Why does the Job Corps increase gender earnings inequality?

Anthony Strittmatter*

University of St. Gallen

November 30, 2016

Abstract

Several studies considering the Job Corps find more positive earnings effects for males than for females. This effect heterogeneity in favor of males contrasts with the results of the majority of other training program evaluations. Applying the translated quantile approach of Bitler, Hoynes, and Domina (2014), I show that the main mechanism behind the surprising findings for the Job Corps operates through existing gender earnings inequality rather than Job Corps trainability differences by gender. A program assignment mechanism that balances the earnings structure could increase earnings opportunities without promoting gender inequality.

JEL-Classification: J38, C21

Keywords: Gender Inequality, Decomposition, Quantile Regression, Program Evaluation

*Previous versions of this manuscript circulated under different titles. The manuscript was presented at the SSES 2012 in Zurich, at the NOeG 2012 in Vienna, at the EALE 2013 in Torino, at the ESEM 2013 in Goteborg, at the IAAE 2014 in London, and at Humboldt University Berlin 2015. I thank the participants for their useful comments. In particular, I am grateful for helpful suggestions from Simone Balestra, Xuan Chen, and Bernd Fitzenberger. I thank Ozkan Eren and Serkan Ozbeklik, who kindly shared their data and programs with me. The usual disclaimer applies. Mail: anthony.strittmatter@unisg.ch

1 Introduction

In this study, I investigate the heterogeneous earnings effects of the Job Corps by gender. In particular, I show that gender differences in labor market opportunities as opposed to gender differences in Job Corps trainability is the main driver of effect heterogeneity. For this purpose, I apply the translated quantile decomposition of Bitler, Hoynes, and Domina (2014). Furthermore, I extend this approach to incorporate average translated effects.

The Job Corps is the largest U.S. labor market program targeting disadvantaged youths. It provides academic, vocational, and social training, as well as health care counseling and job search assistance, for an average duration of eight to nine months. Blanco, Flores, and Flores-Lagunes (2013), Chen and Flores (2015), and Lee (2009) show that the Job Corps has positive wage effects for both genders. This effect is supposed to increase earnings through higher hourly wages and greater labor supply.¹ The female labor supply is expected to respond more positively than the male labor supply to wage changes (see, e.g., the discussion in Bargain, Orsini, and Peichl, 2013, Killingsworth and Heckman, 1986). However, several studies find that the Job Corps increases the gender earnings inequality. In a large-scale experimental evaluation, Schochet, Burghardt, and McConnell (2008) show evidence of higher average earnings gains from the Job Corps for males than for females. Moreover, Eren and Ozbeklik (2014) report more positive quantile earnings effects for males than for females. These findings contrast with the empirical evaluations of many other active labor market programs (ALMPs). Meta studies conducted by Bergemann and van den Berg (2008) and Card, Kluve, and Weber (2015) report evidence of higher ALMP returns for females than for males. In this study, I show the potential mechanisms behind the unexpected findings for the Job Corps.

Frumento, Mealli, Pacini, and Rubin (2012) investigate the effect heterogeneity for different socio-economic groups using a principal stratification approach. They find that individuals with worse initial conditions (in terms of education, labor market experience, race, and gender) profit less from the Job Corps than those with good labor market opportunities.² Expected earnings

¹Chen, Flores, and Flores-Lagunes (2015) report positive employment effects of the Job Corps four years after assignment.

²This is another peculiarity of the Job Corps because for many other ALMPs, the returns do not increase with labor market opportunities (see, e.g., Card, Kluve, and Weber, 2015, Rinne, Schneider, and Uhlendorff, 2011,

can be viewed as an aggregate measure of labor market opportunities. Quantile regressions can identify heterogeneity by earnings. Eren and Ozbeklik (2014) report larger earnings differences for the Job Corps at higher quantiles than at lower quantiles. Such a pattern can be expected if the Job Corps is more effective for groups with better labor market opportunities, that is, those with high earnings even without training. Higher opportunity costs for participants with better earnings opportunities can rationalize this positive relationship between earnings opportunities and Job Corps returns. It is well documented that a larger share of females is located in the lower part than in the upper part of the earnings distribution. Thus, unexpected effect heterogeneity can arise in favor of males from structural gender earnings inequality rather than from differences in Job Corps trainability related to gender. The latter could arise from program contents, for example, that are more aligned to the needs of one gender.

This study contributes to the literature by decomposing the Job Corps' quantile effects into direct and structural components. The inter-gender comparison of conditional quantile treatment effects (CQTEs) can be considered in two parts. First, the earnings functions can be heterogeneous by gender, which is a direct moderator effect of gender. Second, quantiles are relative measures defined separately in each gender group. Thus, the same rank can correspond to different earnings quantiles in each gender group, which may be an additional reason for effect heterogeneity, and is the structural moderator effect of gender. Bitler, Hoynes, and Domina (2014) first explicitly consider the outcome structure by comparing the CQTEs of the different groups. They establish the translated quantile treatment effect (TQTE), which places the CQTE of each subgroup on the same absolute scale, enabling a Machado and Mata (2005) type decomposition to separate the direct and structural moderator effects.

For an example of the relevance of transformed parameters, assume the need to find assignment criteria for the Job Corps that do not increase gender inequality. If structural inequality leads to heterogeneity, a non-discriminatory assignment rule can be used to offer participation to males and females with balanced labor market opportunities. If Job Corps trainability differs by gender, then the curriculum of the Job Corps programs needs to be better aligned with the specific training demands of males and females. Accordingly, by determining how the Job Corps

among others).

increases gender earnings inequality, decision makers can choose an optimal non-discriminatory policy.

As with the existing literature, my findings suggest a larger CQTE for males than for females. However, TQTE shows no strong effect heterogeneity by gender. Structural gender earnings inequality accounts for 82% of the average effect heterogeneity by gender. These findings suggest that the offer to participate in the Job Corps shows no significant disfavor toward females with earnings opportunities that are equal to those of males. However, the Job Corps amplifies the existing structural earnings inequality, which is more favorable for males than females. Randomly offering participation in the Job Corps to eligible individuals increases the average earnings opportunities by 8% after 48 months but also increases the average gender earnings gap by 8% within the eligible group. An allocation scheme that balances the earnings structures of the selected awardees by gender yields approximately similar average returns but has significantly lower impacts on gender earnings inequality.

The remainder of the paper is structured as follows. The next section discusses the econometric background. Section 4 reviews the institutional background of the Job Corps and the related literature. Section 4 describes the experimental data. Section 5 explains the parameters of interest as well as the quantile transformation approach. Section 6 presents the empirical results, and the final section concludes the paper. Online Appendix A provides additional details on the estimation approach. Online Appendix B reports additional robustness checks.

2 Econometrics Background

Assume that earnings is a function of a Job Corps offer D , gender G , and an error term U that incorporates labor market opportunities. The structural earnings function is $\varphi_1(G, U)$ if there is an offer to join the Job Corps and $\varphi_0(G, U)$ if there is no offer. The error term possibly interacts non-separably with gender. The quantile earnings function is $Q_\tau(\varphi_d(G, U))$ for $d \in \{0, 1\}$.³ For an insight, assume that $\varphi_d(\cdot)$ is strictly increasing in its second argument. Then, the conditional quantile function $Q_\tau(\varphi_d(G, U)|G = g)$ equals $\varphi_d(g, Q_\tau(U|G = g))$, where

³ $Q_\tau(A)$ represents the τ th quantile of A .

