0.05	0.45
1993	0.03	-0.02	0.05	2012	0.12	-0.05	0.17
1994	0.08	-0.03	0.11	2013	0.34	-0.05	0.38

Notes: Total, EC and CC respectively denote total difference, endowments component and coefficients component.

Table 13: 25th percentile decomposition, potential earnings for the full population (Frank copula)

Year	Total	EC	CC	Year	Total	EC	CC
1976	0.40	0.00	0.40	1995	0.21	-0.02	0.24
1977	0.36	0.00	0.37	1996	0.37	-0.02	0.39
1978	0.45	0.00	0.45	1997	0.34	-0.02	0.36
1979	0.45	0.00	0.45	1998	0.19	-0.03	0.22
1980	0.44	0.00	0.44	1999	0.40	-0.02	0.42
1981	0.40	-0.01	0.41	2000	0.32	-0.02	0.35
1982	0.41	0.00	0.42	2001	0.31	-0.03	0.34
1983	0.38	-0.01	0.39	2002	0.28	-0.03	0.30
1984	0.22	-0.01	0.23	2003	0.36	-0.03	0.39
1985	0.36	-0.01	0.36	2004	0.35	-0.04	0.39
1986	0.40	-0.01	0.40	2005	0.36	-0.04	0.40
1987	0.45	-0.01	0.46	2006	0.39	-0.04	0.43
1988	0.31	0.00	0.32	2007	0.40	-0.05	0.45
1989	0.28	0.00	0.29	2008	0.43	-0.05	0.47
1990	0.36	-0.01	0.36	2009	0.28	-0.05	0.33
1991	0.26	-0.01	0.28	2010	0.17	-0.05	0.22
1992	0.26	-0.01	0.27	2011	0.39	-0.05	0.43
1993	0.13	-0.01	0.14	2012	0.18	-0.05	0.23
1994	0.16	-0.02	0.18	2013	0.36	-0.05	0.41

Notes: Total, EC and CC respectively denote total difference, endowments component and coefficients component.

Table 14: 50th percentile decomposition, potential earnings for the full population (Frank copula)

Year	Total	EC	CC	Year	Total	EC	CC
1976	0.45	0.01	0.44	1995	0.28	0.00	0.28
1977	0.44	0.01	0.43	1996	0.40	-0.01	0.41
1978	0.53	0.01	0.52	1997	0.37	-0.01	0.37
1979	0.54	0.01	0.53	1998	0.26	-0.01	0.27
1980	0.53	0.01	0.52	1999	0.41	-0.01	0.42
1981	0.50	0.00	0.49	2000	0.35	-0.01	0.36
1982	0.51	0.00	0.51	2001	0.34	-0.02	0.35
1983	0.48	0.00	0.48	2002	0.31	-0.01	0.32
1984	0.33	0.01	0.32	2003	0.37	-0.02	0.39
1985	0.44	0.01	0.44	2004	0.36	-0.03	0.39
1986	0.48	0.01	0.47	2005	0.37	-0.03	0.40
1987	0.53	0.01	0.52	2006	0.41	-0.03	0.44
1988	0.38	0.01	0.37	2007	0.39	-0.04	0.43
1989	0.36	0.01	0.35	2008	0.42	-0.04	0.45
1990	0.41	0.01	0.40	2009	0.30	-0.04	0.34
1991	0.33	0.00	0.33	2010	0.21	-0.04	0.25
1992	0.32	0.00	0.32	2011	0.38	-0.04	0.43
1993	0.22	0.00	0.22	2012	0.23	-0.04	0.27
1994	0.24	0.00	0.24	2013	0.36	-0.04	0.41

Notes: Total, EC and CC respectively denote total difference, endowments component and coefficients component.

Table 15: 75th percentile decomposition, potential earnings for the full population (Frank copula)

Year	Total	EC	CC	Year	Total	EC	CC
1976	0.47	0.02	0.45	1995	0.31	0.02	0.29
1977	0.46	0.02	0.45	1996	0.41	0.01	0.40
1978	0.55	0.02	0.53	1997	0.37	0.01	0.37
1979	0.57	0.02	0.55	1998	0.30	0.00	0.29
1980	0.56	0.02	0.54	1999	0.40	0.01	0.39
1981	0.53	0.01	0.52	2000	0.36	0.00	0.36
1982	0.54	0.02	0.53	2001	0.36	0.00	0.36
1983	0.53	0.02	0.51	2002	0.33	0.00	0.33
1984	0.39	0.02	0.37	2003	0.37	0.00	0.38
1985	0.48	0.02	0.46	2004	0.37	-0.01	0.38
1986	0.51	0.02	0.49	2005	0.38	-0.01	0.39
1987	0.53	0.02	0.51	2006	0.43	-0.02	0.45
1988	0.43	0.03	0.40	2007	0.39	-0.02	0.41
1989	0.39	0.02	0.36	2008	0.41	-0.02	0.44
1990	0.43	0.03	0.40	2009	0.31	-0.03	0.34
1991	0.36	0.02	0.35	2010	0.25	-0.03	0.28
1992	0.35	0.02	0.33	2011	0.37	-0.03	0.40
1993	0.28	0.02	0.26	2012	0.27	-0.03	0.30
1994	0.29	0.02	0.27	2013	0.37	-0.03	0.40

Notes: Total, EC and CC respectively denote total difference, endowments component and coefficients component.

Table 16: 90th percentile decomposition, potential earnings for the full population (Frank copula)

Year	Total	EC	CC	Year	Total	EC	CC
1976	0.48	0.03	0.45	1995	0.32	0.03	0.29
1977	0.47	0.03	0.44	1996	0.40	0.02	0.38
1978	0.55	0.03	0.52	1997	0.37	0.02	0.35
1979	0.57	0.03	0.54	1998	0.32	0.01	0.30
1980	0.55	0.03	0.52	1999	0.40	0.02	0.38
1981	0.52	0.02	0.50	2000	0.38	0.02	0.37
1982	0.54	0.03	0.51	2001	0.41	0.01	0.40
1983	0.54	0.03	0.50	2002	0.37	0.01	0.36
1984	0.42	0.03	0.38	2003	0.40	0.01	0.39
1985	0.48	0.04	0.45	2004	0.40	0.00	0.40
1986	0.49	0.04	0.46	2005	0.41	0.00	0.41
1987	0.51	0.04	0.48	2006	0.48	0.00	0.48
1988	0.42	0.04	0.39	2007	0.41	-0.01	0.42
1989	0.39	0.04	0.36	2008	0.41	-0.01	0.43
1990	0.42	0.04	0.38	2009	0.34	-0.02	0.36
1991	0.38	0.03	0.35	2010	0.27	-0.02	0.30
1992	0.36	0.03	0.32	2011	0.38	-0.02	0.40
1993	0.30	0.03	0.27	2012	0.29	-0.02	0.32
1994	0.31	0.03	0.28	2013	0.38	-0.02	0.40

Notes: Total, EC and CC respectively denote total difference, endowments component and coefficients component.

E.2 Generalized Entropy Measures of the Gap

I report the generalized entropy measures recommended by Maasoumi and Wang (2019) for the three populations considered. For the sake of brevity, I only comment the normalization of the Bhattacharya-Matusita-Hellinger measure, denoted by S_ρ, and the normalized and symmetrized Kullback-Leibler-Theil measure, denoted by Theil.

Table 17 shows the estimates of these measures for the distributions of actual earnings for participants, and both of them experience a steady decrease over time, suggesting an important convergence between the distributions of both genders. These results are similar if one considers the entire population (Table 18), with two main differences: the values of these measures are larger when one considers the entire population, and the reduction of the measures has also been more pronounced. The explanation for these differences lies in the large reduction of the participation rates between genders.

Finally, the behavior of the generalized entropy measures is also more volatile for the potential outcomes than for the actual outcomes. Regardless, the long-term trend is a marked decrease, pointing at a reduction in earnings differences between men and women.

Table 17: Generalized entropy measures of the gap, actual earnings for participants (Frank copula)

Year	S_ρ	Theil	$k=0.1$	$k=0.2$	$k=0.3$	$k=0.4$	$k=0.5$	$k=0.6$	$k=0.7$	$k=0.8$	$k=0.9$
1976	13.5	129.5	8.4	13.2	17.5	21.9	26.9	32.9	40.7	52.7	75.3
1977	12.3	109.4	7.3	11.8	15.8	20.0	24.6	30.0	36.9	47.2	65.8
1978	10.8	51.3	4.8	9.1	13.3	17.5	21.8	26.2	31.1	36.5	43.1
1979	10.9	49.8	4.7	9.1	13.3	17.5	21.8	26.3	31.1	36.4	42.7
1980	10.5	45.7	4.4	8.7	12.8	16.9	21.1	25.4	29.9	34.7	39.9
1981	10.1	60.5	4.9	8.8	12.5	16.3	20.2	24.4	29.2	35.2	44.0
1982	9.7	49.8	4.5	8.4	12.1	15.8	19.6	23.7	28.2	33.5	40.5
1983	8.5	41.0	4.1	7.6	10.9	14.2	17.7	21.3	25.4	30.3	37.3
1984	7.8	49.1	4.1	7.2	10.0	13.0	16.0	19.4	23.4	28.7	37.3
1985	7.1	35.8	3.4	6.2	8.9	11.7	14.5	17.5	20.9	24.9	30.5
1986	6.3	28.1	3.1	5.6	8.2	10.7	13.3	16.1	19.1	22.5	27.5
1987	5.7	23.7	3.0	5.3	7.7	10.0	12.5	15.1	17.9	21.3	26.8
1988	5.3	23.4	2.9	5.1	7.3	9.5	11.8	14.2	17.0	20.4	26.0
1989	5.1	22.8	2.9	5.0	7.1	9.2	11.4	13.8	16.5	20.0	25.9
1990	4.4	19.2	2.9	4.8	6.7	8.6	10.7	13.0	15.6	19.1	25.9
1991	3.8	15.8	2.1	3.7	5.3	7.0	8.6	10.4	12.4	15.0	19.3
1992	3.4	14.2	2.0	3.5	5.0	6.5	8.0	9.7	11.6	14.0	18.3
1993	3.1	13.0	2.3	3.6	5.0	6.5	8.0	9.7	11.7	14.5	20.3
1994	2.7	11.1	2.1	3.3	4.5	5.8	7.2	8.7	10.6	13.2	18.9
1995	2.6	10.5	2.2	3.4	4.6	5.9	7.2	8.8	10.7	13.5	19.8
1996	2.4	9.7	2.2	3.3	4.4	5.6	6.9	8.4	10.3	13.0	19.4
1997	2.3	9.2	2.2	3.2	4.4	5.5	6.8	8.3	10.2	13.0	19.6
1998	2.4	9.7	2.7	3.9	5.2	6.5	8.0	9.7	12.0	15.7	24.7
1999	2.3	9.5	2.3	3.4	4.6	5.8	7.1	8.7	10.7	13.7	20.9
2000	2.1	8.6	2.3	3.3	4.4	5.5	6.8	8.3	10.3	13.3	20.6
2001	1.9	7.6	2.3	3.3	4.3	5.3	6.6	8.0	9.9	13.0	20.8
2002	1.7	6.8	2.5	3.3	4.3	5.3	6.6	8.0	10.0	13.4	22.1
2003	1.6	6.4	2.6	3.4	4.4	5.4	6.6	8.1	10.2	13.8	23.1
2004	1.5	5.9	2.4	3.2	4.1	5.0	6.1	7.5	9.5	12.8	21.5
2005	1.5	5.9	2.3	3.1	4.0	5.0	6.1	7.5	9.4	12.6	21.1
2006	1.5	6.0	2.0	2.8	3.6	4.5	5.5	6.7	8.4	11.1	18.2
2007	1.3	5.2	2.5	3.3	4.1	5.0	6.1	7.5	9.5	13.0	22.4
2008	1.4	5.8	2.4	3.1	4.0	4.9	6.0	7.3	9.3	12.6	21.4
2009	1.3	5.2	2.6	3.4	4.3	5.2	6.4	7.9	10.0	13.7	23.8
2010	1.1	4.6	3.0	3.8	4.7	5.7	6.9	8.6	10.9	15.3	27.4
2011	1.3	5.5	2.7	3.4	4.3	5.3	6.4	7.9	10.0	13.8	24.0
2012	1.1	4.3	3.0	3.8	4.6	5.6	6.8	8.5	10.8	15.2	27.4
2013	1.1	4.3	2.3	3.0	3.7	4.6	5.6	6.8	8.7	12.0	21.0

Table 18: Generalized entropy measures of the gap, actual earnings for the full population (Frank copula)

Year	S_ρ	Theil	$k=0.1$	$k=0.2$	$k=0.3$	$k=0.4$	$k=0.5$	$k=0.6$	$k=0.7$	$k=0.8$	$k=0.9$
1976	11.3	50.2	5.3	9.9	14.4	18.9	23.5	28.3	33.6	39.7	47.7
1977	11.0	51.1	5.3	9.8	14.2	18.5	23.0	27.8	33.1	39.3	48.1
1978	10.9	52.3	5.4	9.9	14.2	18.5	23.0	27.7	33.1	39.5	48.9
1979	10.9	51.2	5.3	9.8	14.1	18.4	22.8	27.6	32.9	39.2	48.1
1980	10.0	45.6	4.9	9.0	13.0	17.0	21.1	25.4	30.2	35.9	43.9
1981	9.2	43.5	4.5	8.3	11.9	15.6	19.4	23.4	27.8	33.1	40.8
1982	8.5	43.2	4.5	7.9	11.3	14.6	18.1	21.9	26.3	31.7	40.1
1983	7.1	34.0	3.8	6.7	9.6	12.4	15.4	18.6	22.3	26.9	34.2
1984	6.2	27.9	3.3	5.9	8.4	10.9	13.6	16.4	19.6	23.5	29.8
1985	6.1	29.9	3.4	5.9	8.4	10.9	13.5	16.3	19.5	23.6	30.6
1986	5.8	25.8	3.2	5.7	8.1	10.5	13.0	15.8	18.8	22.6	28.9
1987	5.4	22.7	3.1	5.3	7.6	9.9	12.3	14.9	17.7	21.4	27.6
1988	5.0	21.7	3.0	5.1	7.2	9.3	11.6	14.0	16.8	20.3	26.8
1989	4.7	20.9	2.9	4.9	6.9	9.0	11.1	13.5	16.2	19.7	26.3
1990	4.6	19.9	2.9	4.9	6.9	9.0	11.1	13.4	16.1	19.7	26.5
1991	4.1	16.9	2.5	4.3	6.0	7.8	9.7	11.7	14.1	17.1	22.8
1992	3.6	14.7	2.3	3.9	5.4	7.0	8.7	10.5	12.6	15.4	20.8
1993	3.2	13.4	2.2	3.7	5.1	6.6	8.1	9.8	11.9	14.6	20.2
1994	3.1	12.5	2.2	3.5	4.9	6.3	7.8	9.4	11.4	14.0	19.5
1995	3.1	12.6	2.2	3.6	5.0	6.4	7.9	9.6	11.6	14.4	20.2
1996	2.9	11.8	2.4	3.7	5.1	6.5	8.0	9.7	11.8	14.9	21.7
1997	2.7	11.1	2.4	3.6	4.9	6.3	7.8	9.4	11.5	14.6	21.5
1998	2.8	11.3	2.6	3.9	5.2	6.6	8.2	10.0	12.2	15.6	23.3
1999	2.8	11.5	2.6	3.9	5.2	6.7	8.2	10.0	12.2	15.6	23.2
2000	2.6	10.5	2.6	3.9	5.2	6.5	8.0	9.8	12.0	15.5	23.7
2001	2.4	9.9	2.7	3.9	5.2	6.6	8.1	9.8	12.1	15.8	24.7
2002	2.3	9.4	2.8	3.9	5.1	6.5	7.9	9.7	12.0	15.7	24.9
2003	2.1	8.5	2.6	3.7	4.8	6.0	7.4	9.0	11.2	14.7	23.5
2004	2.0	8.0	2.5	3.5	4.5	5.7	7.0	8.5	10.6	13.9	22.2
2005	2.0	8.1	2.5	3.5	4.5	5.7	7.0	8.5	10.6	13.9	22.2
2006	1.9	7.7	2.4	3.3	4.4	5.5	6.7	8.2	10.2	13.4	21.4
2007	1.9	7.6	2.5	3.5	4.5	5.7	6.9	8.5	10.6	14.0	22.7
2008	1.8	7.2	2.4	3.3	4.3	5.3	6.5	8.0	10.0	13.2	21.5
2009	1.7	6.8	2.4	3.3	4.3	5.3	6.5	7.9	9.9	13.2	21.8
2010	1.3	5.2	2.3	3.1	3.9	4.8	5.8	7.2	9.1	12.3	21.0
2011	1.3	5.4	2.2	2.9	3.7	4.6	5.6	6.9	8.6	11.6	19.5
2012	1.4	5.6	2.4	3.1	4.0	4.9	6.0	7.4	9.3	12.6	21.3
2013	1.4	5.8	2.2	3.0	3.8	4.8	5.8	7.2	9.0	12.0	20.1

Table 19: Generalized entropy measures of the gap, potential earnings for the full population (Frank copula)