$Q_\tau(U|G = g)$ is the quantile of the error term distribution for those with gender $g \in \{m, f\}$. The CQTEs for males ($g = m$) and females ($g = f$) are, respectively,

$$\begin{aligned}\delta_{CQTE}(\tau, m) &= \varphi_1(m, Q_\tau(U|G = m)) - \varphi_0(m, Q_\tau(U|G = m)) \text{ and} \\ \delta_{CQTE}(\tau, f) &= \varphi_1(f, Q_\tau(U|G = f)) - \varphi_0(f, Q_\tau(U|G = f)).\end{aligned}$$

Thus, even when the earnings functions are homogenous by gender (i.e., $\varphi_d(m, U) = \varphi_d(f, U)$), the CQTEs can be heterogeneous if $Q_\tau(U|G = m) \neq Q_\tau(U|G = f)$ (e.g., this is noted in Abadie, Angrist, and Imbens, 2002 and Bitler, Hoynes, and Domina, 2014, among others). The direct effect operates through heterogeneity caused by the first argument of the earnings functions. The structural effect operates through heterogeneity caused by the conditional quantiles. Simply by comparing $\delta_{CQTE}(\tau, m)$ and $\delta_{CQTE}(\tau, f)$, the direct and structural channels cannot be distinguished.

Bitler, Hoynes, and Domina (2014) use unique reference quantiles for their TQTE approach, which I denote as $Q_\tau(U_r)$. They then transform the CQTE to the following reference quantile:⁴

$$\begin{aligned}\delta_{TQTE}(\tau, m) &= \varphi_1(m, Q_\tau(U_r)) - \varphi_0(m, Q_\tau(U_r)) \text{ and} \\ \delta_{TQTE}(\tau, f) &= \varphi_1(f, Q_\tau(U_r)) - \varphi_0(f, Q_\tau(U_r)).\end{aligned}$$

Heterogeneity arises between $\delta_{TQTE}(\tau, m)$ and $\delta_{TQTE}(\tau, f)$ solely from the first argument of the earnings function. Following the approach by Bitler, Hoynes, and Domina (2014), I propose the quantiles from the potential earnings distribution under non-treatment as the reference distribution. With this reference distribution, I show that TQTE equals CQTE when no structural earnings inequality exists. Furthermore, I show that TQTE equals the (unconditional) quantile treatment effect (QTE) when structural inequality is solely responsible for the effect heterogeneity by gender. This improves the interpretability of TQTE.

⁴Powell (2016) proposes a generalized quantile regression (GQR), which also employs a transformation of conditional quantiles. The main purpose of GQR is to estimate unconditional quantiles under unrestrictive assumptions.

The conditional average treatment effects (CATE) are defined as

$$\delta_{CATE}(m) = \int_0^1 \delta_{CQTE}(\tau, m) d\tau \text{ and } \delta_{CATE}(f) = \int_0^1 \delta_{CQTE}(\tau, f) d\tau.$$

This implies that heterogeneity in the earnings structure can be carried through to average effects. Therefore, I extend the approach in Bitler, Hoynes, and Domina (2014) to incorporate average effects. This approach has analogies in the literature on conditional mean decomposition, such as in the seminal contributions of Blinder (1973) and Oaxaca (1973). These methods enable decomposition of the differences in earnings functions into explained and unexplained components (see Fortin, Lemieux, and Firpo, 2010, for an overview). However, in mean decomposition, the different outcome (here earnings) structures cannot explain heterogeneity. It is not possible to condition on the outcome variable. Bitler, Gelbach, and Hoynes (2014) document that mean comparisons by subgroups can miss important aspects of within-group heterogeneity. In contrast, the Machado–Mata-type decomposition that I propose can explain the differences in earnings functions with the within-group earnings structure itself.

3 Job Corps

The Job Corps was established in the United States in 1964 under the Economic Opportunity Act. The aim of the Job Corps is to provide counseling and to develop individualized programs that enhance participants' job opportunities. In particular, completing a vocational degree and earning General Educational Development (GED) credentials are two major goals of the Job Corps. Job Corps programs follow a holistic approach. They can include physical and mental health care counseling, various vocational and social education schemes, and job search assistance. Courses have theoretical and practical content and training can be partly on the job. Participants are trained in one of approximately 125 Job Corps centers and they typically live on Job Corps campuses during the program. Candidates must pass a screening and recruitment process conducted by certain admission agencies. Eligible participants are disadvantaged youths (aged 16–24 years) in need of additional education and training (possibly

mental training). Candidates must have low incomes to meet these eligibility criteria. Approximately 60,000 eligible youths enroll in the Job Corps each year. In 2008, total program expenditures exceeded 1.58 billion U.S. dollars (Job Corps Annual Report, 2008). Of the total expenditures, training costs account for 40%, support services (e.g., meals, lodging, medical care) costs account for 33%, and administration and infrastructure investment costs account for 26%. In 2008, expenditures per participant were approximately 26,000 U.S. dollars, or about ten times the cost of a typical program under the Workforce Investment Act (Eren and Ozbeklik, 2014).

Numerous studies investigate the effectiveness of Job Corps programs in improving participants' post-program earnings. Early studies using data from the 1970s find that males profited more from the Job Corps than females. Mallar, Kerachsky, Thornton, and Long (1982) find that the Job Corps increases annual earnings by 2,000 U.S. dollars for male and 1,000 U.S. dollars for female participants. Gay and Borus (1980) report negative earnings impacts for black female participants and positive effects for black male participants. Between 1994 and 1996, Mathematica Policy Research conducted a randomized National Job Corps Study (NJCS). This study had an experimental design and, hence, became the benchmark for evaluating the Job Corps. Using survey data collected from NJCS participants in 1998, Schochet, Burghardt, and McConnell (2008) find yearly earnings gains of 1,530 U.S. dollars for males, but only 1,134 U.S. dollars for females, participating in the Job Corps. These earnings differences are statistically not significant. However, Schochet, Burghardt, and McConnell (2008) report a significantly higher yearly earnings gain of 540 U.S. dollars for male than for female participants when using the administrative data of NJCS participants collected from the 1998 Summary Earnings Records (SER).

Furthermore, Schochet, Burghardt, and McConnell (2008) report the effects of receiving an offer to participate in the Job Corps, which could include actual participation in the program. Using the survey data collected from NJCS participants 48 months after randomization, Schochet, Burghardt, and McConnell (2008) find an average yearly earnings gains of 828 U.S. dollars, which translates into an average weekly earnings gains of 16 U.S. dollars. Using the

same data, Eren and Ozbeklik (2014) report weekly earnings gains of 18 U.S. dollars for male and 11 U.S. dollars for female awardees of an offer to join the Job Corps. Quantile differences suggest almost zero effects on the earnings distribution of males below the 4th decile. Above the 4th decile, the quantile earnings differences increase to approximately 30 U.S. dollars and remain at this positive level. For females, the quantile earnings differences reach approximately 20 U.S. dollars at the 15th percentile. Above this percentile, the positive quantile earnings differences remain almost stable. Thus, the offer to join the Job Corps increases gender earnings inequality. In particular, the upper part of the earnings distribution shows increases in inequality.

4 Experimental data

As previously mentioned, Mathematica Policy Research carried out a randomized experiment with the Job Corps. Of the 80,833 eligible individuals who first applied for participation in the program between November 1994 and February 1996, 15,386 were randomly selected for the experiment research sample. From this group, 9,409 were randomly selected to receive an offer to participate in the Job Corps. Out of these, 73% joined the Job Corps program and actually started to participate after an average duration of 1.4 months. The remaining 5,977 individuals were assigned to the control group. Control group members were not eligible for the Job Corps programs for three years following randomization (approximately 1% participated anyway). Nevertheless, most control group members received some sort of alternative training in the form of GED preparation (27%), high school (23%), and vocational, technical, or trade school (21%) (Burghardt et al., 2001). The intensity of the training is higher among treatment than among control group members. During the 48 months following the randomization, 47% of the treatment and 27% of the control group members obtained GED credentials. Likewise, 38% of the treatment and 15% of the control group members obtained a vocational degree. Approximately 5–8% of the eligible individuals in both groups received a high school diploma (Burghardt et al., 2001). To improve the accuracy of the research sample, some subgroups were oversampled (see description in Schochet, 2001). To account for this, I use the following

sampling weights whenever necessary.⁵

Interviews were conducted with members of the research sample 12, 30, and 48 months after randomization. This study considers publicly released interviews of the sample of individuals conducted 48 months after randomization, eliminating observations with missing information (3,951 observations are dropped). In a large-scale bound analysis, Lee (2009) investigates the potential attrition bias in the NJCS data. In his approach, sample attrition does not strongly alter the results (see Remark 2 in Lee, 2009). I use the non-response adjusted sampling weights (*wgt48b*) provided in the NJCS data to account for sample attrition. The final sample includes 10,595 observations, of which 6,372 are in the treatment sample and 4,223 are in the control sample.

The outcome of interest is weekly earnings four years after random assignment, which is more than two years after 92% of the participants completed the program. I investigate the effects of the randomized offer to participate in the Job Corps program. This intention-to-treat effect (ITT) differs from the effect of actual participation. The ITT effect is relevant to investigate the optimal selection rules for awarding offers to join the Job Corps program. The award of an offer to join the Job Corps program is under the control of the decision makers; however, enrollment and actual participation are not under their control. An alternative approach is to investigate the local average treatment effects of actual participation. This approach identifies the effects for only those individuals accepting the award and not for all eligible individuals. Chen, Flores, and Flores-Lagunes (2015) address this limitation with an interesting bound analysis to identify the effects for individuals rejecting the offer to participate. Frumento, Mealli, Pacini, and Rubin (2012) overcome this drawback using a principal stratification approach. I focus only on the ITT effects in this study.

Table 1 reports the means and standard errors of the observed variables. Four years after randomized assignment, individuals with an offer to join the Job Corps earn on average 15 U.S. dollars per week more than individuals without an offer. Furthermore, I report the descriptive statistics for socio-economic characteristics, employment, earnings, and welfare histories, as well

⁵Eren and Ozbeklik (2014) condition key features of the sample design in a regression framework instead of using sampling weights. In unreported robustness checks, I find that my results are not sensitive to different ways of constructing sampling weights.

Table 1: Means and standard errors by offer to join Job Corps (JC).

	Offer the join JC		Not invited to JC		Standardized
	Mean	Std. Err.	Mean	Std. Err.	Difference
	(1)	(2)	(3)	(4)	(5)
Earnings per week in Year 4	210.854	200.359	195.846	187.475	7.740
Ever enrolled in a Job Corps center	0.733	0.442	0.011	0.106	224.320
Socio-economic characteristics					
Female	0.412	0.492	0.406	0.491	1.200
Aged between 16-17 years	0.414	0.493	0.417	0.493	0.590
Aged between 18-19 years	0.317	0.465	0.317	0.465	0.070
Aged between 20-24 years	0.269	0.443	0.266	0.442	0.580
White	0.277	0.447	0.266	0.442	2.470
Black	0.479	0.500	0.478	0.500	0.120
Hispanic	0.171	0.377	0.180	0.384	2.420
Indian	0.074	0.261	0.076	0.265	0.900
Lives with spouse or partner	0.061	0.239	0.063	0.243	0.980
Dummy for bad health	0.127	0.333	0.137	0.344	3.090
0-6 months education program in last year	0.268	0.442	0.254	0.435	2.990
6-12 months education program in last year	0.352	0.477	0.368	0.482	3.360
High school credential	0.228	0.420	0.238	0.426	2.380
Lives in PMSA	0.319	0.466	0.327	0.469	1.610
Lives in MSA	0.459	0.498	0.454	0.498	1.110
Past employment and earnings history					
Ever had job for two weeks or more	0.803	0.398	0.792	0.406	2.750
Worked in year prior to random assignment	0.653	0.476	0.645	0.478	1.750
Has job at random assignment	0.211	0.408	0.206	0.404	1.300
Below 3 months employed in last year	0.190	0.380	0.189	0.379	0.270
3-9 months employed in last year	0.282	0.437	0.276	0.433	1.510
9-12 months employed in last year	0.181	0.373	0.180	0.372	0.230
Yearly earnings less than \$1,000	0.107	0.309	0.111	0.314	1.470
Yearly earnings \$1,000 to \$5,000	0.273	0.446	0.264	0.441	2.060
Yearly earnings \$5,000 to \$10,000	0.138	0.345	0.133	0.339	1.530
Yearly earnings above \$10,000	0.067	0.250	0.067	0.250	0.060
Past welfare history					
Family on welfare when growing up	0.200	0.400	0.193	0.395	1.530
Received food stamps in last year	0.435	0.496	0.435	0.496	0.090
Public or rent-subsidized housing	0.204	0.400	0.193	0.392	2.940
Received AFDC in last year	0.300	0.458	0.297	0.457	0.620
Drugs and crime					
Used hard drugs in last year	0.068	0.251	0.063	0.243	1.970
Smoked marijuana in last year	0.249	0.432	0.243	0.429	1.380
Ever arrested dummy	0.250	0.433	0.253	0.434	0.700
Observations	6,372		4,223		

Note: Time-varying control variables are measured at the time of random assignment. Weights accounting for the sampling design and non-response of the 48-month interview are used (*wgt48b*). JC stands for Job Corps. PMSA stand for Primary Metropolitan Statistical Area. MSA stand for Metropolitan Statistical Area. AFDC stands for Aid to Families with Dependent Children.

as drug use and arrests. For all of these observed control variables, I find little difference between individuals with and without an offer to join the Job Corps. The standardized differences

Table 2: Means and standard errors by gender.

	Males		Females		Standardized Difference (5)
	Mean (1)	Std. Err. (2)	Mean (3)	Std. Err. (4)	
Earnings per week in Year 4	230.310	206.581	164.308	167.071	35.130
Offer to join Job Corps	0.495	0.500	0.502	0.500	1.220
Ever enrolled in a Job Corps center	0.380	0.485	0.351	0.477	6.140
Socio-economic characteristics					
Aged between 16-17 years	0.435	0.496	0.386	0.487	9.960
Aged between 18-19 years	0.310	0.463	0.327	0.469	3.620
Aged between 20-24 years	0.254	0.436	0.287	0.452	7.220
White	0.305	0.460	0.222	0.415	19.010
Black	0.450	0.497	0.519	0.500	13.950
Hispanic	0.169	0.375	0.185	0.388	4.180
Indian	0.076	0.265	0.074	0.261	0.820
Lives with spouse or partner	0.051	0.219	0.078	0.268	11.170
Dummy for bad health	0.121	0.326	0.147	0.355	7.720
0-6 months education program in last year	0.243	0.429	0.287	0.452	9.900
6-12 months education program in last year	0.385	0.486	0.324	0.468	12.660
High school credential	0.198	0.398	0.285	0.451	20.500
Lives in PMSA	0.316	0.465	0.334	0.471	3.770
Lives in MSA	0.440	0.496	0.480	0.500	7.910
Past employment and earnings history					
Ever had job for two weeks or more	0.809	0.393	0.779	0.415	7.490
Worked in year prior to random assignment	0.664	0.472	0.628	0.483	7.430
Has job at random assignment	0.212	0.409	0.203	0.402	2.240
Below 3 months employed in last year	0.190	0.378	0.190	0.382	0.080
3-9 months employed in last year	0.287	0.437	0.268	0.432	4.380
9-12 months employed in last year	0.187	0.377	0.171	0.366	4.360
Yearly earnings less than \$1,000	0.099	0.299	0.123	0.328	7.590
Yearly earnings \$1,000 to \$5,000	0.269	0.444	0.268	0.443	0.340
Yearly earnings \$5,000 to \$10,000	0.143	0.350	0.124	0.330	5.600
Yearly earnings above \$10,000	0.079	0.269	0.050	0.219	11.530
Past welfare history					
Family on welfare when growing up	0.174	0.379	0.229	0.420	13.850
Received food stamps in last year	0.366	0.482	0.534	0.499	34.290
Public or rent-subsidized housing	0.183	0.384	0.222	0.413	9.790
Received AFDC in last year	0.222	0.416	0.408	0.491	40.860
Drugs and crime					
Used hard drugs in last year	0.076	0.265	0.050	0.218	10.620
Smoked marijuana in last year	0.277	0.448	0.200	0.400	18.250
Ever arrested dummy	0.311	0.463	0.164	0.370	35.180
Observations	6,104		4,491		