Year	S_ρ	Theil	$k=0.1$	$k=0.2$	$k=0.3$	$k=0.4$	$k=0.5$	$k=0.6$	$k=0.7$	$k=0.8$	$k=0.9$
1976	11.2	49.5	5.0	9.5	13.9	18.4	22.8	27.5	32.5	38.1	44.8
1977	10.5	50.6	4.8	9.0	13.1	17.2	21.4	25.8	30.6	36.0	43.1
1978	14.0	69.0	6.5	12.1	17.5	22.9	28.4	34.3	40.8	48.3	58.1
1979	14.6	71.3	6.7	12.5	18.2	23.8	29.6	35.7	42.4	50.2	60.1
1980	13.9	64.3	6.1	11.7	17.1	22.5	28.0	33.7	39.9	46.8	55.2
1981	12.4	59.9	5.5	10.4	15.2	19.9	24.7	29.8	35.4	41.6	49.3
1982	12.5	65.4	5.9	10.8	15.6	20.4	25.3	30.6	36.4	43.4	52.8
1983	10.6	50.4	4.8	9.0	13.1	17.2	21.4	25.8	30.5	36.0	42.9
1984	5.9	26.8	2.6	5.0	7.3	9.6	11.9	14.3	16.9	19.8	23.4
1985	8.6	43.3	4.1	7.5	10.9	14.2	17.7	21.3	25.4	30.1	36.7
1986	9.0	39.9	4.0	7.7	11.2	14.8	18.4	22.2	26.2	30.6	36.3
1987	9.5	40.3	4.3	8.1	11.9	15.7	19.6	23.6	27.8	32.5	38.3
1988	6.0	26.2	2.7	5.2	7.6	10.0	12.4	14.9	17.6	20.7	24.7
1989	5.1	22.8	2.4	4.5	6.5	8.6	10.6	12.8	15.2	17.9	21.6
1990	5.8	24.8	2.6	5.0	7.3	9.6	12.0	14.4	17.0	19.9	23.7
1991	4.3	17.8	1.9	3.7	5.4	7.1	8.8	10.6	12.5	14.6	17.3
1992	3.9	16.1	1.8	3.3	4.9	6.5	8.1	9.8	11.5	13.4	15.8
1993	2.4	10.0	1.1	2.1	3.0	4.0	5.0	6.0	7.1	8.3	9.8
1994	2.4	10.0	1.1	2.1	3.1	4.0	5.0	6.1	7.1	8.3	9.8
1995	2.6	10.6	1.2	2.2	3.3	4.3	5.4	6.5	7.7	8.9	10.7
1996	5.0	20.4	2.7	4.8	6.9	9.0	11.2	13.5	16.1	19.2	24.6
1997	4.1	16.9	2.4	4.1	5.9	7.7	9.5	11.5	13.7	16.5	21.4
1998	2.4	9.6	1.6	2.6	3.7	4.7	5.9	7.1	8.5	10.5	14.4
1999	5.1	20.8	3.0	5.1	7.3	9.5	11.8	14.3	17.1	20.6	27.0
2000	3.8	15.6	2.8	4.4	6.2	7.9	9.8	11.9	14.4	17.7	24.8
2001	3.7	14.9	3.1	4.8	6.5	8.3	10.3	12.5	15.2	19.3	28.3
2002	3.0	12.1	2.7	4.1	5.5	7.1	8.7	10.6	12.9	16.4	24.6
2003	4.1	16.8	3.1	4.9	6.8	8.7	10.7	13.0	15.8	19.6	27.6
2004	4.0	16.6	3.1	4.9	6.7	8.6	10.6	12.9	15.6	19.4	27.5
2005	4.2	17.2	3.1	5.0	6.9	8.9	11.0	13.3	16.1	20.0	28.2
2006	5.0	20.7	4.0	6.2	8.6	11.0	13.5	16.4	20.0	25.0	36.0
2007	4.6	19.0	3.3	5.3	7.4	9.5	11.8	14.3	17.2	21.3	29.7
2008	5.4	23.6	3.6	5.9	8.3	10.7	13.2	16.0	19.3	23.8	32.6
2009	2.5	10.2	2.1	3.3	4.5	5.7	7.0	8.5	10.4	13.1	19.2
2010	1.3	5.1	1.6	2.2	2.9	3.6	4.4	5.4	6.7	8.8	14.0
2011	4.4	19.8	3.0	4.9	6.8	8.7	10.8	13.1	15.8	19.5	27.0
2012	1.4	5.6	1.6	2.3	3.0	3.8	4.6	5.7	7.0	9.1	14.4
2013	3.7	15.3	2.9	4.5	6.2	8.0	9.9	12.0	14.6	18.2	25.9

E.3 Specification as in Maasoumi and Wang (2019)

The main specification difference relative to Maasoumi and Wang (2019) is the inclusion of the covariate number of children in the quantile regressions. Tables 20-26 show the estimates when this variable is not included. Relative to the baseline specification, the distributions of actual earnings for participants are almost unaltered for both genders. However, the decomposition presents some differences. Specifically, the coefficients components has been between 0.05 and 0.15 points larger than in the baseline specification, presenting some volatility across time. In contrast, the selection and participation components are smaller. The former accounts to about two thirds of the magnitude of the change in the coefficients component, whereas the latter accounts for approximately the remaining third. Finally, the endowments component has remained almost unaltered.

Table 20: Actual earnings distributions for participants by gender (Frank copula)

	Male						Female					
Year	Mean	P10	P25	P50	P75	P90	Mean	P10	P25	P50	P75	P90
1976	2.72	2.03	2.40	2.77	3.09	3.35	2.31	1.76	2.01	2.31	2.61	2.87
1977	2.73	2.03	2.40	2.78	3.10	3.37	2.33	1.79	2.04	2.33	2.63	2.89
1978	2.73	2.03	2.40	2.78	3.12	3.38	2.33	1.78	2.04	2.33	2.64	2.90
1979	2.75	2.05	2.40	2.79	3.13	3.39	2.34	1.80	2.05	2.34	2.64	2.91
1980	2.73	2.03	2.39	2.78	3.12	3.38	2.33	1.80	2.05	2.33	2.63	2.90
1981	2.68	1.97	2.34	2.73	3.07	3.34	2.30	1.76	2.01	2.30	2.60	2.87
1982	2.67	1.95	2.31	2.71	3.06	3.34	2.28	1.73	1.99	2.28	2.59	2.87
1983	2.65	1.91	2.28	2.70	3.07	3.36	2.29	1.72	1.99	2.29	2.61	2.89
1984	2.65	1.90	2.28	2.70	3.07	3.38	2.31	1.71	1.99	2.32	2.64	2.92
1985	2.66	1.88	2.28	2.71	3.08	3.38	2.32	1.70	1.99	2.33	2.66	2.94
1986	2.66	1.88	2.28	2.71	3.09	3.39	2.33	1.70	2.00	2.34	2.68	2.97
1987	2.68	1.89	2.29	2.73	3.11	3.42	2.36	1.71	2.02	2.37	2.72	3.02
1988	2.68	1.89	2.29	2.72	3.11	3.42	2.37	1.70	2.03	2.38	2.73	3.03
1989	2.68	1.90	2.30	2.73	3.11	3.43	2.37	1.71	2.02	2.38	2.74	3.04
1990	2.67	1.89	2.29	2.71	3.10	3.43	2.37	1.70	2.03	2.39	2.75	3.06
1991	2.64	1.86	2.25	2.67	3.07	3.40	2.37	1.71	2.02	2.38	2.74	3.05
1992	2.62	1.83	2.23	2.67	3.06	3.39	2.37	1.70	2.01	2.38	2.74	3.06
1993	2.62	1.83	2.22	2.66	3.06	3.39	2.38	1.70	2.01	2.39	2.76	3.07
1994	2.60	1.80	2.19	2.64	3.04	3.39	2.37	1.67	2.01	2.39	2.76	3.08
1995	2.60	1.80	2.20	2.64	3.05	3.41	2.37	1.66	1.99	2.38	2.76	3.10
1996	2.61	1.80	2.19	2.63	3.04	3.39	2.36	1.65	1.99	2.37	2.76	3.10
1997	2.61	1.82	2.20	2.63	3.04	3.39	2.37	1.67	2.00	2.38	2.77	3.11
1998	2.63	1.84	2.23	2.66	3.06	3.43	2.40	1.71	2.02	2.41	2.78	3.11
1999	2.67	1.87	2.26	2.68	3.08	3.45	2.42	1.72	2.04	2.43	2.81	3.15
2000	2.68	1.87	2.26	2.69	3.10	3.48	2.44	1.72	2.06	2.45	2.83	3.17
2001	2.70	1.90	2.27	2.70	3.11	3.51	2.46	1.75	2.08	2.47	2.85	3.19
2002	2.70	1.90	2.28	2.70	3.11	3.52	2.48	1.76	2.10	2.48	2.86	3.21
2003	2.70	1.90	2.28	2.70	3.12	3.52	2.49	1.77	2.11	2.49	2.88	3.23
2004	2.69	1.88	2.26	2.69	3.12	3.51	2.49	1.77	2.11	2.50	2.89	3.25
2005	2.68	1.88	2.25	2.69	3.11	3.51	2.48	1.75	2.10	2.49	2.88	3.24
2006	2.68	1.88	2.24	2.67	3.10	3.50	2.47	1.73	2.08	2.47	2.86	3.23
2007	2.67	1.87	2.24	2.67	3.10	3.51	2.48	1.73	2.08	2.48	2.89	3.27
2008	2.68	1.88	2.25	2.67	3.10	3.50	2.49	1.76	2.10	2.50	2.89	3.26
2009	2.66	1.85	2.23	2.66	3.10	3.52	2.47	1.73	2.08	2.47	2.88	3.25
2010	2.67	1.87	2.25	2.68	3.12	3.54	2.48	1.74	2.09	2.49	2.89	3.28
2011	2.66	1.85	2.22	2.67	3.10	3.51	2.47	1.72	2.07	2.48	2.89	3.28
2012	2.64	1.82	2.20	2.65	3.10	3.52	2.46	1.71	2.06	2.47	2.87	3.27
2013	2.64	1.80	2.19	2.64	3.09	3.50	2.46	1.70	2.06	2.47	2.88	3.27

Table 21: Mean decomposition, actual earnings distributions for participants by gender (Frank copula)

Year	Total	EC	CC	SC	PC	Year	Total	EC	CC	SC	PC
1976	0.41	0.01	0.32	0.00	0.08	1995	0.24	0.00	0.18	0.07	-0.02
1977	0.40	0.01	0.30	0.02	0.07	1996	0.24	0.00	0.29	-0.03	-0.02
1978	0.40	0.01	0.36	-0.01	0.04	1997	0.24	0.00	0.27	-0.01	-0.01
1979	0.41	0.01	0.38	-0.02	0.03	1998	0.23	-0.01	0.14	0.11	-0.01
1980	0.40	0.01	0.37	-0.02	0.04	1999	0.25	0.00	0.25	0.01	0.00
1981	0.38	0.00	0.35	-0.01	0.04	2000	0.24	-0.01	0.24	0.00	0.01
1982	0.39	0.01	0.34	0.01	0.03	2001	0.23	-0.01	0.25	-0.02	0.01
1983	0.36	0.01	0.30	0.04	0.02	2002	0.22	-0.01	0.20	0.02	0.01
1984	0.34	0.01	0.18	0.12	0.04	2003	0.21	-0.01	0.28	-0.04	-0.02
1985	0.34	0.01	0.29	0.01	0.02	2004	0.20	-0.02	0.29	-0.06	-0.02
1986	0.33	0.01	0.30	0.00	0.01	2005	0.20	-0.02	0.30	-0.06	-0.02
1987	0.32	0.01	0.35	-0.04	0.00	2006	0.21	-0.02	0.34	-0.10	0.00
1988	0.31	0.01	0.21	0.07	0.02	2007	0.20	-0.03	0.33	-0.08	-0.03
1989	0.31	0.01	0.19	0.09	0.01	2008	0.19	-0.03	0.31	-0.07	-0.02
1990	0.30	0.01	0.28	0.02	-0.01	2009	0.19	-0.03	0.27	-0.02	-0.02
1991	0.27	0.01	0.19	0.05	0.01	2010	0.19	-0.03	0.17	0.08	-0.02
1992	0.25	0.01	0.19	0.05	0.01	2011	0.19	-0.03	0.33	-0.08	-0.03
1993	0.24	0.00	0.13	0.10	0.00	2012	0.18	-0.03	0.20	0.04	-0.03
1994	0.23	0.00	0.16	0.07	-0.01	2013	0.18	-0.03	0.33	-0.10	-0.03

Notes: Total, EC and CC respectively denote total difference, endowments component and coefficients component.

Table 22: 10th percentile decomposition, actual earnings for participants(Frank copula)

Year	Total	EC	CC	SC	PC	Year	Total	EC	CC	SC	PC
1976	0.27	-0.01	0.22	0.00	0.06	1995	0.15	-0.02	0.09	0.10	-0.02
1977	0.24	-0.01	0.19	0.02	0.05	1996	0.15	-0.02	0.21	-0.03	-0.01
1978	0.25	-0.01	0.23	-0.01	0.04	1997	0.16	-0.02	0.21	-0.01	-0.01
1979	0.25	-0.01	0.24	-0.01	0.03	1998	0.13	-0.03	0.02	0.15	-0.01
1980	0.23	-0.01	0.22	-0.01	0.03	1999	0.15	-0.02	0.18	0.01	-0.01
1981	0.21	-0.01	0.21	-0.01	0.03	2000	0.16	-0.02	0.18	-0.01	0.01
1982	0.22	-0.01	0.19	0.01	0.03	2001	0.15	-0.03	0.19	-0.02	0.01
1983	0.19	-0.01	0.15	0.05	0.01	2002	0.14	-0.03	0.13	0.03	0.01
1984	0.19	-0.01	0.04	0.14	0.03	2003	0.13	-0.03	0.22	-0.04	-0.02
1985	0.18	-0.01	0.16	0.01	0.02	2004	0.11	-0.04	0.24	-0.07	-0.02
1986	0.17	-0.01	0.18	0.00	0.01	2005	0.13	-0.04	0.25	-0.07	-0.02
1987	0.18	-0.02	0.24	-0.05	0.00	2006	0.14	-0.04	0.28	-0.10	0.00
1988	0.19	-0.01	0.09	0.09	0.02	2007	0.14	-0.04	0.30	-0.09	-0.03
1989	0.20	-0.01	0.07	0.12	0.01	2008	0.12	-0.04	0.27	-0.09	-0.02
1990	0.19	-0.01	0.19	0.03	-0.02	2009	0.12	-0.05	0.22	-0.03	-0.03
1991	0.15	-0.02	0.09	0.07	0.01	2010	0.13	-0.04	0.07	0.13	-0.02
1992	0.13	-0.02	0.08	0.06	0.01	2011	0.12	-0.04	0.31	-0.11	-0.03
1993	0.13	-0.02	0.01	0.14	0.00	2012	0.11	-0.04	0.12	0.07	-0.04
1994	0.13	-0.02	0.05	0.10	-0.01	2013	0.10	-0.04	0.29	-0.13	-0.03

Notes: Total, EC and CC respectively denote total difference, endowments component and coefficients component.

Table 23: 25th percentile decomposition, actual earnings for participants(Frank copula)

Year	Total	EC	CC	SC	PC	Year	Total	EC	CC	SC	PC
1976	0.38	0.00	0.30	0.00	0.08	1995	0.21	-0.02	0.15	0.10	-0.02
1977	0.36	0.00	0.28	0.02	0.06	1996	0.20	-0.02	0.27	-0.03	-0.02
1978	0.36	0.00	0.33	-0.01	0.04	1997	0.20	-0.02	0.25	-0.01	-0.02
1979	0.35	0.00	0.33	-0.02	0.03	1998	0.21	-0.02	0.09	0.15	-0.01
1980	0.35	0.00	0.33	-0.02	0.04	1999	0.22	-0.02	0.23	0.01	-0.01
1981	0.33	-0.01	0.31	-0.01	0.04	2000	0.21	-0.02	0.22	-0.01	0.01
1982	0.33	0.00	0.29	0.01	0.03	2001	0.19	-0.02	0.22	-0.02	0.01
1983	0.30	-0.01	0.24	0.05	0.02	2002	0.18	-0.02	0.17	0.03	0.01
1984	0.29	0.00	0.11	0.14	0.04	2003	0.17	-0.03	0.26	-0.04	-0.02
1985	0.29	0.00	0.25	0.02	0.03	2004	0.15	-0.04	0.27	-0.07	-0.02
1986	0.28	0.00	0.27	0.00	0.01	2005	0.16	-0.04	0.28	-0.07	-0.02
1987	0.27	-0.01	0.32	-0.04	0.00	2006	0.16	-0.04	0.31	-0.11	0.00
1988	0.26	0.00	0.16	0.09	0.02	2007	0.16	-0.04	0.32	-0.09	-0.03
1989	0.28	0.00	0.14	0.12	0.02	2008	0.14	-0.04	0.29	-0.08	-0.02
1990	0.25	0.00	0.25	0.03	-0.02	2009	0.15	-0.05	0.25	-0.03	-0.03
1991	0.22	-0.01	0.16	0.07	0.01	2010	0.16	-0.04	0.12	0.11	-0.03
1992	0.22	-0.01	0.16	0.06	0.01	2011	0.15	-0.04	0.32	-0.09	-0.04
1993	0.21	-0.01	0.09	0.13	0.00	2012	0.15	-0.04	0.17	0.06	-0.04
1994	0.19	-0.02	0.12	0.09	-0.01	2013	0.13	-0.04	0.32	-0.12	-0.03

Notes: Total, EC and CC respectively denote total difference, endowments component and coefficients component.

Table 24: 50th percentile decomposition, actual earnings for participants(Frank copula)

Year	Total	EC	CC	SC	PC	Year	Total	EC	CC	SC	PC
1976	0.46	0.01	0.34	0.00	0.09	1995	0.26	0.00	0.20	0.08	-0.02
1977	0.44	0.01	0.34	0.02	0.08	1996	0.25	0.00	0.30	-0.03	-0.02
1978	0.45	0.01	0.40	-0.01	0.05	1997	0.25	-0.01	0.29	-0.01	-0.02
1979	0.45	0.01	0.42	-0.02	0.04	1998	0.25	-0.01	0.15	0.11	-0.01
1980	0.45	0.01	0.41	-0.02	0.05	1999	0.25	-0.01	0.26	0.01	-0.01
1981	0.43	0.01	0.40	-0.01	0.05	2000	0.25	-0.01	0.25	0.00	0.01
1982	0.43	0.01	0.38	0.01	0.04	2001	0.23	-0.01	0.25	-0.02	0.01
1983	0.40	0.01	0.33	0.05	0.02	2002	0.22	-0.01	0.20	0.03	0.01
1984	0.38	0.01	0.20	0.13	0.04	2003	0.21	-0.02	0.28	-0.04	-0.02
1985	0.38	0.01	0.33	0.02	0.03	2004	0.19	-0.02	0.30	-0.06	-0.02
1986	0.37	0.01	0.35	0.00	0.01	2005	0.20	-0.02	0.30	-0.06	-0.02
1987	0.36	0.01	0.40	-0.05	0.00	2006	0.20	-0.03	0.34	-0.11	0.00
1988	0.34	0.01	0.23	0.08	0.02	2007	0.19	-0.03	0.33	-0.09	-0.03
1989	0.34	0.01	0.21	0.10	0.02	2008	0.18	-0.04	0.31	-0.07	-0.02
1990	0.32	0.01	0.30	0.02	-0.02	2009	0.19	-0.04	0.28	-0.03	-0.03
1991	0.29	0.00	0.22	0.06	0.01	2010	0.19	-0.04	0.18	0.08	-0.03
1992	0.29	0.01	0.22	0.06	0.01	2011	0.18	-0.04	0.33	-0.08	-0.03
1993	0.27	0.00	0.15	0.11	0.00	2012	0.18	-0.03	0.21	0.05	-0.04
1994	0.25	0.00	0.18	0.08	-0.01	2013	0.17	-0.04	0.34	-0.10	-0.03

Notes: Total, EC and CC respectively denote total difference, endowments component and coefficients component.