Note: Time-varying control variables are measured at the time of random assignment. Weights accounting for the sampling design and non-response of the 48-month interview are used (*wgt48b*). PMSA stand for Primary Metropolitan Statistical Area. MSA stand for Metropolitan Statistical Area. AFDC stands for Aid to Families with Dependent Children.

are always below 3.5.⁶ This confirms that the offer to join the Job Corps was randomized

⁶The standardized difference of variable X between samples A and B is defined as

$$SD = \frac{|\bar{X}_A - \bar{X}_B|}{\sqrt{\frac{1}{2}(Var(X_A) + Var(X_B))}} \cdot 100,$$

appropriately.

Table 2 documents the descriptive statistics of observed variables by gender without conditioning on an offer to join the Job Corps. On average, males earn 66 U.S. dollars per week more than females four years after the randomized offer. Relative to males, females in the sample were less likely to be white, more likely to have high school credentials, less likely used drugs the previous year, and less often arrested. Females received food stamps and Aid to Families with Dependent Children (AFDC) more often than males. The families of females in the sample were on welfare during their childhood more often than the families of males.

5 Empirical approach

5.1 Treatment effects

Let D_i be a binary variable indicating whether or not individual i receives an offer to participate in the Job Corps program ($D_i = 1$ or $D_i = 0$). Following the potential outcome framework in Rubin (1974), $Y_i(1)$ is the potential earnings under assignment to and possible participation in the Job Corps. $Y_i(0)$ is the potential earnings when individual i does not receive the offer to participate in the Job Corps (for $i = 1, \dots, N$). For each individual, only the realized earnings,

$$Y_i = Y_i(1) \cdot D_i + Y_i(0) \cdot (1 - D_i),$$

is observed, implying that either $Y_i(1)$ or $Y_i(0)$ is counterfactual. For an analogy to the earnings functions, potential earnings under treatment are $Y_i(1) = \varphi_1(G_i, U_i)$ and potential earnings under non-treatment are $Y_i(0) = \varphi_0(G_i, U_i)$. The individual-specific causal effects are

$$\delta_i = Y_i(1) - Y_i(0) = \varphi_1(G_i, U_i) - \varphi_0(G_i, U_i).$$

where $\bar{X}_A$ denotes the mean of sample A and $\bar{X}_B$ denotes the mean of sample B . Rosenbaum and Rubin (1983) consider an absolute standardized difference exceeding 20 to be "large."

We cannot identify individual-specific causal effects using cross-sectional data without strong assumptions such as rank preservation. Rank preservation implies that individuals do not change their rank because of their treatment status (see discussion in Firpo, 2007).⁷

Nevertheless, quantile differences can be identified without assuming rank preservation. Assume that $F_{Y(d)}(y) = Pr(Y_i(d) \leq y)$ denotes the cumulative distribution function of $Y_i(d)$. The τ th quantile of $Y_i(d)$ is

$$Q_{Y(d)}(\tau) = \inf\{y : F_{Y(d)}(y) \geq \tau\},$$

for the real $\tau \in (0, 1)$. The potential quantile $Q_{Y(d)}(\tau)$ is the minimum value of $Y_i(d)$ such that at least the share τ of the earnings distribution lies below this value. Quantile treatment effects (QTEs) are defined as

$$\delta_{QTE}(\tau) = Q_{Y(1)}(\tau) - Q_{Y(0)}(\tau),$$

the difference between the potential quantiles. QTEs identify the horizontal differences in potential earnings distributions at specific ranks.

Furthermore, assume that $F_{Y(d)|G}(y|g) = Pr(Y_i(d) \leq y | G_i = g)$ denotes the potential earnings distribution of $Y_i(d)$ conditional on gender g for $g \in \{m, f\}$. Then, the τ th quantile of $Y_i(d)$ conditional on gender is

$$Q_{Y(d)|G}(\tau|g) = \inf\{y : F_{Y(d)|G}(y|g) \geq \tau\}.$$

The potential conditional quantile $Q_{Y(d)|G}(\tau|g)$ is the minimum value of $Y_i(d)$ such that at least the share τ of the conditional earnings distribution lies below this value. The CQTEs are defined as

$$\delta_{CQTE}(\tau, g) = Q_{Y(1)|G}(\tau|g) - Q_{Y(0)|G}(\tau|g),$$

the difference between the conditional potential quantiles.

⁷Chernozhukov and Hansen (2005) use the term rank invariance for this assumption. They also introduce rank similarity, which is a weaker assumption than rank preservation because it allows for random deviations in ranks. To overcome limitations in point estimations, Heckman, Smith, and Clements (1997) suggest estimating bounds on the individual-specific causal effects under monotonicity assumptions.

Finally, the average treatment effect (ATE) is the expected value of δ_i ,

$$\delta_{ATE} = \int_0^1 \delta_{QTE}(\tau) d\tau = E[Y_i(1) - Y_i(0)],$$

and the CATE is

$$\delta_{CATE}(g) = \int_0^1 \delta_{CQTE}(\tau, g) d\tau = E[Y_i(1) - Y_i(0) | G_i = g].$$

Both average effects can be identified without assuming rank preservation.

5.2 Quantile transformations

One non-separable method to incorporate the earnings structure is to fix the value of the non-treated outcome at which CQTEs are measured. This requires a unique reference distribution, the quantiles of which are used as fixed anchor points. The CQTEs of different groups are transformed to these reference quantiles to place each CQTE on the same absolute scale. The required quantile transformations can be conceptualized through a relative distribution (see Morris, Bernhardt, and Handcock, 1994). Although the choice of reference distribution is obviously crucial, generally there is no best choice. I propose using the potential earnings distribution under non-treatment as a reference distribution and define the relative distribution as

$$R(\tau, g) = F_{Y(0)|G}(Q_{Y(0)}(\tau) | g).$$

This relative distribution can be interpreted as the proportion of the conditional population whose earnings are below the τ^{th} potential earnings quantile of the population under non-treatment. Subsequently, I demonstrate why this relative distribution allows for a meaningful decomposition of CQTE. Nevertheless, many alternative measures are available for the relative distribution, and these can also be meaningful (e.g., the observed outcome distribution).

The translated quantile treatment effect (TQTE) is defined as

$$\delta_{TQTE}(\tau, g) = Q_{Y(1)|G}(R(\tau, g) | g) - Q_{Y(0)|G}(R(\tau, g) | g), \quad (1)$$

implying that TQTE measures the horizontal distance between two conditional potential outcome distributions. Note that the only difference between CQTE and TQTE is that the location has shifted from τ to $R(\tau, g)$. TQTEs make a mapping between the conditional and reference distributions. If $F_{Y(0)|G}(y|g)$ is strictly increasing in the interval $(0, 1)$ such that $Q_{Y(0)|G}(F_{Y(0)|G}(y|g)|g) = y$, then TQTE can be rearranged as

$$\delta_{TQTE}(\tau, g) = Q_{Y(1)|G}(F_{Y(0)|G}(Q_{Y(0)}(\tau)|g)|g) - Q_{Y(0)}(\tau). \quad (2)$$

This parameter is strongly related to the changes-in-changes estimator proposed by Athey and Imbens (2006). The structural conditional quantile treatment effect (SQTE) can be defined as

$$\delta_{SQTE}(\tau, g) = \delta_{CQTE}(\tau, g) - \delta_{TQTE}(\tau, g).$$

This describes the difference between CQTE and TQTE. In the following section, I illustrate the interpretation of TQTE and SQTE in two important special cases.