Table 25: 75th percentile decomposition, actual earnings for participants(Frank copula)

Year	Total	EC	CC	SC	PC	Year	Total	EC	CC	SC	PC
1976	0.48	0.02	0.35	0.00	0.10	1995	0.29	0.02	0.23	0.06	-0.02
1977	0.47	0.02	0.36	0.02	0.08	1996	0.28	0.01	0.31	-0.03	-0.02
1978	0.48	0.02	0.42	-0.01	0.05	1997	0.27	0.01	0.29	-0.01	-0.02
1979	0.49	0.02	0.45	-0.02	0.04	1998	0.28	0.01	0.19	0.09	-0.01
1980	0.49	0.02	0.44	-0.02	0.05	1999	0.27	0.01	0.26	0.01	0.00
1981	0.48	0.02	0.43	-0.02	0.05	2000	0.28	0.01	0.27	0.00	0.01
1982	0.47	0.02	0.40	0.01	0.04	2001	0.26	0.00	0.27	-0.02	0.01
1983	0.46	0.02	0.38	0.04	0.02	2002	0.26	0.00	0.22	0.02	0.01
1984	0.44	0.02	0.26	0.11	0.04	2003	0.24	0.00	0.29	-0.03	-0.02
1985	0.42	0.02	0.36	0.02	0.03	2004	0.23	-0.01	0.31	-0.05	-0.02
1986	0.41	0.03	0.37	0.00	0.01	2005	0.22	-0.01	0.31	-0.05	-0.02
1987	0.39	0.03	0.41	-0.04	0.00	2006	0.24	-0.01	0.36	-0.11	0.00
1988	0.39	0.03	0.28	0.07	0.02	2007	0.21	-0.02	0.34	-0.08	-0.02
1989	0.37	0.02	0.25	0.08	0.02	2008	0.21	-0.02	0.32	-0.07	-0.02
1990	0.35	0.03	0.32	0.02	-0.01	2009	0.22	-0.03	0.29	-0.02	-0.03
1991	0.33	0.02	0.25	0.05	0.01	2010	0.22	-0.03	0.21	0.06	-0.03
1992	0.32	0.02	0.25	0.05	0.01	2011	0.21	-0.03	0.34	-0.07	-0.03
1993	0.30	0.02	0.20	0.09	0.00	2012	0.22	-0.03	0.25	0.03	-0.04
1994	0.29	0.02	0.22	0.06	-0.01	2013	0.21	-0.03	0.35	-0.09	-0.03

Notes: Total, EC and CC respectively denote total difference, endowments component and coefficients component.

Table 26: 90th percentile decomposition, actual earnings for participants(Frank copula)

Year	Total	EC	CC	SC	PC	Year	Total	EC	CC	SC	PC
1976	0.49	0.04	0.36	0.00	0.09	1995	0.31	0.03	0.24	0.05	-0.02
1977	0.48	0.03	0.36	0.02	0.07	1996	0.30	0.02	0.31	-0.02	-0.01
1978	0.48	0.03	0.42	-0.01	0.05	1997	0.29	0.02	0.29	-0.01	-0.01
1979	0.49	0.03	0.44	-0.02	0.03	1998	0.32	0.02	0.23	0.07	0.00
1980	0.48	0.03	0.43	-0.02	0.04	1999	0.30	0.02	0.28	0.01	0.00
1981	0.47	0.03	0.41	-0.02	0.05	2000	0.31	0.02	0.29	0.00	0.01
1982	0.47	0.03	0.40	0.01	0.03	2001	0.32	0.02	0.31	-0.02	0.01
1983	0.47	0.04	0.38	0.04	0.02	2002	0.31	0.01	0.26	0.02	0.01
1984	0.46	0.04	0.28	0.10	0.04	2003	0.29	0.01	0.32	-0.03	-0.02
1985	0.44	0.04	0.36	0.01	0.02	2004	0.27	0.00	0.34	-0.06	-0.02
1986	0.41	0.04	0.36	0.00	0.01	2005	0.27	0.00	0.34	-0.05	-0.02
1987	0.40	0.04	0.40	-0.04	0.00	2006	0.27	0.00	0.40	-0.12	0.00
1988	0.39	0.04	0.28	0.06	0.01	2007	0.25	-0.01	0.36	-0.07	-0.03
1989	0.39	0.04	0.27	0.07	0.01	2008	0.24	-0.01	0.33	-0.07	-0.02
1990	0.36	0.04	0.32	0.01	-0.01	2009	0.27	-0.01	0.32	-0.02	-0.02
1991	0.35	0.03	0.26	0.05	0.01	2010	0.26	-0.02	0.25	0.05	-0.03
1992	0.33	0.03	0.25	0.04	0.01	2011	0.23	-0.02	0.35	-0.06	-0.03
1993	0.32	0.04	0.21	0.07	0.00	2012	0.26	-0.02	0.28	0.03	-0.03
1994	0.31	0.03	0.24	0.05	-0.01	2013	0.23	-0.02	0.36	-0.09	-0.03

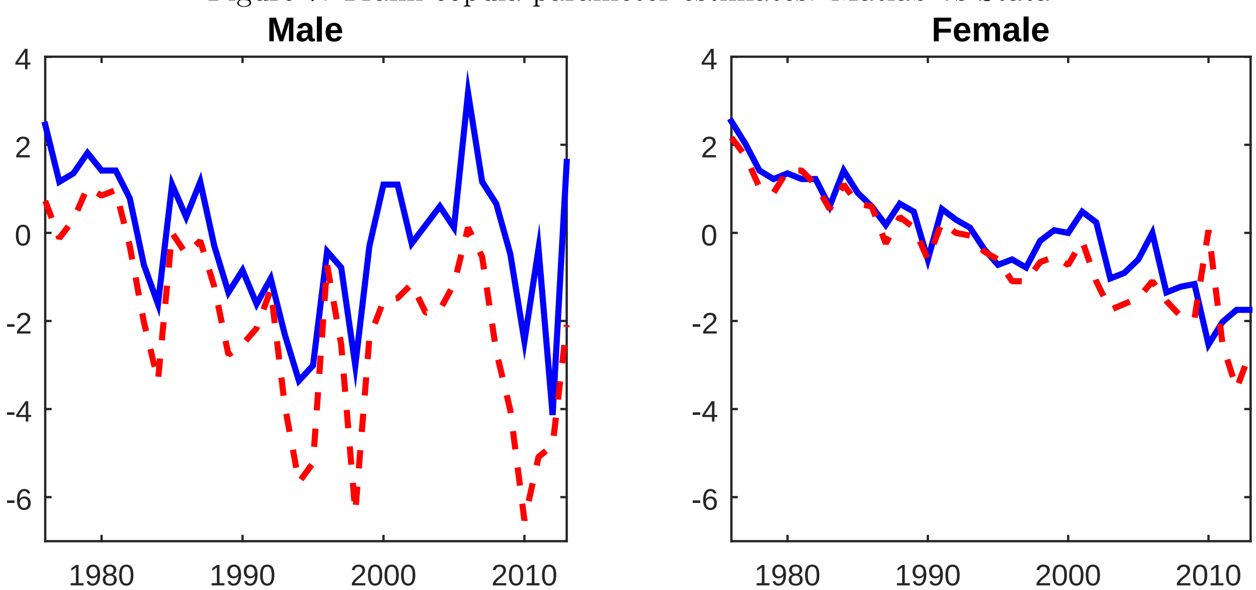
Notes: Total, EC and CC respectively denote total difference, endowments component and coefficients component.

E.4 Stata vs Matlab estimates

The QRS estimates presented in this paper differ from those in Maasoumi and Wang (2019). As mentioned in Section 6, this is partly due to some differences in the implementation of the estimator. However, an important factor to explain the differences in the results is the worse performance of Stata relative to Matlab.[29] To see this, consider the estimation of the QRS estimates using the exact specification and implementation used by Maasoumi and Wang (2019) in both statistical packages.

Both sets of estimates of the Frank copula for both genders are shown in Figure 7. The original estimates are slightly smaller in general, which means that the estimated amount of self-selection is slightly more positive.

Figure 7: Frank copula parameter estimates: Matlab vs Stata



Notes: the solid thick blue lines denote the Matlab estimates and the dashed red lines denote the Stata estimates.

To understand why the two sets of estimates are different, let us focus on the estimates for male workers in 1976. These estimates minimize the value of the criterion function defined in Equation 18, which in turn depend on the estimates $\hat{\beta}_d(\tau; t)$ given by Equation 17. Fix the

[29]Specifically, the Stata estimates are obtained with the codes available at `https://www.journals.uchicago.edu/doi/suppl/10.1086/701788/suppl_file/2012616data.zip`, which are based on the Stata package qreg, while the Matlab estimates are obtained using the rq.m code available at `http://www.econ.uiuc.edu/~roger/research/rq/rq.m`.

value of the Frank copula to be the first one considered in the estimation grid: $t = -42.889$. The minimized values of Equation 17 for the quantiles used in the estimation of θ_d for both sets of estimates are shown in Table 27. These are larger for the Stata estimates than for the Matlab ones. As such, when the slope parameter estimates are plugged into Equation 18, this is different for the estimates with both statistical packages. As a consequence, the estimates of the copula parameters are different.

Table 27: Minimized value of check function

	τ				
	0.3	0.4	0.5	0.6	0.7
Matlab estimates	1216.086	2044.821	2672.171	3051.696	3090.995
Stata estimates	1216.519	2045.354	2673.005	3052.189	3091.509

Notes: male sample, year 1976, Frank copula with $t = -42.889$.

One final aspect regards the computational time required to obtain the estimates with each statistical package. Whereas those in Stata are obtained using the simplex method, the Matlab estimates are based on the interior point method, which are substantially faster (Portnoy and Koenker, 1997; Chernozhukov et al., 2022). Therefore, this allows to obtain a larger set of estimates, including more flexible models, such as those with heterogeneous copulas.

F Additional Tables and Figures

Table 28: Kendall's τ correlation coefficients for the Gaussian copula

	QRS		H2S			QRS		H2S	
Year	Male	Female	Male	Female	Year	Male	Female	Male	Female
1976	0.26	0.13	0.31	0.10	1995	-0.34	-0.22	-0.28	-0.22
1977	0.13	0.08	0.17	0.07	1996	0.04	-0.21	0.06	-0.22
1978	0.23	-0.01	0.24	-0.02	1997	-0.01	-0.21	-0.10	-0.17
1979	0.21	-0.02	0.22	-0.02	1998	-0.27	-0.21	-0.34	-0.23
1980	0.24	0.01	0.22	-0.01	1999	-0.06	-0.22	-0.01	-0.26
1981	0.19	0.01	0.22	0.02	2000	0.07	-0.15	0.06	-0.10
1982	0.13	-0.04	0.14	-0.02	2001	0.13	-0.13	0.13	-0.14
1983	-0.05	-0.09	-0.01	-0.10	2002	0.03	-0.14	-0.01	-0.11
1984	-0.15	0.02	-0.08	0.01	2003	0.04	-0.25	0.02	-0.23
1985	0.10	-0.05	0.09	-0.05	2004	0.04	-0.22	0.03	-0.21
1986	0.08	-0.10	0.10	-0.10	2005	0.07	-0.22	0.08	-0.25
1987	0.12	-0.18	0.06	-0.19	2006	0.30	-0.17	0.19	-0.18
1988	-0.09	-0.10	-0.04	-0.10	2007	0.08	-0.30	0.03	-0.28
1989	-0.17	-0.10	-0.12	-0.03	2008	0.10	-0.33	0.06	-0.37
1990	-0.09	-0.22	-0.09	-0.22	2009	-0.15	-0.26	-0.14	-0.26
1991	-0.09	-0.10	-0.07	-0.11	2010	-0.34	-0.28	-0.23	-0.25
1992	-0.11	-0.12	-0.07	-0.10	2011	-0.10	-0.33	-0.10	-0.34
1993	-0.31	-0.14	-0.22	-0.13	2012	-0.33	-0.29	-0.32	-0.36
1994	-0.24	-0.17	-0.24	-0.19	2013	0.01	-0.26	-0.05	-0.31

Notes: Kendall's τ correlation coefficients of the copula estimates by year and gender. QRS and H2S respectively denote the estimates of the quantile regression with selection and Heckman 2-stage estimators.

Table 29: 10th percentile decomposition, actual earnings for participants (Frank copula)

Year	Total	EC	CC	SC	PC	Year	Total	EC	CC	SC	PC
1976	0.29	-0.01	0.30	-0.03	0.04	1995	0.16	-0.03	0.20	0.04	-0.05
1977	0.26	-0.02	0.26	-0.02	0.02	1996	0.16	-0.03	0.34	-0.11	-0.05
1978	0.27	-0.01	0.35	-0.06	0.00	1997	0.17	-0.03	0.31	-0.08	-0.04
1979	0.27	-0.01	0.35	-0.07	-0.01	1998	0.14	-0.03	0.18	0.04	-0.05
1980	0.25	-0.01	0.32	-0.06	0.00	1999	0.17	-0.03	0.37	-0.12	-0.05
1981	0.23	-0.02	0.29	-0.05	0.00	2000	0.16	-0.03	0.30	-0.08	-0.03
1982	0.23	-0.01	0.31	-0.06	-0.01	2001	0.16	-0.03	0.31	-0.09	-0.02
1983	0.21	-0.02	0.30	-0.05	-0.02	2002	0.14	-0.03	0.27	-0.07	-0.03
1984	0.21	-0.02	0.16	0.06	0.00	2003	0.13	-0.03	0.35	-0.13	-0.05
1985	0.20	-0.02	0.27	-0.05	-0.01	2004	0.12	-0.04	0.35	-0.14	-0.05
1986	0.18	-0.02	0.31	-0.08	-0.03	2005	0.13	-0.04	0.37	-0.15	-0.05
1987	0.19	-0.02	0.38	-0.12	-0.05	2006	0.15	-0.04	0.40	-0.17	-0.03
1988	0.20	-0.01	0.26	-0.02	-0.03	2007	0.14	-0.04	0.43	-0.18	-0.07
1989	0.21	-0.01	0.23	0.01	-0.02	2008	0.12	-0.05	0.46	-0.21	-0.08
1990	0.20	-0.01	0.32	-0.04	-0.07	2009	0.12	-0.05	0.31	-0.08	-0.05
1991	0.16	-0.02	0.22	-0.01	-0.03	2010	0.14	-0.05	0.21	0.01	-0.04
1992	0.15	-0.02	0.20	-0.01	-0.03	2011	0.13	-0.05	0.45	-0.21	-0.07
1993	0.14	-0.02	0.11	0.09	-0.03	2012	0.11	-0.04	0.21	0.00	-0.06
1994	0.14	-0.03	0.15	0.05	-0.04	2013	0.10	-0.05	0.38	-0.19	-0.05

Notes: Total, EC, CC, SC and PC respectively denote total difference, endowments component, coefficients component, selection component and participation component.

Table 30: 25th percentile decomposition, actual earnings for participants (Frank copula)

Year	Total	EC	CC	SC	PC	Year	Total	EC	CC	SC	PC
1976	0.39	0.00	0.38	-0.03	0.04	1995	0.21	-0.02	0.26	0.03	-0.06
1977	0.37	-0.01	0.36	-0.02	0.03	1996	0.20	-0.02	0.39	-0.11	-0.06
1978	0.37	0.00	0.44	-0.06	0.00	1997	0.21	-0.02	0.36	-0.08	-0.05
1979	0.36	0.00	0.43	-0.07	-0.01	1998	0.21	-0.03	0.25	0.04	-0.05
1980	0.36	-0.01	0.42	-0.06	0.00	1999	0.22	-0.02	0.42	-0.12	-0.06
1981	0.34	-0.01	0.39	-0.05	0.00	2000	0.21	-0.02	0.35	-0.08	-0.03
1982	0.34	-0.01	0.41	-0.06	-0.01	2001	0.19	-0.03	0.34	-0.09	-0.02
1983	0.31	-0.01	0.39	-0.05	-0.03	2002	0.18	-0.02	0.31	-0.07	-0.03
1984	0.29	-0.01	0.24	0.05	0.00	2003	0.17	-0.03	0.39	-0.13	-0.06
1985	0.29	-0.01	0.36	-0.05	-0.01	2004	0.15	-0.04	0.39	-0.14	-0.05
1986	0.29	-0.01	0.40	-0.08	-0.03	2005	0.16	-0.04	0.40	-0.15	-0.05
1987	0.28	-0.01	0.46	-0.12	-0.05	2006	0.16	-0.04	0.43	-0.19	-0.04
1988	0.27	0.00	0.32	-0.02	-0.03	2007	0.16	-0.04	0.45	-0.18	-0.07
1989	0.28	0.00	0.30	0.01	-0.02	2008	0.15	-0.05	0.47	-0.21	-0.07
1990	0.26	-0.01	0.37	-0.04	-0.07	2009	0.16	-0.05	0.33	-0.07	-0.05
1991	0.23	-0.01	0.28	-0.01	-0.03	2010	0.16	-0.05	0.25	0.01	-0.05
1992	0.23	-0.01	0.28	-0.01	-0.03	2011	0.15	-0.05	0.43	-0.17	-0.06
1993	0.21	-0.01	0.18	0.08	-0.03	2012	0.15	-0.05	0.25	0.00	-0.06
1994	0.19	-0.02	0.20	0.05	-0.04	2013	0.13	-0.05	0.41	-0.17	-0.05

Notes: Total, EC, CC, SC and PC respectively denote total difference, endowments component, coefficients component, selection component and participation component.

Table 31: 50th percentile decomposition, actual earnings for participants (Frank copula)

Year	Total	EC	CC	SC	PC	Year	Total	EC	CC	SC	PC
1976	0.47	0.01	0.44	-0.03	0.05	1995	0.26	0.00	0.29	0.03	-0.05
1977	0.45	0.01	0.43	-0.02	0.03	1996	0.26	-0.01	0.41	-0.10	-0.05
1978	0.46	0.01	0.52	-0.06	0.00	1997	0.25	-0.01	0.38	-0.07	-0.05
1979	0.46	0.01	0.53	-0.07	-0.01	1998	0.25	-0.01	0.28	0.03	-0.05
1980	0.46	0.01	0.52	-0.06	0.00	1999	0.25	-0.01	0.42	-0.10	-0.05
1981	0.44	0.00	0.49	-0.05	0.00	2000	0.25	-0.01	0.36	-0.07	-0.03
1982	0.44	0.01	0.50	-0.06	-0.01	2001	0.23	-0.01	0.35	-0.08	-0.02
1983	0.41	0.00	0.48	-0.05	-0.03	2002	0.22	-0.01	0.32	-0.06	-0.03
1984	0.39	0.01	0.33	0.05	0.00	2003	0.21	-0.02	0.39	-0.11	-0.05
1985	0.39	0.01	0.44	-0.05	-0.01	2004	0.19	-0.03	0.39	-0.13	-0.05
1986	0.37	0.01	0.47	-0.08	-0.03	2005	0.20	-0.03	0.40	-0.13	-0.05
1987	0.36	0.01	0.52	-0.12	-0.05	2006	0.20	-0.03	0.44	-0.18	-0.03
1988	0.34	0.01	0.37	-0.01	-0.03	2007	0.19	-0.04	0.43	-0.15	-0.06
1989	0.35	0.01	0.35	0.01	-0.02	2008	0.18	-0.04	0.45	-0.18	-0.06
1990	0.32	0.01	0.40	-0.03	-0.06	2009	0.19	-0.04	0.34	-0.06	-0.05
1991	0.30	0.00	0.33	-0.01	-0.03	2010	0.19	-0.04	0.27	0.01	-0.04
1992	0.29	0.00	0.32	-0.01	-0.03	2011	0.18	-0.04	0.42	-0.14	-0.06
1993	0.27	0.00	0.24	0.06	-0.03	2012	0.18	-0.04	0.28	0.00	-0.06
1994	0.25	0.00	0.25	0.04	-0.04	2013	0.17	-0.04	0.41	-0.15	-0.05

Notes: Total, EC, CC, SC and PC respectively denote total difference, endowments component, coefficients component, selection component and participation component.