First, assume that the structural inequality between males and females is zero. This implies that $F_{Y(0)|G}(y|g) = F_{Y(0)}(y)$ for all y and $Q_{Y(0)|G}(\tau|g) = Q_{Y(0)}(\tau)$ for all τ . Further, assume that $Y_i(0)$ is continuously distributed such that $F_{Y(0)|G}(Q_{Y(0)|G}(\tau|g)|g) = \tau$. Now, equation (1) can be rewritten as

$$\begin{aligned} \delta_{TQTE}(\tau, g) &= Q_{Y(1)|G}(F_{Y(0)|G}(Q_{Y(0)|G}(\tau|g)|g)|g) - Q_{Y(0)|G}(F_{Y(0)|G}(Q_{Y(0)|G}(\tau|g)|g)|g), \\ &= Q_{Y(1)|G}(\tau|g) - Q_{Y(0)|G}(\tau|g). \end{aligned}$$

This shows that $\delta_{TQTE}(\tau, g) = \delta_{CQTE}(\tau, g)$ and $\delta_{SQTE}(\tau, g) = 0$.⁸ Thus, CQTEs have no structural component and the entire effect heterogeneity can be directly associated with gender; that is, $\delta_{TQTE}(\tau, f) - \delta_{TQTE}(\tau, m) = \delta_{CQTE}(\tau, f) - \delta_{CQTE}(\tau, m)$ and $\delta_{SQTE}(\tau, f) - \delta_{SQTE}(\tau, m) =$

⁸The continuity assumption can be relaxed. Without assuming continuity, Athey and Imbens (2006) show that $Q_{Y(0)|G}(F_{Y(0)|G}(Q_{Y(0)|G}(\tau|g)|g)|g) = Q_{Y(0)|G}(\tau|g)$ (see Lemma A.1 in Athey and Imbens, 2006). Further, the condition $Q_{Y(1)|G}(F_{Y(0)|G}(Q_{Y(0)|G}(\tau|g)|g)|g) = Q_{Y(1)|G}(\tau|g)$ will hold if $F_{Y(0)|G}(y|g)$ is at least as smooth as $F_{Y(1)|G}(y|g)$. This means that the step size to the next higher rank must satisfy the condition $F_{Y(0)|G}(Q_{Y(0)|G}(\tau|g)|g) \leq F_{Y(1)|G}(Q_{Y(1)|G}(\tau|g)|g)$ for all τ . However, the assumption of continuity is more intuitive and appears reasonable in many applications.

0.

Second, assume that the entire gender effect heterogeneity can be explained by structural earnings differences. This implies that the transformed quantiles are equal for females and males,

$$Q_{Y(1)|G}(F_{Y(0)|G}(Q_{Y(0)}(\tau)|f)|f) = Q_{Y(1)|G}(F_{Y(0)|G}(Q_{Y(0)}(\tau)|m)|m), \quad (3)$$

because conditional potential quantiles would not be heterogeneous by gender after rank adjustment. If $Y_i(d)$ is continuously distributed such that $F_{Y(d)|G}(Q_{Y(d)|G}(\tau|g)|g) = \tau$, equation (3) can be rearranged as

$$F_{Y(0)|G}(Q_{Y(0)}(\tau)|f) = F_{Y(1)|G}(Q_{Y(1)|G}(F_{Y(0)|G}(Q_{Y(0)}(\tau)|m)|m)|f).$$

The law of iterative expectations for distribution implies that $F_{Y(1)}(y) = Pr(G_i = f)F_{Y(1)|G}(y|f) + Pr(G_i = m)F_{Y(1)|G}(y|m)$. Thus,

$$\begin{aligned} Pr(G_i = f)F_{Y(0)|G}(Q_{Y(0)}(\tau)|f) &= F_{Y(1)}(Q_{Y(1)|G}(F_{Y(0)|G}(Q_{Y(0)}(\tau)|m)|m)) \\ &\quad - Pr(G_i = m)F_{Y(1)|G}(Q_{Y(1)|G}(F_{Y(0)|G}(Q_{Y(0)}(\tau)|m)|m)|m). \end{aligned}$$

Further rearrangements using the previous continuity assumption leads to

$$\begin{aligned} Pr(G_i = f)F_{Y(0)|G}(Q_{Y(0)}(\tau)|f) + Pr(G_i = m)F_{Y(0)|G}(Q_{Y(0)}(\tau)|m) \\ = F_{Y(1)}(Q_{Y(1)|G}(F_{Y(0)|G}(Q_{Y(0)}(\tau)|m)|m)). \end{aligned}$$

The left-hand side of this equation is nothing but the marginal distribution $F_{Y(0)}(y) = Pr(G_i = f)F_{Y(0)|G}(y|f) + Pr(G_i = m)F_{Y(0)|G}(y|m)$ evaluated at $Q_{Y(0)}(y)$. This leads to the equality $\tau = F_{Y(1)}(Q_{Y(1)|G}(F_{Y(0)|G}(Q_{Y(0)}(\tau)|m)|m))$ under the previous continuity assumption. If $F_{Y(1)}(\tau)$ is strictly increasing in the interval $(0, 1)$ such that $Q_{Y(1)}(F_{Y(1)}(y)) = y$, this result implies that

$$Q_{Y(1)}(\tau) = Q_{Y(1)|G}(F_{Y(0)|G}(Q_{Y(0)}(\tau)|g)|g)$$

under condition (3).⁹ When this finding is plugged into equation (2), $\delta_{TQTE}(\tau, g) = \delta_{QTE}(\tau)$ and $\delta_{TQTE}(\tau, f) - \delta_{TQTE}(\tau, m) = 0$. Further, $\delta_{SQTE}(\tau, g) = \delta_{CQTE}(\tau, g) - \delta_{QTE}(\tau)$, implying that the entire inter-gender CQTE difference is assigned to the structural effect $\delta_{SQTE}(\tau, f) - \delta_{SQTE}(\tau, m) = \delta_{CQTE}(\tau, f) - \delta_{CQTE}(\tau, m)$.

In sum, TQTE equals CQTE when the structural earnings difference is zero. In contrast, TQTE equals QTE when the heterogeneity between CQTE for males and females can be fully explained by structural earnings inequality. This is a meaningful interpretation for TQTE because QTE represents the unconditional quantile effects.

Finally, the translated conditional average treatment effect (TATE) can be shown as

$$\delta_{TATE}(g) = \int_0^1 \delta_{TQTE}(\tau, g) d\tau = E[Q_{Y(1)|G}(F_{Y(0)|G}(Y_i(0)|g)|g)] - E[Y_i(0)],$$

as an analogy to equation (2), whereas the structural conditional average treatment effect (SATE) can be shown as

$$\delta_{SATE}(g) = \int_0^1 \delta_{SQTE}(\tau, g) d\tau = \delta_{CATE}(g) - \delta_{TATE}(g).$$

These parameters enable the average decomposition to account for differences in outcome structures. The interpretations of TATE and SATE are analogous to the interpretations of TQTE and SQTE, respectively. When structural earnings inequality equals zero, then TATE equals CATE and SATE equals zero. When the difference between the CATE for males and females can be fully explained by structural earnings inequality, TATE equals ATE and SATE is the residual between CATE and ATE.

⁹This quantile-quantile stability condition is strongly related to the rank stability condition $F_{Y(1)|G}(Q_{Y(1)}(\tau)|g) = F_{Y(0)|G}(Q_{Y(0)}(\tau)|g)$, which is used in Bitler, Gelbach, and Hoynes (2008) and Frandsen and Lefgren (2015) as a falsification test for rank preservation. Essentially, their findings imply that the assumption of rank preservation must be rejected if the quantile-quantile stability condition fails. TQTEs implicitly test this condition. Although the results of Section 6 cannot reject the quantile-quantile stability condition, possibly other variables than gender are associated with rank reorganization.

5.3 Estimation

For an extensive study on estimating counterfactual distributions under different identification strategies and model specifications, see, e.g., Chernozhukov, Fernández-Val, and Melly (2013) and Fortin, Lemieux, and Firpo (2010). In this study, I employ the quantile regression estimator proposed by Koenker and Bassett (1978). This estimator minimizes the weighted absolute deviations. I run a bivariate regression of the treatment dummy on earnings. The potential earnings quantiles are identified under the randomized research design. The most challenging estimation steps are quantile transformations. The first application of quantile transformation in labor economics dates back to at least Juhn, Murphy, and Pierce (1991). I follow the estimation approach proposed by Athey and Imbens (2006). They replace the potential outcome distributions by empirical distribution functions. Transformed quantiles can be obtained using a plug-in approach between the empirical quantiles and ranks of the empirical distribution functions. The average effects are estimated by replacing the expected values with the transformed sample averages that incorporate quantile transformation. Online Appendix A explains each estimation step in greater detail. Athey and Imbens (2006) show that the resulting estimates are $\sqrt{n}$ -consistent and asymptotically normal. I use a non-parametric bootstrap approach to calculate the standard errors. Chaisemartin and D'Haultfeuille (2015) show the validity of the bootstrap approach for this estimation. Melly and Santangelo (2015) extend the quantile transformation approach to incorporate control variables.

Additionally, I report the rank invariant Kolmogorov–Smirnov test statistics for all effects. The test statistics are

$$T_{KS} = \sqrt{n} \cdot \sup_{\tau} (|\hat{\delta}(\tau)|),$$

where $\hat{\delta}(\tau)$ indicates the estimate of some quantile difference $\delta(\tau)$. The null hypothesis is $\delta(\tau) = 0$ for all τ . Thus, this is a test for significant effects at some quantile or, in other words, for the inequality between some potential earning distributions. Further, I report signed versions of the Kolmogorov–Smirnov statistics, which enable testing for positive or negative first-order stochastic dominance. Positive stochastic dominance is defined as $\delta(\tau) \geq 0$ for all τ . Accordingly, under (positive) stochastic dominance, $\delta(\tau)$ is never negative regardless of

the quantile at which the effect is measured. In contrast, negative stochastic dominance is defined as $\delta(\tau) \leq 0$ for all τ . Accordingly, under negative stochastic dominance, $\delta(\tau)$ is never positive, regardless of the quantile at which the effect is measured.¹⁰ For all Kolmogorov–Smirnov-type test statistics, I estimate p-values using the recentered bootstrapping approach proposed by Chernozhukov and Fernández-Val (2005). This non-pivotal method computes p-values by avoiding restrictive assumptions. Chernozhukov and Fernández-Val (2005) show that recentering can substantially improve the finite-sample power of the Kolmogorov–Smirnov test relative to its uncentered counterpart.

6 Empirical results

6.1 QTE and CQTE

The empirical results suggest that quantile differences increase with rank. Figure 1 reports the QTE of the Job Corps on weekly earnings 48 months after the randomized assignment for all percentiles. Additionally, column 1 of Table 3 documents the deciles of the QTE, including standard errors. Figure 2 shows the potential earnings distributions. The QTEs lower than the 2nd decile do not differ from zero because of the unemployed with no earnings. Between the 2nd and 5th deciles, the earnings gains are approximately 10 U.S. dollars. The earnings gains further increase to approximately 20 U.S. dollars between the 5th and 9th deciles. Beyond the 9th decile, the quantile earnings differences reach values higher than 100 U.S. dollars. The Kolmogorov–Smirnov test results at the bottom of Table 3 support the findings that the Job Corps has a significant positive effect on earnings distribution. The potential earnings distribution under treatment stochastically dominates the potential earnings distribution under non-treatment.

The patterns found for CQTE by gender are somewhat different and are reported in Figure 3 and columns 2 and 3 of Table 3. For males, the CQTEs are close to zero until the 4th decile. They jump to an approximate 25 U.S.-dollar positive earnings above the 4th decile.

¹⁰In particular, I use the test statistics for positive stochastic dominance $T_{PSD} = \sqrt{n} \cdot \sup_{\tau}(\hat{\delta}(\tau))$ and negative stochastic dominance $T_{NSD} = \sqrt{n} \cdot \inf_{\tau}(\hat{\delta}(\tau))$.

Figure 1: QTEs of the Job Corps on weekly earnings (in U.S. dollars) 48 months after randomized assignment.

Note: The lines report the point estimates calculated separately for all percentiles. Crosses on the lines indicate significant effects at the 10% level. Hollow circles on the lines indicate significant effects at the 5% level. Full diamonds on the lines indicate significant effects at the 1% level.

Figure 2: Potential earnings distribution of all eligible Job Corps candidates.

Note: The treatment is a offer to join the Job Corps program. The outcome is weekly earnings (in U.S. dollars) 48 months after randomized assignment. The figure is truncated by earnings of 600 U.S. dollars per week.

Figure 3: CQTE of the Job Corps program by gender on weekly earnings (in U.S. dollars) 48 months after randomized assignment.

Note: The lines report the point estimates calculated separately for all percentiles. Crosses on the lines indicate significant effects at the 10% level. Hollow circles on the lines indicate significant effects at the 5% level. Full diamonds on the lines indicate significant effects at the 1% level.

Figure 4: Difference between the Job Corps CQTEs for females and males.

Note: The outcome is weekly earnings (in U.S. dollars) 48 months after randomized assignment. The lines report the point estimates calculated separately for all percentiles. Crosses on the lines indicate significant effects at the 10% level. Hollow circles on the lines indicate significant effects at the 5% level. Full diamonds on the lines indicate significant effects at the 1% level.

Table 3: QTE and CQTE of the Job Corps program by gender on weekly earnings (in U.S. dollars) 48 months after randomized assignment.

	QTE	CQTE		Difference between
	(1)	Females (2)	Males (3)	(2) and (3) (4)
20% Quantile	13.01*** (4.166)	4.425 (3.326)	8.011 (8.133)	-3.585 (8.730)
30% Quantile	7.555 (5.943)	16.68*** (6.134)	2.743 (8.735)	13.94 (10.79)
40% Quantile	8.414 (6.356)	19.12** (8.163)	7.055 (9.690)	12.06 (12.84)
50% Quantile	16.21*** (4.490)	9.059 (9.225)	25.08*** (7.387)	-16.03 (11.70)
60% Quantile	17.93*** (4.833)	11.99** (6.218)	17.25*** (7.749)	-5.261 (9.929)
70% Quantile	17.14*** (4.950)	22.36*** (6.536)	23.43*** (3.966)	-1.070 (7.605)
80% Quantile	23.28*** (6.424)	12.08* (8.259)	22.64*** (8.334)	-10.56 (11.70)
90% Quantile	16.51 (11.01)	22.38* (12.38)	10.03 (11.47)	12.35 (16.70)
KS statistic	11,231*** [0.002]	3,610 [0.508]	16,029*** [0.004]	12,419* [0.086]
PSD statistic	11,231*** [0.002]	3,610 [0.259]	16,029*** [0.001]	1,507 [0.811]
NSD statistic	0 [0.983]	0 [0.978]	0 [0.991]	-12,419** [0.033]

Note: *** indicates significance at the 1% level, ** indicates significance at the 5% level, and * indicates significance at the 10% level. Bootstrapped standard errors are in parentheses and are obtained from 1,999 bootstrap replications. "KS statistic" indicates the Kolmogorov–Smirnov statistic. "PSD statistic" indicates the Kolmogorov–Smirnov-type test statistic for positive stochastic dominance. "NSD statistic" indicates the Kolmogorov–Smirnov-type test statistic for negative stochastic dominance. Bootstrapped p-values (1,999 replications) are in squared brackets.