Table 32: 75th percentile decomposition, actual earnings for participants (Frank copula)

Year	Total	EC	CC	SC	PC	Year	Total	EC	CC	SC	PC
1976	0.49	0.02	0.45	-0.04	0.05	1995	0.29	0.02	0.30	0.02	-0.04
1977	0.48	0.02	0.45	-0.02	0.03	1996	0.28	0.01	0.40	-0.08	-0.04
1978	0.49	0.02	0.53	-0.06	0.00	1997	0.28	0.01	0.36	-0.06	-0.04
1979	0.50	0.02	0.56	-0.07	-0.01	1998	0.28	0.00	0.29	0.02	-0.04
1980	0.50	0.02	0.55	-0.07	0.00	1999	0.27	0.01	0.39	-0.08	-0.04
1981	0.48	0.01	0.52	-0.06	0.00	2000	0.27	0.00	0.36	-0.06	-0.03
1982	0.48	0.02	0.53	-0.06	-0.01	2001	0.26	0.00	0.36	-0.08	-0.02
1983	0.46	0.02	0.51	-0.04	-0.02	2002	0.26	0.00	0.33	-0.05	-0.02
1984	0.44	0.02	0.37	0.04	0.00	2003	0.24	0.00	0.38	-0.09	-0.04
1985	0.43	0.02	0.46	-0.04	-0.01	2004	0.23	-0.01	0.38	-0.10	-0.04
1986	0.41	0.02	0.49	-0.07	-0.03	2005	0.23	-0.01	0.39	-0.11	-0.04
1987	0.40	0.02	0.51	-0.10	-0.04	2006	0.24	-0.02	0.44	-0.16	-0.03
1988	0.39	0.03	0.40	-0.01	-0.02	2007	0.21	-0.02	0.41	-0.12	-0.05
1989	0.37	0.02	0.36	0.01	-0.02	2008	0.21	-0.03	0.44	-0.15	-0.05
1990	0.36	0.02	0.40	-0.02	-0.05	2009	0.22	-0.03	0.34	-0.04	-0.04
1991	0.33	0.02	0.35	-0.01	-0.02	2010	0.22	-0.03	0.28	0.01	-0.04
1992	0.32	0.02	0.33	-0.01	-0.02	2011	0.21	-0.03	0.40	-0.11	-0.05
1993	0.31	0.02	0.27	0.05	-0.02	2012	0.22	-0.03	0.30	0.00	-0.05
1994	0.29	0.02	0.28	0.03	-0.03	2013	0.21	-0.03	0.40	-0.12	-0.04

Notes: Total, EC, CC, SC and PC respectively denote total difference, endowments component, coefficients component, selection component and participation component.

Table 33: 90th percentile decomposition, actual earnings for participants (Frank copula)

Year	Total	EC	CC	SC	PC	Year	Total	EC	CC	SC	PC
1976	0.49	0.03	0.45	-0.04	0.04	1995	0.31	0.03	0.30	0.01	-0.04
1977	0.49	0.03	0.44	-0.02	0.03	1996	0.29	0.02	0.38	-0.07	-0.04
1978	0.49	0.03	0.52	-0.06	0.00	1997	0.29	0.02	0.35	-0.05	-0.03
1979	0.49	0.03	0.54	-0.08	-0.01	1998	0.31	0.02	0.31	0.02	-0.03
1980	0.49	0.03	0.53	-0.07	0.00	1999	0.29	0.02	0.38	-0.07	-0.04
1981	0.47	0.03	0.50	-0.06	0.00	2000	0.31	0.02	0.37	-0.06	-0.02
1982	0.48	0.03	0.51	-0.05	-0.01	2001	0.32	0.02	0.39	-0.07	-0.02
1983	0.48	0.04	0.50	-0.04	-0.02	2002	0.30	0.01	0.36	-0.05	-0.02
1984	0.46	0.04	0.38	0.04	0.00	2003	0.28	0.01	0.39	-0.08	-0.04
1985	0.44	0.04	0.45	-0.04	-0.01	2004	0.27	0.00	0.40	-0.10	-0.04
1986	0.41	0.04	0.46	-0.06	-0.02	2005	0.27	0.00	0.41	-0.10	-0.04
1987	0.40	0.04	0.48	-0.08	-0.04	2006	0.27	0.00	0.47	-0.17	-0.03
1988	0.39	0.04	0.38	-0.01	-0.02	2007	0.25	-0.01	0.42	-0.11	-0.05
1989	0.39	0.04	0.36	0.01	-0.02	2008	0.23	-0.01	0.43	-0.14	-0.05
1990	0.36	0.04	0.38	-0.02	-0.04	2009	0.27	-0.01	0.36	-0.04	-0.04
1991	0.35	0.03	0.35	-0.01	-0.02	2010	0.25	-0.02	0.30	0.01	-0.04
1992	0.33	0.03	0.32	-0.01	-0.02	2011	0.23	-0.02	0.39	-0.10	-0.05
1993	0.32	0.04	0.27	0.04	-0.02	2012	0.25	-0.02	0.31	0.00	-0.05
1994	0.31	0.03	0.28	0.02	-0.03	2013	0.22	-0.02	0.40	-0.11	-0.04

Notes: Total, EC, CC, SC and PC respectively denote total difference, endowments component, coefficients component, selection component and participation component.

Table 34: 10th percentile decomposition, actual earnings for the full population (Frank copula)

Year	Total	EC	CC	SC	PC	Year	Total	EC	CC	SC	PC
1976	0.00	0.00	0.00	0.00	0.00	1995	0.00	0.00	0.00	0.00	0.00
1977	0.00	0.00	0.00	0.00	0.00	1996	0.00	0.00	0.00	0.00	0.00
1978	0.00	0.00	0.00	0.00	0.00	1997	0.00	0.00	0.00	0.00	0.00
1979	0.00	0.00	0.00	0.00	0.00	1998	0.00	0.00	0.00	0.00	0.00
1980	0.00	0.00	0.00	0.00	0.00	1999	0.00	0.00	0.00	0.00	0.00
1981	0.00	0.00	0.00	0.00	0.00	2000	0.00	0.00	0.00	0.00	0.00
1982	0.00	0.00	0.00	0.00	0.00	2001	0.00	0.00	0.00	0.00	0.00
1983	0.00	0.00	0.00	0.00	0.00	2002	0.00	0.00	0.00	0.00	0.00
1984	0.00	0.00	0.00	0.00	0.00	2003	0.00	0.00	0.00	0.00	0.00
1985	0.00	0.00	0.00	0.00	0.00	2004	0.00	0.00	0.00	0.00	0.00
1986	0.00	0.00	0.00	0.00	0.00	2005	0.00	0.00	0.00	0.00	0.00
1987	0.00	0.00	0.00	0.00	0.00	2006	0.00	0.00	0.00	0.00	0.00
1988	0.00	0.00	0.00	0.00	0.00	2007	0.00	0.00	0.00	0.00	0.00
1989	0.00	0.00	0.00	0.00	0.00	2008	0.00	0.00	0.00	0.00	0.00
1990	0.00	0.00	0.00	0.00	0.00	2009	0.00	0.00	0.00	0.00	0.00
1991	0.00	0.00	0.00	0.00	0.00	2010	0.00	0.00	0.00	0.00	0.00
1992	0.00	0.00	0.00	0.00	0.00	2011	0.00	0.00	0.00	0.00	0.00
1993	0.00	0.00	0.00	0.00	0.00	2012	0.00	0.00	0.00	0.00	0.00
1994	0.00	0.00	0.00	0.00	0.00	2013	0.00	0.00	0.00	0.00	0.00

Notes: Total, EC, CC, SC and PC respectively denote total difference, endowments component, coefficients component, selection component and participation component.

Table 35: 25th percentile decomposition, actual earnings for the full population (Frank copula)

Year	Total	EC	CC	SC	PC	Year	Total	EC	CC	SC	PC
1976	0.00	0.00	0.00	0.00	0.00	1995	0.00	0.00	0.00	0.00	0.00
1977	0.00	0.00	0.00	0.00	0.00	1996	0.00	0.00	0.00	0.00	0.00
1978	0.00	0.00	0.00	0.00	0.00	1997	0.00	0.00	0.00	0.00	0.00
1979	0.00	0.00	0.00	0.00	0.00	1998	0.00	0.00	0.00	0.00	0.00
1980	0.00	0.00	0.00	0.00	0.00	1999	0.00	0.00	0.00	0.00	0.00
1981	0.00	0.00	0.00	0.00	0.00	2000	0.00	0.00	0.00	0.00	0.00
1982	0.00	0.00	0.00	0.00	0.00	2001	0.00	0.00	0.00	0.00	0.00
1983	0.00	0.00	0.00	0.00	0.00	2002	0.00	0.00	0.00	0.00	0.00
1984	0.00	0.00	0.00	0.00	0.00	2003	0.00	0.00	0.00	0.00	0.00
1985	0.00	0.00	0.00	0.00	0.00	2004	0.00	0.00	0.00	0.00	0.00
1986	0.00	0.00	0.00	0.00	0.00	2005	0.00	0.00	0.00	0.00	0.00
1987	0.00	0.00	0.00	0.00	0.00	2006	0.00	0.00	0.00	0.00	0.00
1988	0.00	0.00	0.00	0.00	0.00	2007	0.00	0.00	0.00	0.00	0.00
1989	0.00	0.00	0.00	0.00	0.00	2008	0.00	0.00	0.00	0.00	0.00
1990	0.00	0.00	0.00	0.00	0.00	2009	0.00	0.00	0.00	0.00	0.00
1991	0.00	0.00	0.00	0.00	0.00	2010	0.00	0.00	0.00	0.00	0.00
1992	0.00	0.00	0.00	0.00	0.00	2011	0.00	0.00	0.00	0.00	0.00
1993	0.00	0.00	0.00	0.00	0.00	2012	0.00	0.00	0.00	0.00	0.00
1994	0.00	0.00	0.00	0.00	0.00	2013	0.00	0.00	0.00	0.00	0.00

Notes: Total, EC, CC, SC and PC respectively denote total difference, endowments component, coefficients component, selection component and participation component.

Table 36: 50th percentile decomposition, actual earnings for the full population (Frank copula)

Year	Total	EC	CC	SC	PC	Year	Total	EC	CC	SC	PC
1976	2.50	0.00	0.42	-0.03	2.11	1995	2.26	-0.02	0.28	0.03	1.97
1977	2.51	0.00	0.40	-0.02	2.12	1996	2.27	-0.02	0.40	-0.10	1.98
1978	2.50	0.00	0.48	-0.05	2.07	1997	2.28	-0.02	0.37	-0.07	2.00
1979	2.52	0.00	0.48	-0.06	2.10	1998	1.15	-0.03	0.28	0.03	0.87
1980	2.50	0.00	0.46	-0.05	2.09	1999	0.98	-0.02	0.43	-0.10	0.68
1981	2.42	0.00	0.43	-0.05	2.04	2000	0.86	-0.02	0.36	-0.07	0.58
1982	2.34	0.00	0.42	-0.05	1.96	2001	0.80	-0.02	0.34	-0.08	0.56
1983	2.24	0.00	0.40	-0.05	1.89	2002	0.88	-0.02	0.32	-0.06	0.64
1984	2.23	0.00	0.26	0.05	1.92	2003	1.09	-0.03	0.40	-0.12	0.84
1985	2.27	0.01	0.37	-0.04	1.93	2004	2.30	-0.04	0.39	-0.13	2.07
1986	2.30	0.01	0.41	-0.07	1.95	2005	2.29	-0.04	0.41	-0.14	2.06
1987	2.32	0.01	0.47	-0.11	1.95	2006	1.43	-0.04	0.42	-0.16	1.21
1988	2.35	0.01	0.34	-0.02	2.01	2007	1.07	-0.05	0.45	-0.16	0.83
1989	2.36	0.01	0.32	0.01	2.02	2008	0.91	-0.06	0.48	-0.19	0.68
1990	2.38	0.00	0.39	-0.03	2.02	2009	2.25	-0.06	0.34	-0.06	2.03
1991	2.33	-0.01	0.31	-0.01	2.04	2010	2.15	-0.07	0.25	0.01	1.96
1992	2.29	-0.01	0.30	-0.01	2.01	2011	2.11	-0.07	0.43	-0.16	1.92
1993	2.26	-0.01	0.21	0.07	2.00	2012	2.13	-0.07	0.27	0.00	1.93
1994	2.23	-0.02	0.23	0.04	1.98	2013	2.13	-0.07	0.41	-0.15	1.95

Notes: Total, EC, CC, SC and PC respectively denote total difference, endowments component, coefficients component, selection component and participation component.

Table 37: 75th percentile decomposition, actual earnings for the full population (Frank copula)

Year	Total	EC	CC	SC	PC	Year	Total	EC	CC	SC	PC
1976	0.89	0.02	0.45	-0.03	0.44	1995	0.46	0.01	0.30	0.02	0.14
1977	0.86	0.02	0.45	-0.02	0.40	1996	0.45	0.00	0.40	-0.08	0.12
1978	0.84	0.02	0.54	-0.06	0.33	1997	0.43	0.00	0.37	-0.05	0.12
1979	0.81	0.02	0.56	-0.07	0.30	1998	0.43	0.00	0.30	0.02	0.11
1980	0.77	0.02	0.55	-0.06	0.26	1999	0.43	0.00	0.40	-0.08	0.11
1981	0.75	0.02	0.52	-0.05	0.27	2000	0.42	0.00	0.36	-0.06	0.12
1982	0.76	0.02	0.52	-0.05	0.27	2001	0.40	0.00	0.35	-0.07	0.13
1983	0.74	0.02	0.51	-0.04	0.26	2002	0.40	-0.01	0.33	-0.05	0.12
1984	0.70	0.02	0.37	0.04	0.27	2003	0.38	-0.01	0.38	-0.09	0.11
1985	0.67	0.02	0.46	-0.04	0.23	2004	0.37	-0.02	0.38	-0.10	0.11
1986	0.66	0.02	0.49	-0.07	0.21	2005	0.38	-0.02	0.39	-0.11	0.12
1987	0.64	0.03	0.52	-0.10	0.19	2006	0.37	-0.02	0.43	-0.15	0.12
1988	0.58	0.03	0.40	-0.01	0.17	2007	0.36	-0.03	0.42	-0.12	0.10
1989	0.56	0.02	0.36	0.01	0.16	2008	0.34	-0.03	0.44	-0.15	0.08
1990	0.54	0.02	0.40	-0.02	0.14	2009	0.36	-0.04	0.34	-0.05	0.10
1991	0.51	0.01	0.35	-0.01	0.16	2010	0.35	-0.04	0.28	0.01	0.10
1992	0.49	0.01	0.33	-0.01	0.15	2011	0.35	-0.04	0.41	-0.12	0.10
1993	0.46	0.01	0.26	0.05	0.14	2012	0.36	-0.04	0.30	0.00	0.10
1994	0.46	0.01	0.27	0.03	0.14	2013	0.36	-0.04	0.40	-0.12	0.11

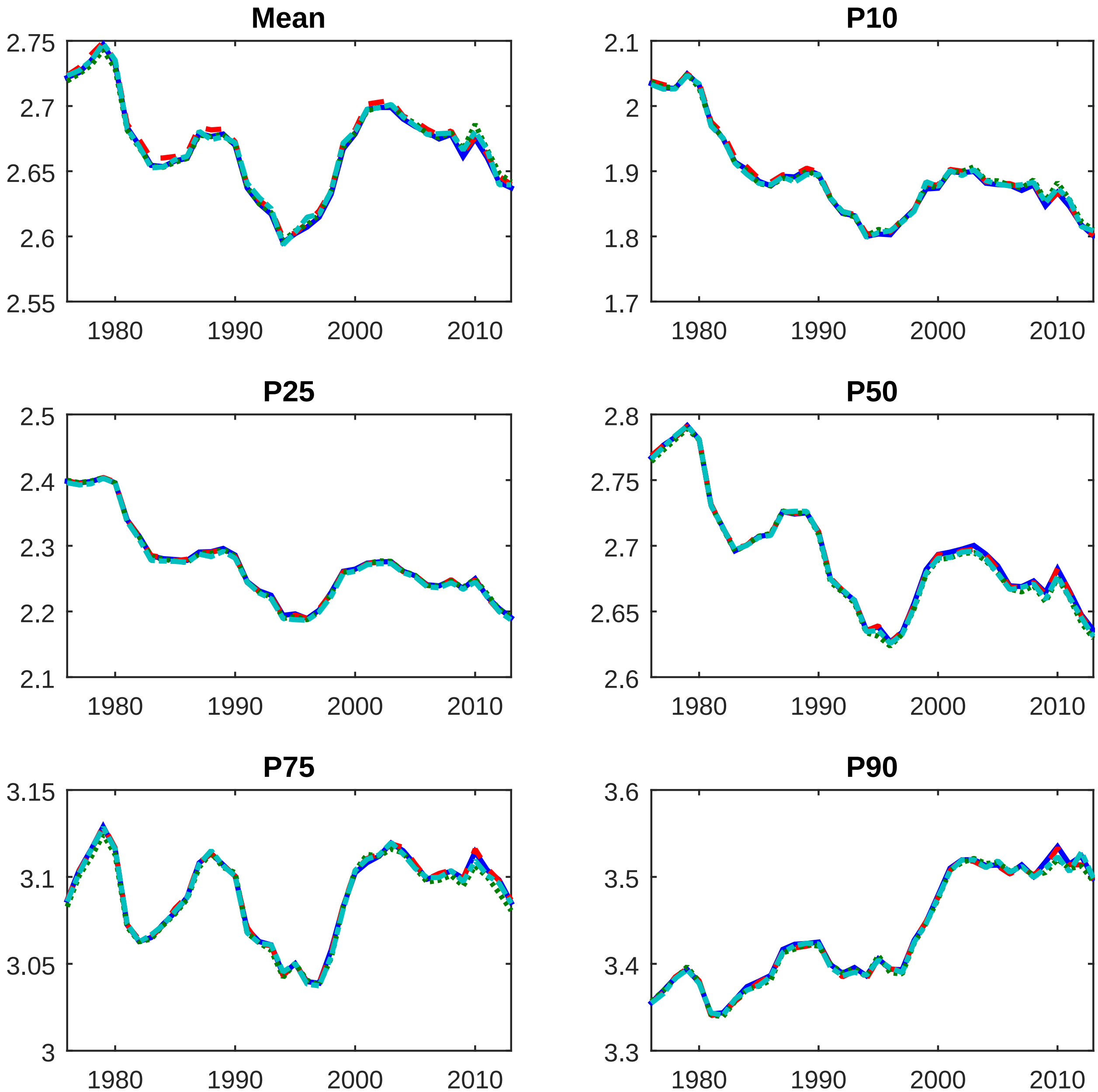
Notes: Total, EC, CC, SC and PC respectively denote total difference, endowments component, coefficients component, selection component and participation component.

Table 38: 90th percentile decomposition, actual earnings for the full population (Frank copula)

Year	Total	EC	CC	SC	PC	Year	Total	EC	CC	SC	PC
1976	0.68	0.03	0.45	-0.03	0.23	1995	0.39	0.03	0.29	0.01	0.06
1977	0.67	0.03	0.44	-0.01	0.21	1996	0.38	0.02	0.37	-0.06	0.06
1978	0.67	0.03	0.53	-0.06	0.17	1997	0.37	0.01	0.34	-0.05	0.06
1979	0.67	0.03	0.55	-0.07	0.16	1998	0.38	0.01	0.30	0.02	0.05
1980	0.64	0.03	0.53	-0.06	0.15	1999	0.38	0.02	0.38	-0.07	0.05
1981	0.63	0.03	0.51	-0.05	0.15	2000	0.38	0.01	0.35	-0.05	0.06
1982	0.63	0.03	0.52	-0.05	0.14	2001	0.38	0.01	0.37	-0.07	0.07
1983	0.62	0.03	0.51	-0.04	0.11	2002	0.38	0.01	0.35	-0.05	0.07
1984	0.58	0.04	0.38	0.04	0.13	2003	0.37	0.00	0.39	-0.08	0.05
1985	0.56	0.04	0.45	-0.04	0.11	2004	0.35	0.00	0.39	-0.09	0.06
1986	0.54	0.04	0.46	-0.06	0.10	2005	0.35	-0.01	0.39	-0.09	0.05
1987	0.52	0.04	0.48	-0.08	0.08	2006	0.35	-0.01	0.45	-0.15	0.06
1988	0.50	0.04	0.39	-0.01	0.08	2007	0.34	-0.02	0.41	-0.10	0.05
1989	0.47	0.03	0.35	0.01	0.08	2008	0.32	-0.02	0.43	-0.13	0.04
1990	0.46	0.04	0.38	-0.02	0.06	2009	0.35	-0.02	0.36	-0.04	0.05
1991	0.44	0.03	0.34	-0.01	0.08	2010	0.32	-0.03	0.30	0.01	0.05
1992	0.41	0.03	0.32	-0.01	0.07	2011	0.32	-0.02	0.40	-0.09	0.04
1993	0.40	0.03	0.27	0.04	0.06	2012	0.34	-0.02	0.32	0.00	0.05
1994	0.39	0.03	0.28	0.02	0.06	2013	0.32	-0.03	0.40	-0.10	0.05

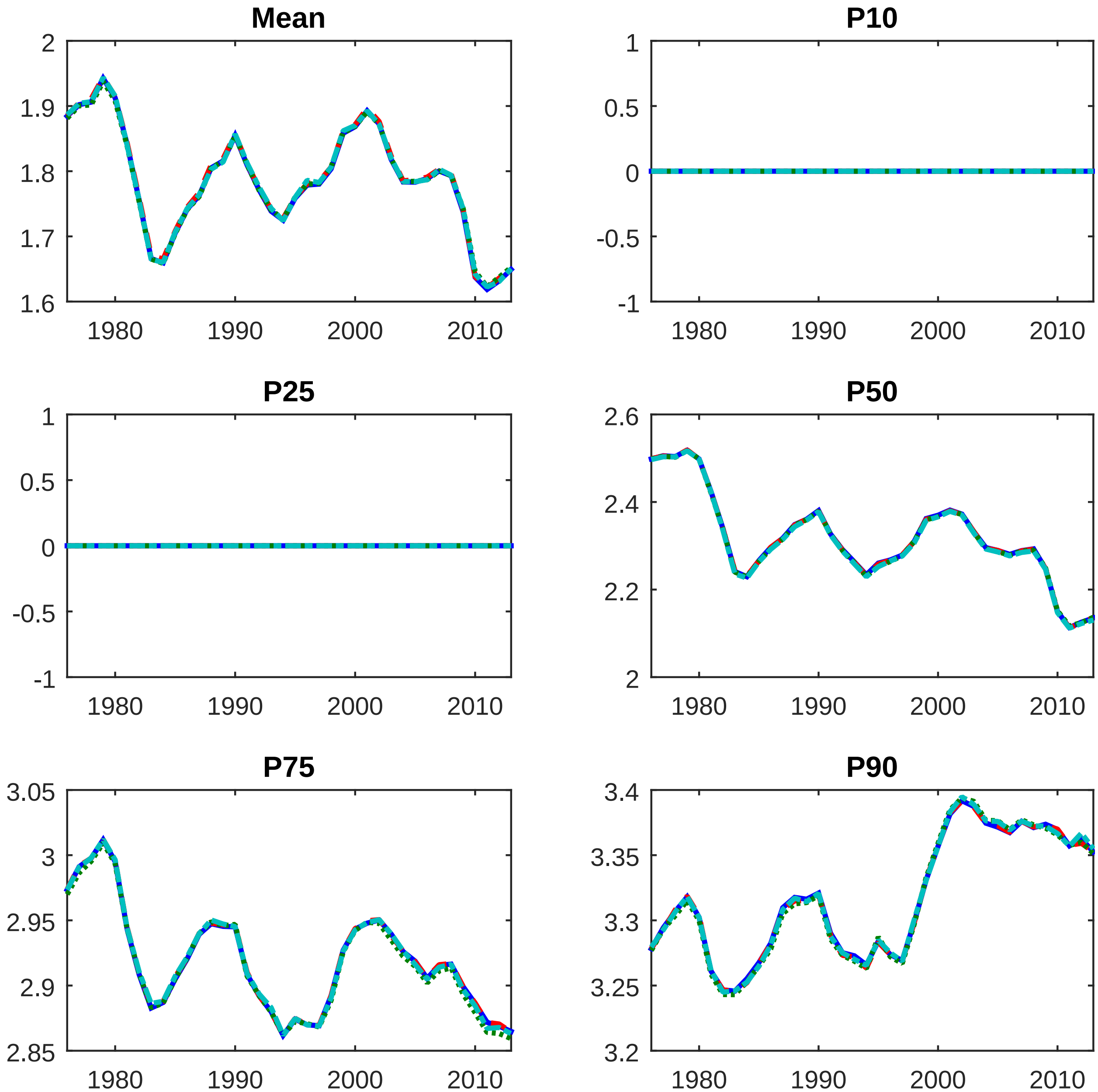
Notes: Total, EC, CC, SC and PC respectively denote total difference, endowments component, coefficients component, selection component and participation component.

Figure 8: Actual earnings distributions for male participants



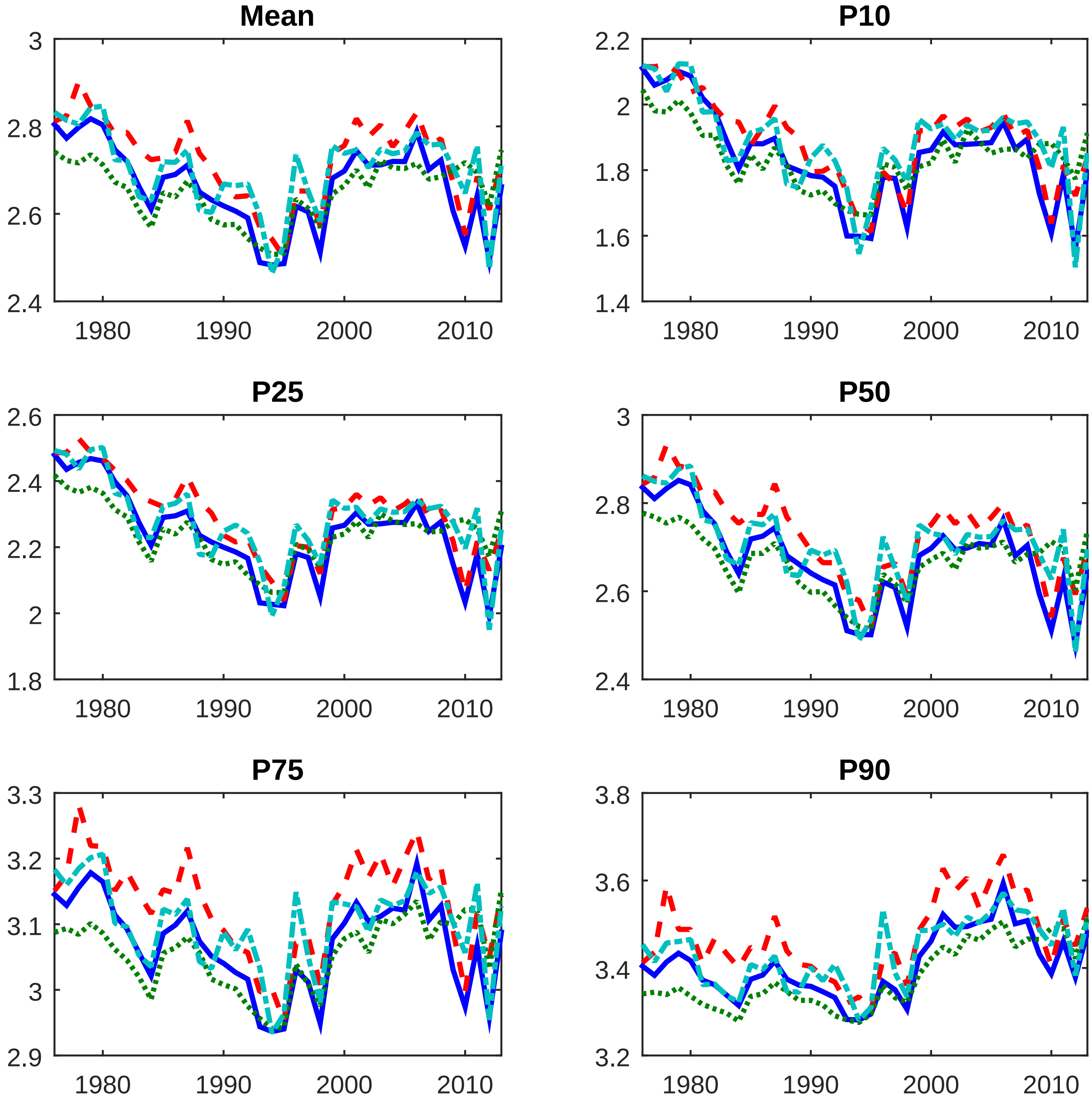
Notes: the solid thick blue line denotes the estimate with the Frank copula for all individuals; the solid thin red line denotes the estimate with the Gaussian copula for all individuals; the dashed green line denotes the estimate with the Frank copula, heterogeneous across race; the dotted orange line denotes the estimate with the Frank copula, heterogeneous across education level; the dashed-dotted cyan line denotes the estimate with the Frank copula, heterogeneous across marital status.

Figure 9: Actual earnings distributions for the full male population



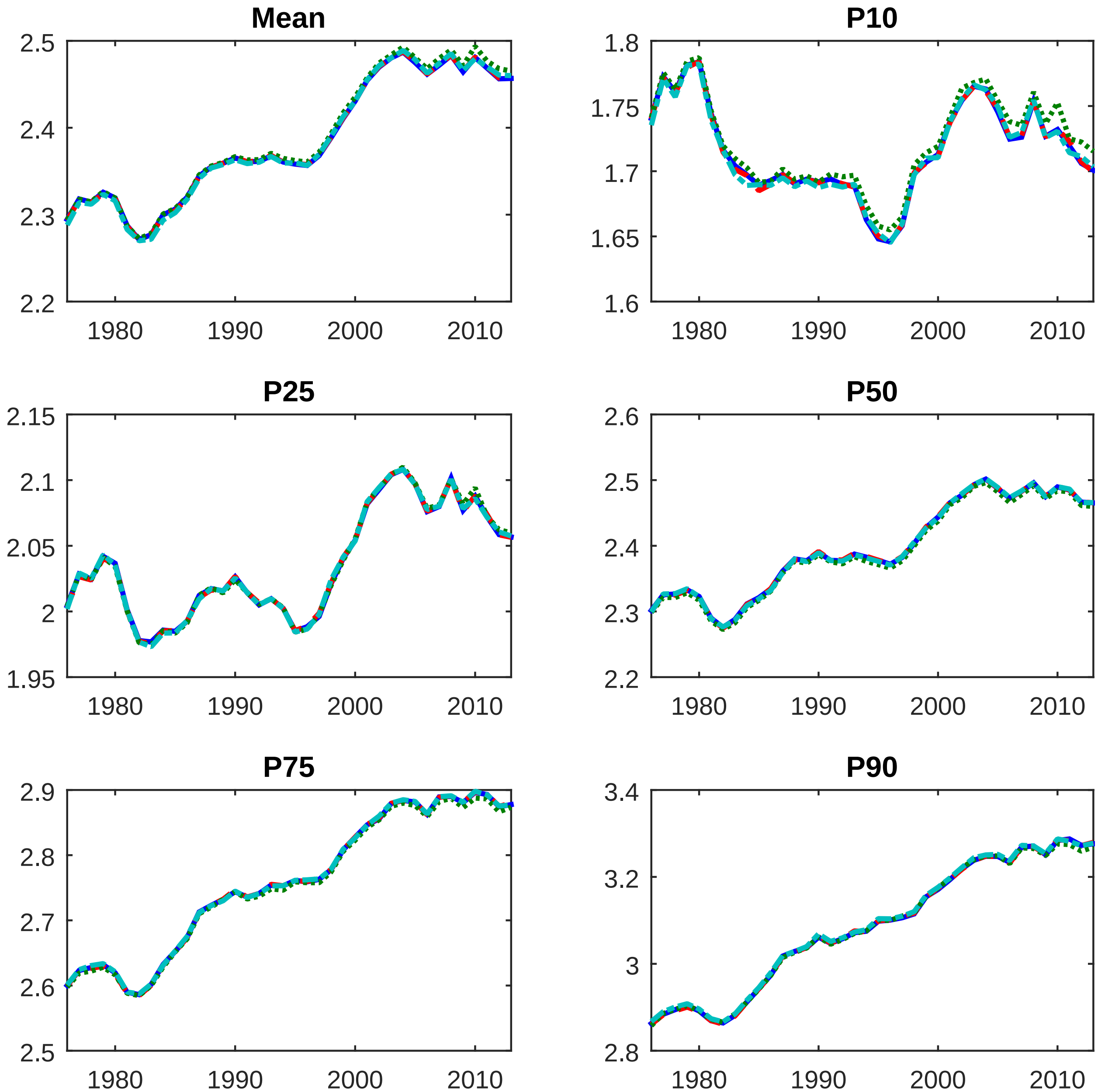
Notes: the solid thick blue line denotes the estimate with the Frank copula for all individuals; the solid thin red line denotes the estimate with the Gaussian copula for all individuals; the dashed green line denotes the estimate with the Frank copula, heterogeneous across race; the dotted orange line denotes the estimate with the Frank copula, heterogeneous across education level; the dashed-dotted cyan line denotes the estimate with the Frank copula, heterogeneous across marital status.

Figure 10: Potential earnings distributions for the full male population



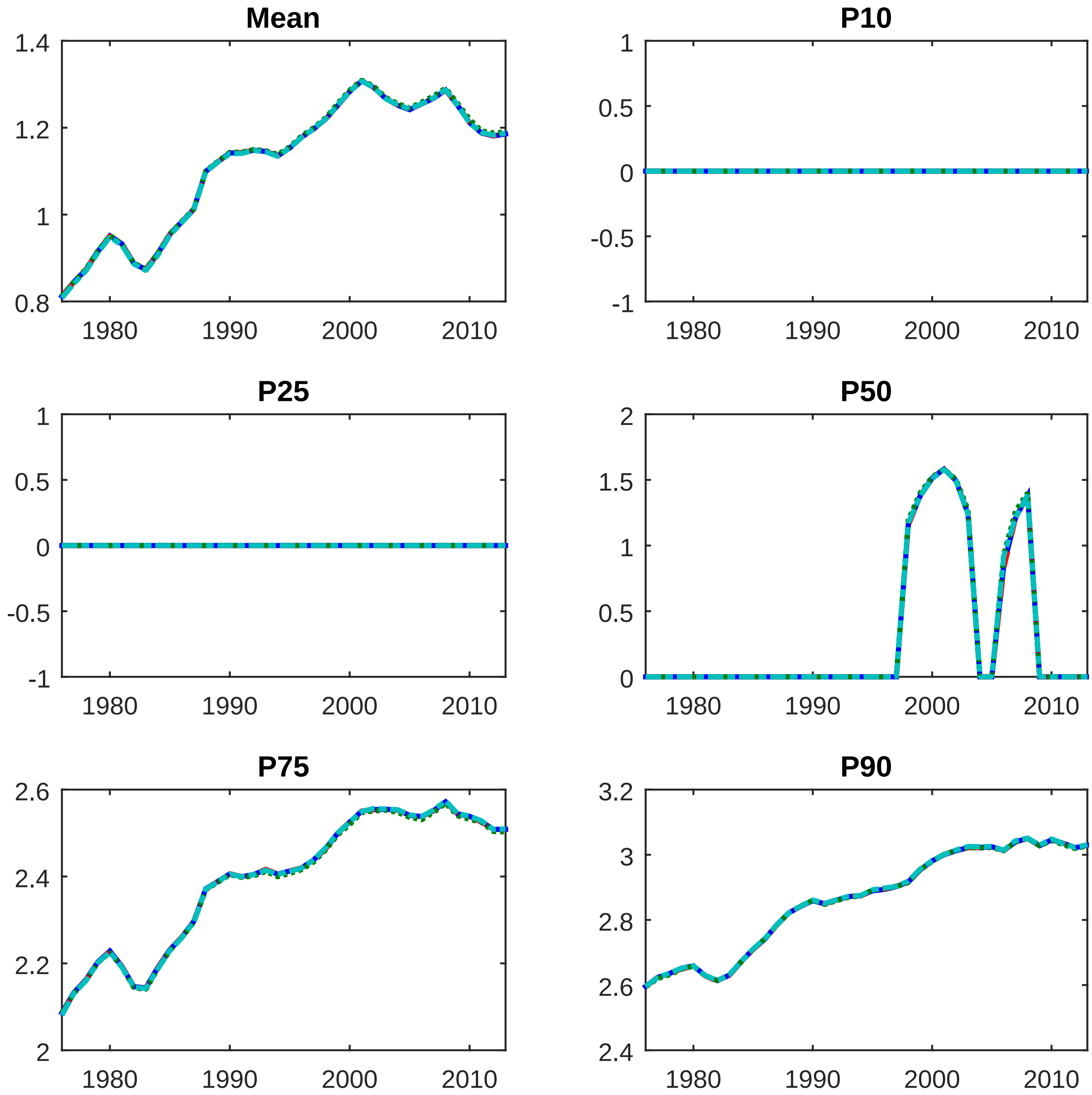
Notes: the solid thick blue line denotes the estimate with the Frank copula for all individuals; the solid thin red line denotes the estimate with the Gaussian copula for all individuals; the dashed green line denotes the estimate with the Frank copula, heterogeneous across race; the dotted orange line denotes the estimate with the Frank copula, heterogeneous across education level; the dashed-dotted cyan line denotes the estimate with the Frank copula, heterogeneous across marital status.