Beyond the 9th decile, the quantile earnings differences exceed 150 U.S. dollars. For females, the CQTEs are zero until the 2nd decile. At the 2nd decile, the quantile earnings differences are approximately 13 U.S. dollars and increase with rank to approximately 25 U.S. dollars. These results are consistent with the findings of Eren and Ozbeklik (2014). The Kolmogorov–Smirnov test results in Table 3 suggest that the Job Corps program has significant positive effects on the earnings distribution of males, but no significant effects on the earnings distribution of females. For males, the potential earnings distribution under treatment stochastically dominates the potential earnings distribution under non-treatment.

Figure 4 and column 4 of Table 3 report the differences between the CQTEs for females and

males. The negative differences at the 56th and 57th percentiles are significant at the 1% level. The negative differences at the 58th percentile are significant at the 5% level. The negative differences at the 18th, 46th, and 99th percentiles are significant at the 10% level. Further, the Kolmogorov–Smirnov statistics in Table 3 report the differences between the CQTEs for females and males, which are significant at the 10% level. The test for negative stochastic dominance suggests that the CQTEs for females are always equal to or smaller than the CQTEs for males (5% level). This suggests that the returns to the offer to participate in the Job Corps are larger for males than for females. Accordingly, the Job Corps offers increase gender earnings inequality.

6.2 TQTE and SQTE

Figure 5 reports the potential earnings distributions by gender. The earnings distributions for females are higher than the earnings distributions for males, regardless of whether measured under treatment or non-treatment. Accordingly, females earn less than males at all quantiles. Possibly the differences between the CQTEs for females and males can be explained by the structural differences between females’ and males’ earnings distributions.

Figure 6 reports the transformed relative ranks on the ordinate and the untransformed ranks on the abscissa. The relative ranks would lie on the 45° line in the absence of structural earnings inequality. The average rank difference between females and males is 12.1 percentage points (the minimum is 2.5 and the maximum is 18.4 percentage points). The relative rank of females is always lower than the 45° line. The relative rank of males is always lower than the 45° line. For example, males (females) at the 40% percentile in the conditional earnings distribution would have rank 48 (rank 34) in the population distribution. Accordingly, the relative distribution shifts the ranks of males upward and the ranks of females downward. This implies that TQTE shifts the CQTE for males to the right and that for females to the left.

Figure 7 and columns 1 and 2 of Table 4 document the findings for TQTEs by gender. The reference distribution is the earnings distribution of all eligible Job Corps candidates under non-treatment. This distribution is documented in Figure 2. I do not consider the parts of

Figure 5: Potential earnings distributions by gender.

Note: The treatment is a offer to join the Job Corps program. The outcome is weekly earnings (in U.S. dollars) 48 months after the randomized assignment. The figure is truncated by 600 U.S. dollars in earnings per week.

Figure 6: Rank transformation.

Note: I consider the relative distribution ranks only in the interval $[0.01, 0.99]$ and exclude all ranks below $\hat{F}_{Y(0)}(0)$ (for explanations, see Online Appendix A).

Figure 7: TQTEs of the Job Corps by gender on weekly earnings (in U.S. dollars) 48 months after randomized assignment.

Note: The lines report the point estimates separately calculated for all percentiles. Crosses on the lines indicate significant effects at the 10% level. Hollow circles on the lines indicate significant effects at the 5% level. Full diamonds on the lines indicate significant effects at the 1% level. I consider only the relative distribution ranks in the interval $[0.01, 0.99]$ and exclude all ranks below $\hat{F}_{Y(0)}(0)$ (for explanations, see Online Appendix A).

the earnings distribution with zero earnings for quantile transformation because the continuity assumptions – described in Section 5.2 – are violated at the mass point with zero earnings. This basically excludes all ranks lower than the 21st percentile. Further, I exclude the relative distributions ranks higher than the 99th percentile, which prevents estimating extreme quantiles.

Figure 7 shows patterns of TQTE that are more aligned between males and females. The TQTE for males is fairly close to zero when below the median. Above the median, the quantile earnings differences remain at approximately the 25 U.S.-dollar level. The large difference between the TQTE and CQTE for males at the 99th percentile decreases by approximately 30% (from 156 to 112 U.S. dollars). The CQTE and TQTE for females are overall quite similar. Below the median TQTE are approximately 15 U.S. dollars. The increase to 25 U.S. dollars occurs earlier for TQTE than for CQTE. The Kolmogorov–Smirnov-type test results at the bottom of Table 4 report that the Job Corps influences the potential earnings distribution

Table 4: TQTEs and SQTEs of the Job Corps by gender on weekly earnings (in U.S. dollars) 48 months after randomized assignment.

	TQTE		Difference between (1) and (2)	SQTE		Difference between (4) and (5)
	Females	Males		Females	Males	
	(1)	(2)	(3)	(4)	(5)	(6)
30% Quantile	14.83*	0.650	14.13	1.853	2.093	-0.240
	(8.499)	(8.289)	(12.107)	(4.431)	(4.024)	(6.208)
40% Quantile	14.18*	4.863	9.314	4.940	2.192	2.748
	(7.994)	(9.096)	(12.11)	(4.614)	(4.865)	(6.795)
50% Quantile	7.461	23.25***	-15.79	1.598	1.830	-0.232
	(6.829)	(7.846)	(10.92)	(6.060)	(6.628)	(7.034)
60% Quantile	19.13***	17.20***	1.928	-7.133	0.056	-7.189
	(7.391)	(5.886)	(8.863)	(5.299)	(4.425)	(6.587)
70% Quantile	13.45*	25.92***	-12.47	8.913*	-2.485	11.40
	(7.649)	(7.326)	(11.17)	(5.201)	(5.290)	(7.737)
80% Quantile	16.56	24.31***	-7.747	-4.479	-1.669	-2.810
	(10.30)	(7.734)	(13.11)	(6.720)	(4.268)	(7.741)
90% Quantile	20.30	14.79	5.501	2.082	-4.767	6.850
	(17.26)	(12.86)	(20.56)	(12.10)	(7.161)	(13.57)
KS statistic	2,659	11,490**	2,455	1,417	4,539	2,056
	[0.741]	[0.011]	[0.909]	[0.988]	[0.310]	[0.973]
PSD statistic	2,659	11,490***	1,610	1,417	4,539	1,258
	[0.378]	[0.010]	[0.727]	[0.844]	[0.112]	[0.982]
NSD statistic	430	67	-2,455	-1,361	-1,369	-2,056
	[0.993]	[0.993]	[0.551]	[0.833]	[0.904]	[0.743]

Note: *** indicates significance at the 1% level, ** indicates significance at the 5% level, and * indicates significance at the 10% level. Bootstrapped standard errors are in parentheses and are obtained from 1,999 bootstrap replications. "KS statistic" indicates the Kolmogorov–Smirnov statistic. "PSD statistic" indicates the Kolmogorov–Smirnov-type test statistic for positive stochastic dominance. "NSD statistic" indicates the Kolmogorov–Smirnov-type test statistic for negative stochastic dominance. Bootstrapped p-values (1,999 replications) are in squared brackets.

of males, but not that of females. These findings are consistent with the Kolmogorov–Smirnov-type tests applied to CQTE.