Figure 11: Actual earnings distributions for female participants



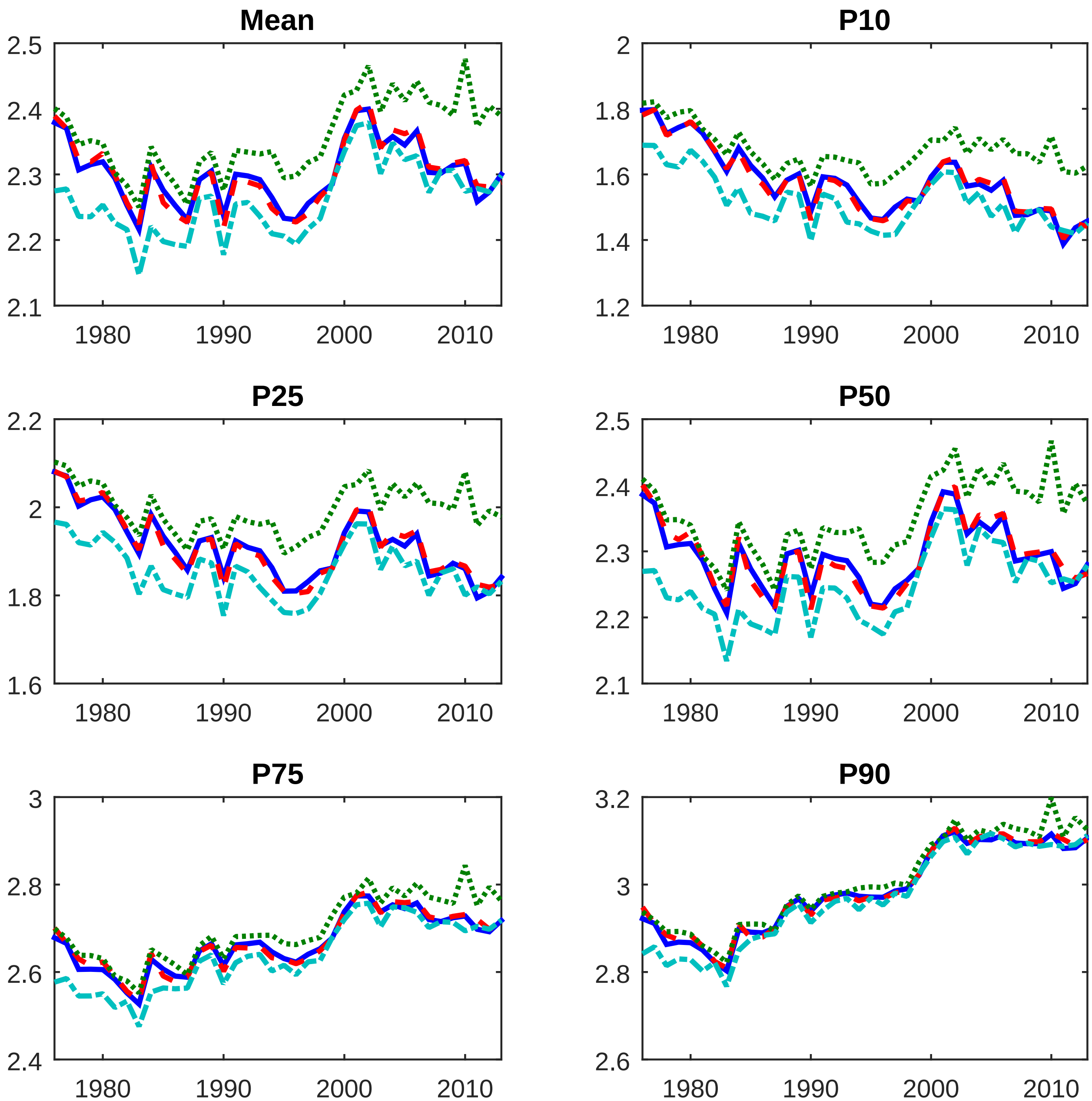
Notes: the solid thick blue line denotes the estimate with the Frank copula for all individuals; the solid thin red line denotes the estimate with the Gaussian copula for all individuals; the dashed green line denotes the estimate with the Frank copula, heterogeneous across race; the dotted orange line denotes the estimate with the Frank copula, heterogeneous across education level; the dashed-dotted cyan line denotes the estimate with the Frank copula, heterogeneous across marital status.

Figure 12: Actual earnings distributions for the full female population



Notes: the solid thick blue line denotes the estimate with the Frank copula for all individuals; the solid thin red line denotes the estimate with the Gaussian copula for all individuals; the dashed green line denotes the estimate with the Frank copula, heterogeneous across race; the dotted orange line denotes the estimate with the Frank copula, heterogeneous across education level; the dashed-dotted cyan line denotes the estimate with the Frank copula, heterogeneous across marital status.

Figure 13: Potential earnings distributions for the full female population



Notes: the solid thick blue line denotes the estimate with the Frank copula for all individuals; the solid thin red line denotes the estimate with the Gaussian copula for all individuals; the dashed green line denotes the estimate with the Frank copula, heterogeneous across race; the dotted orange line denotes the estimate with the Frank copula, heterogeneous across education level; the dashed-dotted cyan line denotes the estimate with the Frank copula, heterogeneous across marital status.

Table 39: Self-selection decomposition (Gaussian copula)

Year	Total	EC	SC	PC	Year	Total	EC	SC	PC
1976	0.2	-0.1	-2.7	3.1	1995	0.0	0.1	2.6	-2.6
1977	0.9	-0.1	-0.9	1.9	1996	-7.9	0.0	-5.5	-2.4
1978	-4.9	-0.1	-4.7	-0.1	1997	-6.5	0.0	-4.3	-2.3
1979	-5.0	-0.1	-4.5	-0.4	1998	-1.1	0.0	1.2	-2.3
1980	-4.6	-0.1	-4.7	0.1	1999	-5.7	0.0	-3.3	-2.4
1981	-3.6	0.0	-3.8	0.2	2000	-6.1	0.0	-4.6	-1.5
1982	-4.7	0.0	-4.0	-0.7	2001	-6.6	0.0	-5.3	-1.3
1983	-2.4	0.0	-1.0	-1.5	2002	-5.1	0.0	-3.7	-1.5
1984	4.6	0.0	4.3	0.3	2003	-8.8	0.0	-6.4	-2.4
1985	-4.4	0.0	-3.6	-0.8	2004	-8.1	0.0	-5.9	-2.2
1986	-5.8	0.0	-4.2	-1.5	2005	-8.9	0.0	-6.7	-2.2
1987	-9.5	0.0	-7.0	-2.6	2006	-12.1	0.0	-10.4	-1.6
1988	-1.4	0.0	-0.1	-1.3	2007	-11.4	0.0	-8.4	-3.0
1989	0.3	0.0	1.6	-1.2	2008	-12.5	0.0	-9.4	-3.0
1990	-5.7	0.0	-2.8	-2.9	2009	-4.6	0.1	-2.3	-2.3
1991	-1.6	0.0	-0.3	-1.3	2010	-0.4	0.2	1.6	-2.2
1992	-1.5	0.0	-0.1	-1.4	2011	-8.6	0.0	-6.0	-2.6
1993	2.2	0.0	3.8	-1.7	2012	-1.5	0.2	0.9	-2.5
1994	-0.5	0.0	1.5	-2.0	2013	-9.2	0.0	-6.8	-2.4

Notes: Total, EC, SC and PC respectively denote total difference, endowments component, selection component and participation component.

Table 40: Actual earnings distributions for participants by gender (Gaussian copula)

	Male						Female					
Year	Mean	P10	P25	P50	P75	P90	Mean	P10	P25	P50	P75	P90
1976	2.72	2.03	2.40	2.77	3.09	3.35	2.30	1.74	2.00	2.30	2.60	2.86
1977	2.73	2.03	2.40	2.78	3.10	3.37	2.32	1.77	2.03	2.33	2.62	2.88
1978	2.74	2.03	2.40	2.78	3.12	3.38	2.31	1.76	2.02	2.33	2.63	2.89
1979	2.75	2.05	2.40	2.79	3.13	3.39	2.33	1.78	2.04	2.33	2.63	2.90
1980	2.73	2.03	2.40	2.78	3.12	3.38	2.32	1.78	2.04	2.32	2.62	2.89
1981	2.69	1.97	2.34	2.73	3.07	3.34	2.29	1.74	2.00	2.29	2.59	2.87
1982	2.67	1.95	2.31	2.71	3.06	3.34	2.27	1.72	1.98	2.28	2.59	2.86
1983	2.65	1.91	2.28	2.70	3.07	3.36	2.28	1.71	1.98	2.29	2.60	2.88
1984	2.65	1.90	2.28	2.70	3.07	3.38	2.30	1.70	1.99	2.31	2.63	2.91
1985	2.66	1.88	2.28	2.71	3.08	3.38	2.31	1.69	1.99	2.32	2.65	2.94
1986	2.66	1.88	2.28	2.71	3.09	3.39	2.32	1.69	1.99	2.33	2.67	2.97
1987	2.68	1.89	2.29	2.73	3.11	3.42	2.34	1.70	2.01	2.36	2.71	3.02
1988	2.68	1.89	2.29	2.72	3.11	3.42	2.35	1.69	2.02	2.38	2.72	3.03
1989	2.68	1.90	2.30	2.73	3.11	3.43	2.36	1.70	2.02	2.38	2.73	3.04
1990	2.67	1.89	2.29	2.71	3.10	3.42	2.36	1.69	2.03	2.39	2.74	3.06
1991	2.64	1.86	2.25	2.67	3.07	3.40	2.36	1.69	2.01	2.38	2.74	3.05
1992	2.62	1.83	2.23	2.67	3.06	3.39	2.36	1.69	2.01	2.38	2.74	3.06
1993	2.62	1.83	2.23	2.66	3.06	3.40	2.37	1.69	2.01	2.39	2.75	3.07
1994	2.59	1.80	2.19	2.64	3.04	3.39	2.36	1.66	2.00	2.38	2.75	3.08
1995	2.60	1.81	2.20	2.64	3.05	3.41	2.36	1.65	1.99	2.38	2.76	3.10
1996	2.61	1.80	2.19	2.63	3.04	3.39	2.35	1.65	1.99	2.37	2.76	3.10
1997	2.61	1.82	2.20	2.63	3.04	3.39	2.37	1.66	2.00	2.38	2.76	3.11
1998	2.63	1.84	2.23	2.66	3.06	3.43	2.39	1.70	2.02	2.40	2.78	3.12
1999	2.66	1.87	2.26	2.68	3.08	3.45	2.41	1.71	2.04	2.43	2.81	3.16
2000	2.68	1.87	2.26	2.69	3.10	3.48	2.43	1.71	2.05	2.44	2.83	3.17
2001	2.70	1.90	2.27	2.70	3.11	3.51	2.45	1.74	2.08	2.47	2.85	3.20
2002	2.70	1.90	2.28	2.70	3.11	3.52	2.47	1.76	2.09	2.48	2.86	3.22
2003	2.70	1.90	2.28	2.70	3.12	3.52	2.48	1.77	2.10	2.49	2.88	3.24
2004	2.69	1.88	2.26	2.69	3.12	3.51	2.48	1.76	2.11	2.50	2.88	3.25
2005	2.68	1.88	2.25	2.68	3.11	3.51	2.47	1.75	2.10	2.49	2.88	3.25
2006	2.68	1.88	2.24	2.67	3.10	3.50	2.46	1.73	2.08	2.47	2.86	3.23
2007	2.68	1.87	2.24	2.67	3.10	3.51	2.47	1.73	2.08	2.48	2.89	3.27
2008	2.68	1.88	2.25	2.67	3.10	3.50	2.48	1.76	2.10	2.50	2.89	3.28
2009	2.66	1.85	2.23	2.66	3.10	3.52	2.46	1.73	2.08	2.48	2.88	3.26
2010	2.67	1.87	2.25	2.68	3.12	3.54	2.48	1.73	2.09	2.49	2.90	3.29
2011	2.66	1.85	2.22	2.67	3.11	3.52	2.47	1.72	2.07	2.49	2.89	3.30
2012	2.64	1.82	2.21	2.65	3.10	3.53	2.45	1.71	2.06	2.47	2.88	3.28
2013	2.64	1.80	2.19	2.64	3.09	3.50	2.45	1.70	2.06	2.47	2.88	3.28

Table 41: Actual earnings distributions for the full population by gender (Gaussian copula)

	Male						Female					
Year	Mean	P10	P25	P50	P75	P90	Mean	P10	P25	P50	P75	P90
1976	1.89	0.00	0.00	2.50	2.97	3.28	0.81	0.00	0.00	0.00	2.09	2.60
1977	1.90	0.00	0.00	2.51	2.99	3.29	0.85	0.00	0.00	0.00	2.13	2.62
1978	1.91	0.00	0.00	2.50	3.00	3.31	0.87	0.00	0.00	0.00	2.16	2.64
1979	1.94	0.00	0.00	2.52	3.01	3.32	0.92	0.00	0.00	0.00	2.20	2.65
1980	1.92	0.00	0.00	2.50	3.00	3.30	0.95	0.00	0.00	0.00	2.23	2.66
1981	1.84	0.00	0.00	2.42	2.94	3.26	0.93	0.00	0.00	0.00	2.19	2.63
1982	1.76	0.00	0.00	2.34	2.91	3.25	0.89	0.00	0.00	0.00	2.15	2.61
1983	1.66	0.00	0.00	2.24	2.88	3.25	0.87	0.00	0.00	0.00	2.14	2.63
1984	1.66	0.00	0.00	2.23	2.89	3.26	0.91	0.00	0.00	0.00	2.19	2.67
1985	1.71	0.00	0.00	2.27	2.90	3.27	0.95	0.00	0.00	0.00	2.23	2.71
1986	1.74	0.00	0.00	2.30	2.92	3.28	0.98	0.00	0.00	0.00	2.26	2.74
1987	1.76	0.00	0.00	2.32	2.94	3.31	1.01	0.00	0.00	0.00	2.30	2.79
1988	1.80	0.00	0.00	2.35	2.95	3.32	1.10	0.00	0.00	0.00	2.37	2.82
1989	1.81	0.00	0.00	2.36	2.95	3.32	1.12	0.00	0.00	0.00	2.39	2.84
1990	1.85	0.00	0.00	2.38	2.94	3.32	1.14	0.00	0.00	0.00	2.41	2.86
1991	1.81	0.00	0.00	2.33	2.91	3.29	1.14	0.00	0.00	0.00	2.40	2.85
1992	1.77	0.00	0.00	2.29	2.89	3.28	1.15	0.00	0.00	0.00	2.40	2.86
1993	1.74	0.00	0.00	2.26	2.88	3.27	1.14	0.00	0.00	0.00	2.42	2.87
1994	1.72	0.00	0.00	2.23	2.86	3.27	1.13	0.00	0.00	0.00	2.41	2.87
1995	1.76	0.00	0.00	2.26	2.87	3.29	1.15	0.00	0.00	0.00	2.41	2.89
1996	1.78	0.00	0.00	2.27	2.87	3.27	1.18	0.00	0.00	0.00	2.42	2.89
1997	1.78	0.00	0.00	2.28	2.87	3.27	1.19	0.00	0.00	0.00	2.44	2.90
1998	1.80	0.00	0.00	2.31	2.89	3.30	1.22	0.00	0.00	1.15	2.46	2.92
1999	1.86	0.00	0.00	2.36	2.93	3.33	1.25	0.00	0.00	1.38	2.50	2.95
2000	1.87	0.00	0.00	2.37	2.94	3.36	1.28	0.00	0.00	1.51	2.53	2.98
2001	1.89	0.00	0.00	2.38	2.95	3.38	1.31	0.00	0.00	1.58	2.55	3.00
2002	1.87	0.00	0.00	2.37	2.95	3.39	1.29	0.00	0.00	1.49	2.55	3.01
2003	1.82	0.00	0.00	2.33	2.94	3.39	1.26	0.00	0.00	1.23	2.55	3.02
2004	1.78	0.00	0.00	2.30	2.93	3.38	1.25	0.00	0.00	0.00	2.55	3.02
2005	1.78	0.00	0.00	2.29	2.92	3.37	1.24	0.00	0.00	0.00	2.54	3.03
2006	1.79	0.00	0.00	2.28	2.91	3.37	1.25	0.00	0.00	0.86	2.54	3.01
2007	1.80	0.00	0.00	2.29	2.92	3.38	1.26	0.00	0.00	1.21	2.55	3.04
2008	1.79	0.00	0.00	2.29	2.92	3.37	1.28	0.00	0.00	1.38	2.57	3.05
2009	1.74	0.00	0.00	2.25	2.90	3.37	1.25	0.00	0.00	0.00	2.54	3.03
2010	1.63	0.00	0.00	2.15	2.89	3.37	1.21	0.00	0.00	0.00	2.54	3.05
2011	1.62	0.00	0.00	2.11	2.87	3.36	1.18	0.00	0.00	0.00	2.53	3.04
2012	1.63	0.00	0.00	2.13	2.87	3.37	1.18	0.00	0.00	0.00	2.51	3.02
2013	1.65	0.00	0.00	2.13	2.86	3.35	1.18	0.00	0.00	0.00	2.51	3.03

Table 42: Potential earnings distributions for the full population by gender (Gaussian copula)

	Male						Female					
Year	Mean	P10	P25	P50	P75	P90	Mean	P10	P25	P50	P75	P90
1976	2.81	2.11	2.48	2.84	3.15	3.42	2.37	1.80	2.07	2.37	2.67	2.92
1977	2.77	2.05	2.43	2.80	3.12	3.38	2.37	1.80	2.07	2.37	2.66	2.91
1978	2.81	2.09	2.47	2.84	3.17	3.43	2.31	1.73	2.01	2.31	2.61	2.87
1979	2.81	2.10	2.46	2.85	3.18	3.44	2.31	1.74	2.02	2.31	2.61	2.87
1980	2.81	2.09	2.47	2.85	3.17	3.43	2.32	1.76	2.03	2.32	2.61	2.87
1981	2.75	2.02	2.40	2.78	3.12	3.38	2.29	1.73	2.00	2.29	2.58	2.85
1982	2.72	1.98	2.36	2.75	3.09	3.37	2.25	1.66	1.94	2.24	2.55	2.82
1983	2.63	1.85	2.24	2.66	3.03	3.32	2.22	1.61	1.90	2.22	2.53	2.81
1984	2.59	1.77	2.18	2.62	3.01	3.30	2.31	1.68	1.98	2.31	2.63	2.90
1985	2.70	1.90	2.30	2.73	3.10	3.39	2.28	1.62	1.93	2.27	2.61	2.89
1986	2.69	1.88	2.29	2.72	3.10	3.39	2.26	1.59	1.91	2.26	2.60	2.89
1987	2.72	1.91	2.32	2.76	3.13	3.43	2.23	1.52	1.86	2.23	2.59	2.90
1988	2.64	1.80	2.23	2.67	3.07	3.37	2.30	1.59	1.93	2.31	2.66	2.95
1989	2.61	1.76	2.19	2.64	3.04	3.35	2.30	1.60	1.93	2.31	2.67	2.96
1990	2.64	1.81	2.22	2.66	3.06	3.37	2.24	1.47	1.84	2.24	2.62	2.93
1991	2.60	1.77	2.18	2.62	3.02	3.34	2.30	1.59	1.93	2.30	2.67	2.97
1992	2.58	1.74	2.16	2.61	3.01	3.33	2.30	1.58	1.91	2.29	2.67	2.97
1993	2.48	1.58	2.02	2.51	2.94	3.28	2.29	1.55	1.90	2.29	2.67	2.98
1994	2.49	1.61	2.04	2.52	2.95	3.28	2.26	1.50	1.86	2.26	2.64	2.97
1995	2.46	1.53	1.98	2.48	2.93	3.28	2.23	1.44	1.81	2.22	2.63	2.96
1996	2.62	1.79	2.19	2.63	3.04	3.37	2.23	1.45	1.82	2.23	2.63	2.97
1997	2.61	1.78	2.17	2.61	3.02	3.35	2.25	1.48	1.83	2.25	2.64	2.98
1998	2.52	1.63	2.06	2.53	2.96	3.31	2.27	1.51	1.86	2.26	2.66	2.99
1999	2.64	1.80	2.21	2.64	3.04	3.39	2.29	1.51	1.87	2.29	2.68	3.02
2000	2.71	1.87	2.27	2.70	3.11	3.47	2.35	1.58	1.94	2.35	2.74	3.07
2001	2.75	1.92	2.31	2.73	3.14	3.53	2.39	1.62	1.98	2.38	2.77	3.10
2002	2.71	1.88	2.27	2.69	3.10	3.49	2.39	1.62	1.98	2.38	2.77	3.11
2003	2.72	1.88	2.27	2.70	3.11	3.50	2.34	1.54	1.91	2.33	2.74	3.09
2004	2.71	1.86	2.26	2.69	3.11	3.49	2.36	1.55	1.93	2.35	2.76	3.10
2005	2.71	1.88	2.27	2.70	3.12	3.51	2.34	1.54	1.91	2.34	2.75	3.10
2006	2.81	1.96	2.34	2.78	3.22	3.64	2.37	1.57	1.94	2.36	2.76	3.11
2007	2.71	1.87	2.25	2.69	3.11	3.51	2.29	1.43	1.83	2.28	2.72	3.08
2008	2.72	1.89	2.27	2.70	3.12	3.51	2.29	1.43	1.84	2.29	2.71	3.09
2009	2.59	1.70	2.13	2.58	3.02	3.42	2.31	1.47	1.88	2.30	2.73	3.09
2010	2.49	1.53	1.99	2.49	2.96	3.37	2.31	1.46	1.86	2.30	2.73	3.11
2011	2.61	1.74	2.15	2.60	3.04	3.44	2.26	1.35	1.79	2.26	2.70	3.08
2012	2.47	1.51	1.96	2.46	2.95	3.37	2.27	1.41	1.81	2.26	2.70	3.08
2013	2.64	1.77	2.17	2.62	3.07	3.46	2.29	1.42	1.83	2.28	2.71	3.10

Table 43: Self-selection decomposition, heterogeneous copula by race

Year	Total	EC	SC	PC	Year	Total	EC	SC	PC
1976	1.3	-0.2	-2.6	4.1	1995	-1.9	0.0	1.1	-3.0
1977	-2.3	-0.2	-4.2	2.1	1996	-10.4	-0.2	-7.1	-3.0
1978	-8.9	-0.5	-9.0	0.6	1997	-10.0	-0.4	-6.6	-3.0
1979	-6.8	-0.4	-6.3	-0.1	1998	-4.5	-0.3	-1.5	-2.7
1980	-5.0	-0.5	-5.2	0.7	1999	-11.7	0.0	-8.7	-3.0
1981	-5.6	-0.1	-6.0	0.5	2000	-9.2	-0.1	-7.3	-1.8
1982	-8.1	-0.3	-7.4	-0.4	2001	-9.7	-0.1	-8.2	-1.4
1983	-9.3	-0.3	-7.7	-1.4	2002	-7.8	-0.2	-6.2	-1.5
1984	-2.9	0.0	-3.6	0.7	2003	-13.9	-0.2	-11.0	-2.7
1985	-7.7	-0.3	-6.3	-1.1	2004	-10.7	0.0	-8.3	-2.4
1986	-10.4	-0.1	-8.2	-2.1	2005	-12.4	-0.2	-9.8	-2.3
1987	-15.0	0.2	-12.5	-2.8	2006	-13.6	-0.1	-11.6	-1.8
1988	-7.7	0.1	-6.2	-1.6	2007	-13.6	-0.2	-10.3	-3.1
1989	-5.1	0.1	-3.7	-1.5	2008	-15.2	-0.3	-11.7	-3.3
1990	-8.4	-0.2	-4.5	-3.7	2009	-9.1	-0.1	-6.3	-2.7
1991	-4.4	-0.2	-2.5	-1.7	2010	-2.8	0.0	-0.4	-2.4
1992	-5.7	0.0	-3.8	-1.9	2011	-11.7	-0.2	-8.6	-2.9
1993	-2.6	0.2	-0.7	-2.0	2012	-8.1	-0.1	-5.2	-2.8
1994	-4.4	-0.3	-1.6	-2.5	2013	-13.8	-0.2	-10.8	-2.8

Notes: Total, EC, SC and PC respectively denote total difference, endowments component, selection component and participation component; coefficients scaled by 100.

Table 44: Self-selection decomposition, heterogeneous copula by education

Year	Total	EC	SC	PC	Year	Total	EC	SC	PC
1976	6.2	0.5	1.5	4.3	1995	0.6	0.4	1.7	-1.5
1977	4.4	0.6	0.9	2.8	1996	-5.4	0.1	-4.3	-1.3
1978	3.0	0.5	1.3	1.2	1997	-3.6	0.3	-2.8	-1.1
1979	2.3	0.3	0.9	1.1	1998	-1.0	0.3	0.0	-1.3
1980	2.8	0.5	1.2	1.1	1999	-2.4	0.3	-2.0	-0.7
1981	0.4	0.7	-0.9	0.6	2000	-1.2	0.3	-1.3	-0.2
1982	0.4	0.7	-0.7	0.4	2001	-2.6	0.2	-2.2	-0.6
1983	-0.6	0.8	-0.7	-0.8	2002	0.3	0.2	0.3	-0.2
1984	7.6	0.6	5.8	1.2	2003	-6.9	0.1	-5.2	-1.7
1985	-0.4	0.8	-1.2	0.1	2004	-4.6	0.0	-3.6	-1.0
1986	-2.9	0.8	-2.8	-0.9	2005	-5.7	0.0	-4.5	-1.2
1987	-6.0	0.5	-4.4	-2.1	2006	-4.0	-0.1	-3.8	-0.1
1988	-0.3	0.4	0.1	-0.8	2007	-4.6	-0.1	-3.3	-1.2
1989	2.4	0.5	2.6	-0.6	2008	-7.0	-0.3	-5.2	-1.6
1990	-1.4	0.4	0.4	-2.2	2009	-6.4	-0.3	-4.9	-1.2
1991	0.6	0.5	0.6	-0.6	2010	-2.3	-0.4	-1.8	-0.2
1992	1.9	0.5	1.9	-0.6	2011	-8.2	-0.4	-6.3	-1.4
1993	2.2	0.4	2.5	-0.7	2012	-0.5	-0.3	0.3	-0.5
1994	3.0	0.4	3.0	-0.4	2013	-8.6	-0.5	-7.1	-1.1

Notes: Total, EC, SC and PC respectively denote total difference, endowments component, selection component and participation component; coefficients scaled by 100.

Table 45: Self-selection decomposition, heterogeneous copula by marital status

Year	Total	EC	SC	PC	Year	Total	EC	SC	PC
1976	-8.4	-0.2	-6.3	-2.0	1995	-4.2	0.0	0.4	-4.7
1977	-8.4	-0.1	-5.0	-3.2	1996	-15.9	0.0	-11.6	-4.2
1978	-10.1	-0.1	-6.2	-3.8	1997	-10.1	0.0	-5.6	-4.5
1979	-12.2	-0.1	-7.5	-4.6	1998	-6.2	0.0	-1.9	-4.3
1980	-11.7	-0.1	-7.9	-3.7	1999	-11.9	0.0	-8.2	-3.7
1981	-6.9	0.0	-3.7	-3.2	2000	-8.9	0.0	-5.8	-3.1
1982	-7.0	0.0	-3.6	-3.4	2001	-7.4	0.0	-4.4	-3.0
1983	-7.3	0.0	-2.8	-4.5	2002	-5.9	0.0	-2.8	-3.1
1984	-3.2	0.0	0.4	-3.6	2003	-13.1	0.0	-9.3	-3.8
1985	-10.0	0.0	-5.4	-4.5	2004	-10.0	-0.2	-5.9	-3.9
1986	-10.8	0.0	-5.9	-4.9	2005	-10.9	0.0	-6.4	-4.5
1987	-13.0	0.0	-8.4	-4.7	2006	-12.4	-0.1	-8.4	-4.0
1988	-1.6	0.0	1.8	-3.3	2007	-14.8	-0.1	-10.3	-4.5
1989	-0.9	0.0	2.5	-3.4	2008	-14.1	-0.2	-10.0	-4.0
1990	-10.6	0.0	-5.5	-5.1	2009	-11.6	-0.1	-8.3	-3.2
1991	-8.0	0.0	-4.7	-3.3	2010	-10.6	-0.3	-7.1	-3.2
1992	-8.4	0.0	-4.8	-3.6	2011	-15.3	-0.3	-11.7	-3.3
1993	-5.4	0.0	-1.1	-4.3	2012	-0.3	0.2	3.5	-4.0
1994	-1.6	0.0	2.5	-4.1	2013	-11.8	-0.3	-7.7	-3.7

Notes: Total, EC, SC and PC respectively denote total difference, endowments component, selection component and participation component; coefficients scaled by 100.

Table 46: Mean decomposition, actual earnings for participants, heterogeneous copula by race

Year	Total	EC	CC	SC	PC	Year	Total	EC	CC	SC	PC
1976	0.43	0.01	0.40	-0.04	0.05	1995	0.24	0.00	0.28	0.02	-0.05
1977	0.41	0.01	0.43	-0.05	0.03	1996	0.25	0.00	0.42	-0.12	-0.05
1978	0.43	0.01	0.52	-0.11	0.01	1997	0.25	-0.01	0.41	-0.11	-0.05
1979	0.42	0.01	0.50	-0.08	0.00	1998	0.25	-0.01	0.32	-0.02	-0.04
1980	0.42	0.01	0.47	-0.07	0.01	1999	0.26	0.00	0.45	-0.14	-0.05
1981	0.40	0.00	0.47	-0.08	0.01	2000	0.25	-0.01	0.40	-0.12	-0.03
1982	0.40	0.01	0.50	-0.10	-0.01	2001	0.25	-0.01	0.41	-0.13	-0.02
1983	0.38	0.01	0.50	-0.10	-0.02	2002	0.23	-0.01	0.37	-0.10	-0.02
1984	0.36	0.01	0.39	-0.05	0.01	2003	0.22	-0.01	0.46	-0.18	-0.05
1985	0.36	0.01	0.45	-0.09	-0.02	2004	0.20	-0.02	0.40	-0.14	-0.04
1986	0.35	0.01	0.48	-0.12	-0.03	2005	0.21	-0.02	0.44	-0.16	-0.04
1987	0.34	0.01	0.56	-0.19	-0.04	2006	0.22	-0.03	0.46	-0.19	-0.03
1988	0.33	0.01	0.43	-0.09	-0.02	2007	0.20	-0.03	0.47	-0.18	-0.05
1989	0.32	0.01	0.39	-0.06	-0.02	2008	0.20	-0.03	0.49	-0.20	-0.06
1990	0.31	0.01	0.43	-0.07	-0.06	2009	0.20	-0.04	0.39	-0.11	-0.05
1991	0.28	0.00	0.34	-0.04	-0.03	2010	0.19	-0.04	0.28	-0.01	-0.04
1992	0.27	0.01	0.35	-0.06	-0.03	2011	0.19	-0.04	0.44	-0.16	-0.05
1993	0.25	0.00	0.29	-0.01	-0.03	2012	0.19	-0.03	0.36	-0.09	-0.05
1994	0.24	0.00	0.30	-0.03	-0.04	2013	0.18	-0.04	0.46	-0.19	-0.05

Notes: Total, EC, CC, SC and PC respectively denote total difference, endowments component, coefficients component, selection component and participation component.

Table 47: Mean decomposition, actual earnings for participants, heterogeneous copula by education

Year	Total	EC	CC	SC	PC	Year	Total	EC	CC	SC	PC
1976	0.43	0.01	0.34	0.02	0.06	1995	0.24	0.00	0.23	0.03	-0.02
1977	0.41	0.01	0.35	0.01	0.04	1996	0.25	0.00	0.34	-0.07	-0.02
1978	0.42	0.01	0.37	0.02	0.02	1997	0.24	-0.01	0.31	-0.04	-0.02
1979	0.42	0.01	0.38	0.01	0.01	1998	0.24	-0.01	0.27	0.00	-0.02
1980	0.41	0.01	0.37	0.02	0.01	1999	0.25	0.00	0.30	-0.03	-0.01
1981	0.39	0.01	0.39	-0.01	0.01	2000	0.24	0.00	0.27	-0.02	-0.01
1982	0.40	0.01	0.39	-0.01	0.00	2001	0.24	-0.01	0.29	-0.03	-0.01
1983	0.38	0.01	0.38	-0.01	-0.01	2002	0.22	-0.01	0.23	0.01	0.00
1984	0.35	0.01	0.25	0.08	0.02	2003	0.22	-0.01	0.34	-0.08	-0.03
1985	0.35	0.01	0.35	-0.02	0.00	2004	0.20	-0.02	0.29	-0.05	-0.02
1986	0.34	0.01	0.38	-0.04	-0.01	2005	0.21	-0.02	0.32	-0.07	-0.02
1987	0.33	0.01	0.42	-0.06	-0.03	2006	0.21	-0.02	0.30	-0.06	-0.01
1988	0.32	0.01	0.32	0.00	-0.01	2007	0.20	-0.03	0.30	-0.05	-0.03
1989	0.32	0.01	0.28	0.04	-0.01	2008	0.19	-0.03	0.32	-0.08	-0.03
1990	0.30	0.01	0.32	0.01	-0.03	2009	0.20	-0.03	0.33	-0.08	-0.03
1991	0.27	0.01	0.27	0.01	-0.01	2010	0.19	-0.03	0.26	-0.03	-0.01
1992	0.26	0.01	0.23	0.03	-0.01	2011	0.19	-0.03	0.35	-0.09	-0.03
1993	0.25	0.01	0.21	0.04	-0.01	2012	0.18	-0.03	0.24	0.00	-0.02
1994	0.23	0.00	0.19	0.05	-0.01	2013	0.18	-0.04	0.36	-0.12	-0.03

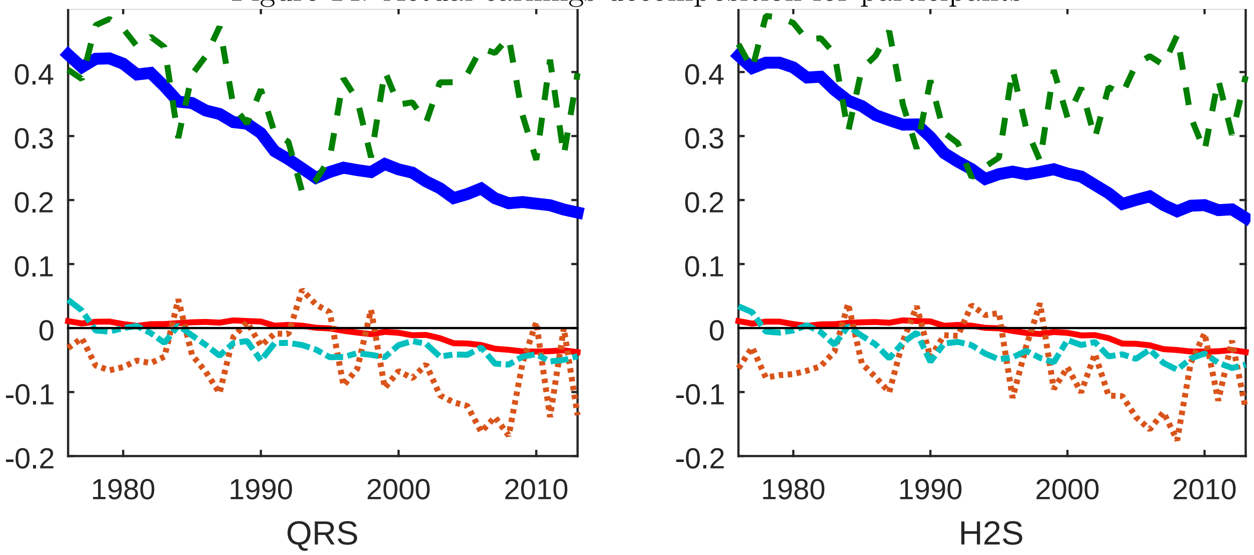
Notes: Total, EC, CC, SC and PC respectively denote total difference, endowments component, coefficients component, selection component and participation component.

Table 48: Mean decomposition, actual earnings for participants, heterogeneous copula by marital status

Year	Total	EC	CC	SC	PC	Year	Total	EC	CC	SC	PC
1976	0.44	0.01	0.53	-0.08	-0.03	1995	0.24	0.00	0.33	0.00	-0.08
1977	0.41	0.00	0.52	-0.07	-0.04	1996	0.26	-0.01	0.54	-0.20	-0.07
1978	0.42	0.00	0.56	-0.09	-0.05	1997	0.25	0.00	0.44	-0.10	-0.09
1979	0.42	0.01	0.59	-0.10	-0.07	1998	0.24	-0.01	0.36	-0.03	-0.08
1980	0.42	0.01	0.57	-0.11	-0.05	1999	0.26	-0.01	0.46	-0.13	-0.06
1981	0.40	0.00	0.49	-0.05	-0.04	2000	0.25	-0.01	0.40	-0.09	-0.05
1982	0.40	0.00	0.50	-0.05	-0.05	2001	0.24	-0.01	0.37	-0.07	-0.05
1983	0.38	0.00	0.50	-0.04	-0.07	2002	0.23	-0.01	0.34	-0.05	-0.05
1984	0.36	0.00	0.41	0.00	-0.05	2003	0.22	-0.01	0.46	-0.16	-0.07
1985	0.36	0.00	0.52	-0.09	-0.08	2004	0.20	-0.02	0.39	-0.10	-0.07
1986	0.34	0.01	0.52	-0.09	-0.08	2005	0.21	-0.02	0.44	-0.12	-0.09
1987	0.34	0.00	0.55	-0.13	-0.08	2006	0.22	-0.03	0.46	-0.14	-0.07
1988	0.32	0.00	0.34	0.03	-0.05	2007	0.20	-0.03	0.51	-0.18	-0.09
1989	0.32	0.01	0.33	0.03	-0.06	2008	0.19	-0.03	0.47	-0.16	-0.07
1990	0.31	0.01	0.48	-0.09	-0.09	2009	0.20	-0.04	0.43	-0.14	-0.06
1991	0.28	0.00	0.40	-0.07	-0.05	2010	0.20	-0.04	0.41	-0.11	-0.06
1992	0.27	0.00	0.41	-0.08	-0.06	2011	0.20	-0.04	0.49	-0.19	-0.06
1993	0.25	0.00	0.36	-0.03	-0.08	2012	0.18	-0.04	0.25	0.04	-0.08
1994	0.23	0.00	0.27	0.04	-0.07	2013	0.18	-0.04	0.42	-0.13	-0.07

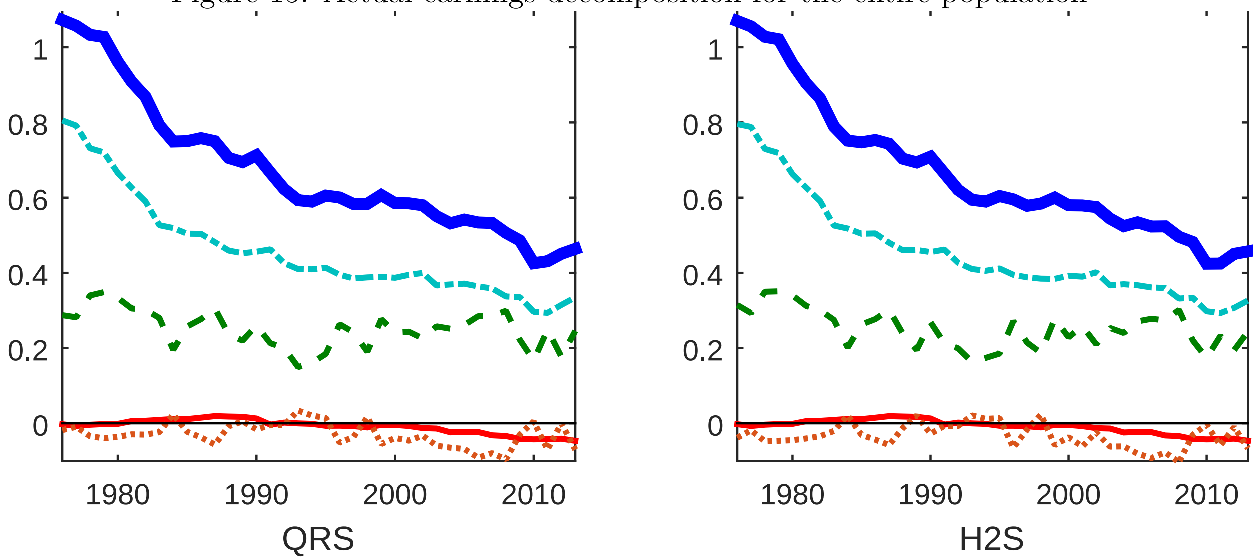
Notes: Total, EC, CC, SC and PC respectively denote total difference, endowments component, coefficients component, selection component and participation component.

Figure 14: Actual earnings decomposition for participants



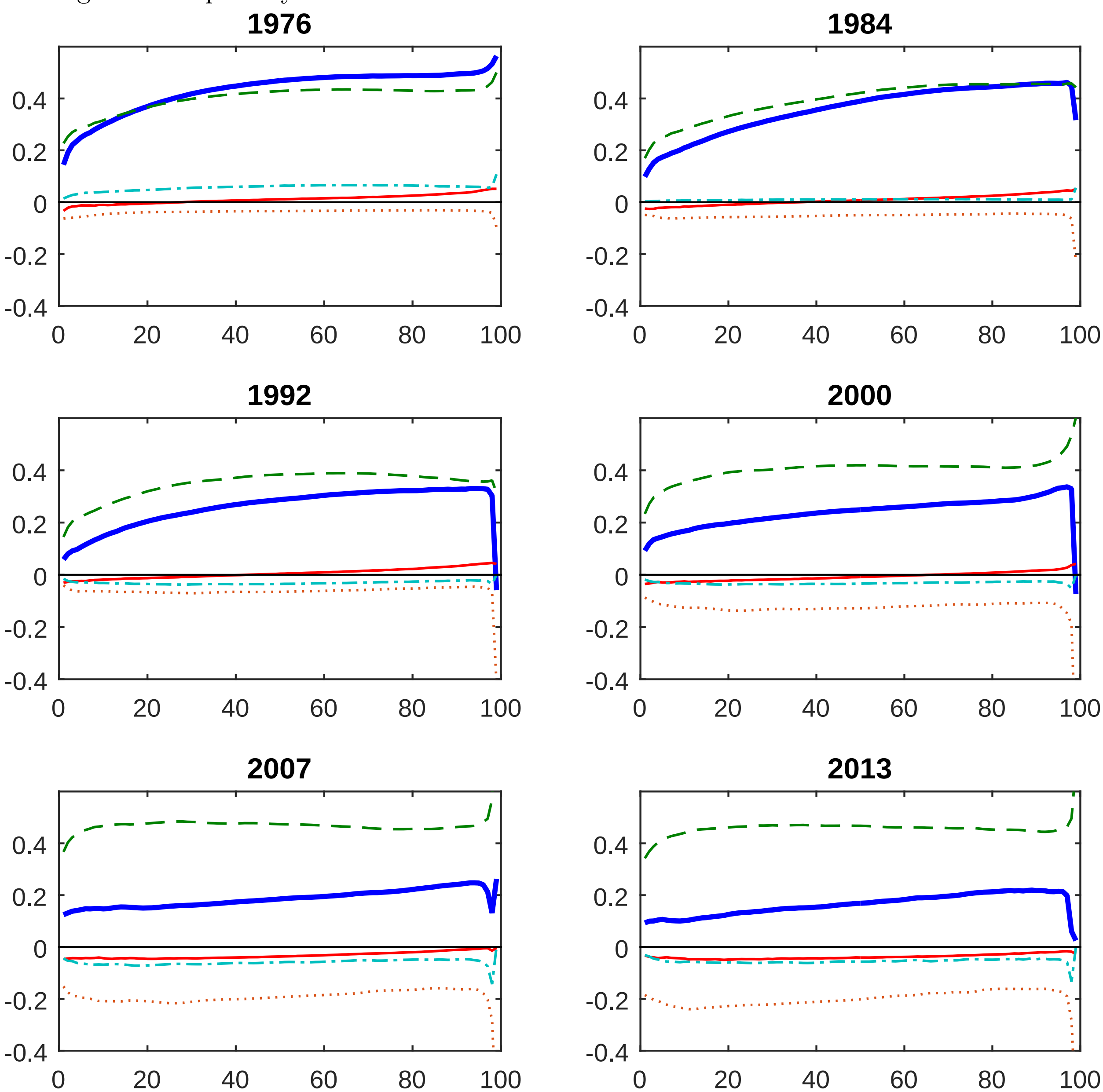
Notes: QRS and H2S stand for Quantile Regression with Selection and Heckman 2-Stage estimators; the solid thick blue line denotes the total gap between male and female workers; the solid thin red line denotes the endowments component; the dashed green line denotes the coefficients component; the dotted orange line denotes the selection component; the dashed-dotted cyan line denotes the participation component.

Figure 15: Actual earnings decomposition for the entire population



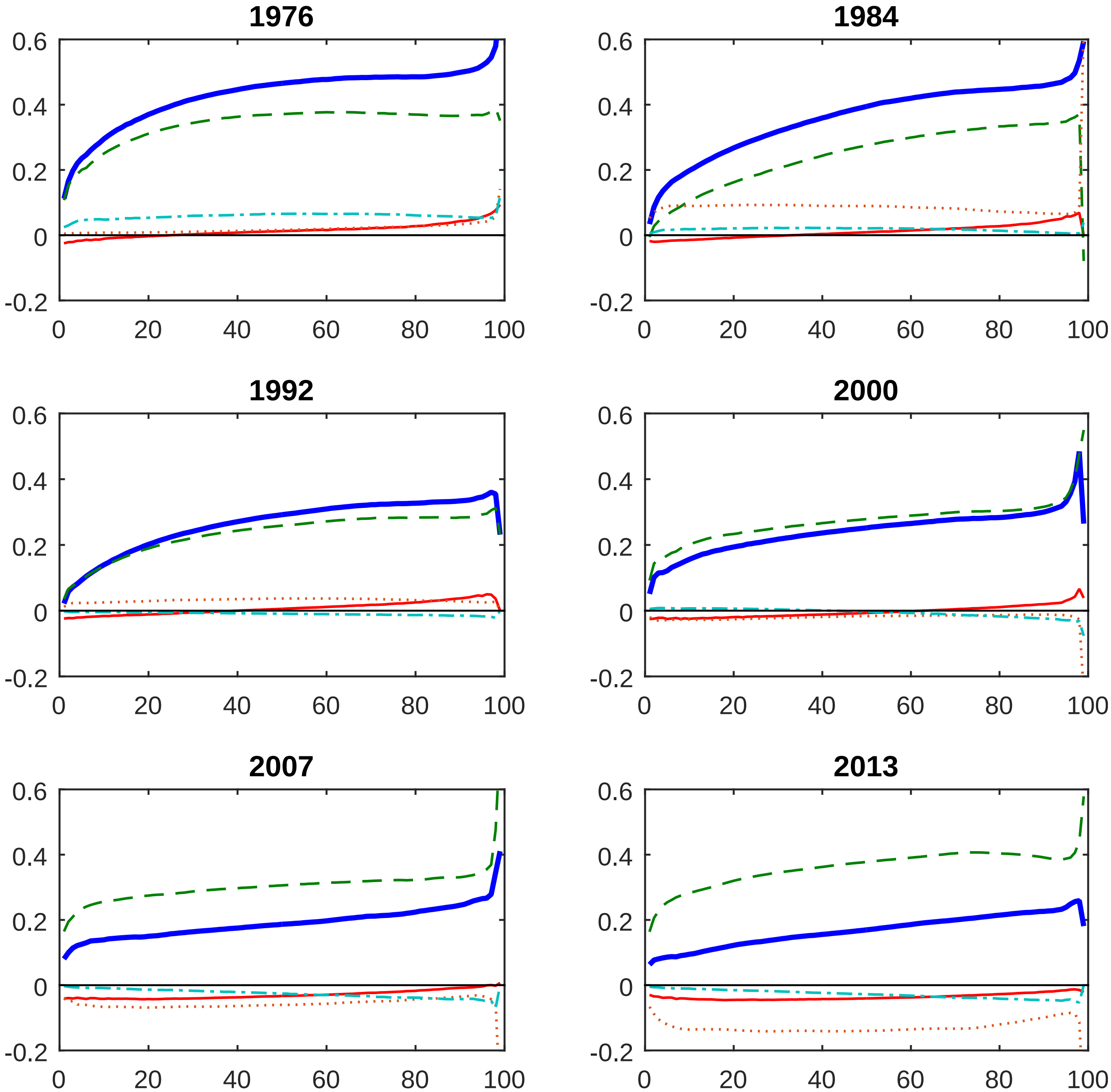
Notes: QRS and H2S stand for Quantile Regression with Selection and Heckman 2-Stage estimators; the solid thick blue line denotes the total gap between male and female workers; the solid thin red line denotes the endowments component; the dashed green line denotes the coefficients component; the dotted orange line denotes the selection component; the dashed-dotted cyan line denotes the participation component.

Figure 16: Unconditional quantiles decompositions, actual earnings for participants, heterogeneous copula by race



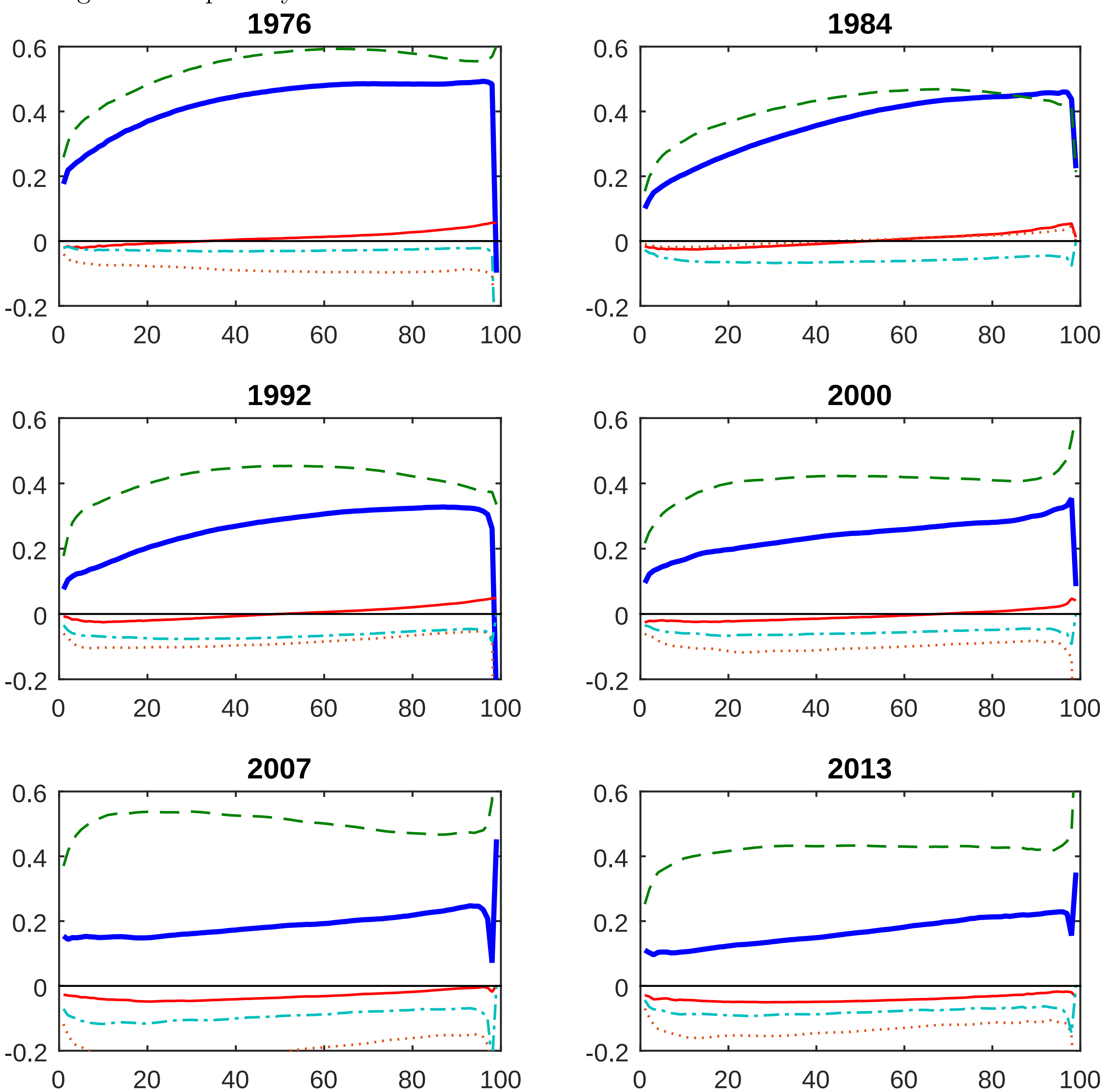
Notes: the solid thick blue line denotes the total gap between male and female workers; the solid thin red line denotes the endowments component; the dashed green line denotes the coefficients component; the dotted orange line denotes the selection component; the dashed-dotted cyan line denotes the participation component.

Figure 17: Unconditional quantiles decompositions, actual earnings for participants, heterogeneous copula by education



Notes: the solid thick blue line denotes the total gap between male and female workers; the solid thin red line denotes the endowments component; the dashed green line denotes the coefficients component; the dotted orange line denotes the selection component; the dashed-dotted cyan line denotes the participation component.

Figure 18: Unconditional quantiles decompositions, actual earnings for participants, heterogeneous copula by marital status



Notes: the solid thick blue line denotes the total gap between male and female workers; the solid thin red line denotes the endowments component; the dashed green line denotes the coefficients component; the dotted orange line denotes the selection component; the dashed-dotted cyan line denotes the participation component.