Figure 7 and column 3 of Table 4 report the quantile differences between the TQTEs for females and males. The differences oscillate around zero. The highly significant negative differences at the 56th and 57th percentiles, which are documented for CQTEs, decline for TQTEs by 45% (from approximately -22 to -12 U.S. dollars). For TQTEs, I do not find statistically significant differences between females and males. Table 3 and 4 also report standard errors. The standard errors of TQTE are not systematically higher than the standard errors of CQTE. Therefore, the insignificant findings are not caused by a more imprecise estimation of TQTE relative to CQTE. Furthermore, the bottom rows of Table 4 report the Kolmogorov–Smirnov-type

Figure 8: Difference between Job Corps TQTEs for females and males.

Note: The outcome is weekly earnings (in U.S. dollars) 48 months after randomized assignment. The lines report the point estimates separately calculated for all percentiles. All point estimates are not statistically different from zero. I consider the relative distribution ranks only in the interval $[0.01, 0.99]$ and exclude all ranks below $\hat{F}_{Y(0)}(0)$ (for explanations, see Online Appendix A).

statistics. I find no evidence of difference between the TQTEs for females and males. I also find no evidence that the TQTEs for females dominate the effects for males or *vice versa*. Overall, these results imply that the direct effects of the Job Corps are not strongly heterogeneous by gender.

Finally, Figure 9 reports the SQTEs by gender and Figure 10 reports the heterogeneity between SQTEs by gender. SQTE is the difference between CQTE and TQTE. Figure 9 reports the strongly significant SQTEs. For males, the CQTEs are larger than TQTEs between the 4th and 5th deciles. For females, the positive and negative differences are mixed. Overall, strong differences can be seen between CQTEs and TQTEs in Figure 9. The quantile differences between the SQTEs for males and females are significantly different from zero. Figure 10 reports the negative effects at the 46th and 58th percentiles, which are significant at the 1% level, and those at the 45th and 59th percentiles, which are significant at the 5% level. Further, positive effects are reported at the 71st and 74th percentiles, which are significant at the 10% level. Thus, the earnings structure affects CQTE heterogeneity by gender. However, the

Figure 9: SQTEs of the Job Corps by gender on weekly earnings (in U.S. dollars) 48 months after randomized assignment.

Note: The lines report the point estimates separately calculated for all percentiles. Crosses on the lines indicate significant effects at the 10% level. Hollow circles on the lines indicate significant effects at the 5% level. Full diamonds on the lines indicate significant effects at the 1% level. I consider the relative distribution ranks only in the interval $[0.01, 0.99]$ and exclude all ranks below $\hat{F}_{Y(0)}(0)$ (for explanations, see Online Appendix A).

Kolmogorov–Smirnov-type statistics at the bottom of Table 4 do not report significant effects.

In sum, the quantile earnings differences increase with rank. This favors males because males typically have higher earnings distribution ranks than females. Accordingly, the Job Corps program increases the existing inequality in structural earnings. After accounting for expected earnings opportunities, I find no strong evidence that the Job Corps is more aligned with the training needs of males than females.

6.3 Average effects

Individuals awarded offers to participate in the Job Corps have an average earnings gain of 15 U.S. dollars, or 8%. Females with no award earn 63 U.S. dollars or 29% less than males with no award of an offer. In the following discussion, I call this the gender earnings gap. The offer to participate in the Job Corps increases the earnings of females by 12 U.S. dollars and the earnings of males by 18 U.S. dollars. The average effect is approximately 5 U.S. dollars

Figure 10: Difference between SQTEs of the Job Corps for females and males.

Note: The outcome is weekly earnings (in U.S. dollars) 48 months after randomized assignment. The lines report the point estimates separately calculated for all percentiles. Crosses on the lines indicate significant effects at the 10% level. Hollow circles on the lines indicate significant effects at the 5% level. Full diamonds on the lines indicate significant effects at the 1% level. I consider the relative distribution ranks only in the interval $[0.01, 0.99]$ and exclude all ranks below $\hat{F}_{Y(0)}(0)$ (for explanations, see Online Appendix A).

lower for females than for males, corresponding to approximately 42% ($= 5/12 \cdot 100\%$) higher earnings gains for males than for females. These higher earnings increase the gender earnings gap by 8% ($= 5/63 \cdot 100\%$). Eren and Ozbeklik (2014) and Schochet, Burghardt, and McConnell (2008) report results of a similar magnitude. In the experimental survey data, the average effect heterogeneity is not significantly different from zero.

TATE accounts for a 1 U.S.-dollar earnings difference between females and males. SATE accounts for a 4 U.S.-dollar earnings difference between females and males. Accordingly, the structural earnings inequality accounts for 82% and the direct gender effects account for only 18% of the increase in the average gender earnings gap. Despite the insignificant average effects, a strong structural earnings inequality can drive the effect heterogeneity by gender. Job Corps trainability differences by gender do not seem to be an important channel of effect heterogeneity. If the offer to participate in the Job Corps is awarded to males and females in such a manner that it aligns the unconditional earnings distribution, then the average effects

Table 5: ATE, CATE, TATE, and SATE of the Job Corps on weekly earnings (in U.S. dollars) 48 months after randomized assignment.

Gender	Potential Outcome Levels		Average Effects	Average Differences between Females and Males
	Y(1)	Y(0)		
	(1)	(2)	(3)	(4)
ATE				
Both Gender	210.9	195.8	15.01*** (3.85)	-
CATE				
Females	170.4	158.1	12.31** (5.26)	-5.22 (7.55)
Males	239.2	221.6	17.54*** (5.32)	-
TATE				
Females	212.3	195.8	16.39** (6.98)	-0.95 (9.22)
Males	213.6	195.8	17.34*** (5.88)	-
SATE				
Females	-	-	-4.08* (2.26)	-4.28 (2.96)
Males	-	-	0.19 (1.70)	-

Note: *** indicates significance at the 1% level, ** indicates significance at the 5% level, and * indicates significance at the 10% level. Bootstrapped standard errors are in parentheses. They are obtained from 499 bootstrap replications.

are larger for both genders (average TATE is larger than average CATE). At the same time, gender earnings inequality increases by only 2% ($= 1/63 \cdot 100\%$). Accordingly, the proposed selection rule appears efficient in terms of larger average earnings gains and smaller increases in gender earnings gap.

6.4 Alternative reference and relative distributions

Online Appendix B.1 reports TQTE and SQTE when the reference quantiles are from the potential outcome distribution under treatment ($Q_{Y(1)}(\tau)$), the observed outcome distribution ($Q_Y(\tau)$), the potential outcome distribution of males under non-treatment ($Q_{Y(0)|G}(\tau|m)$), and the potential outcome distribution of females under non-treatment ($Q_{Y(0)|G}(\tau|f)$). The results do not change qualitatively when a different reference distribution is used (see Figures B.1.1.1-B.1.4.4).

Online Appendix B.2 reports TQTE and SQTE when the relative distribution is defined under treatment $R(\tau, g) = F_{Y(1)|G}(Q_{Y\tau}(\tau)|g)$. Again, I attempt many possible reference distributions (see Figures B.2.1.1–B.2.5.4). The results do not change qualitatively when the relative distribution is defined under treatment. Accordingly, the results are not sensitive to the choice of the reference and relative distributions.

7 Conclusions

The literature finds that the Job Corps increases gender earnings inequality. These findings are surprising because most other ALMP evaluations show more positive effects for females than for males (see, e.g., Card, Kluve, and Weber, 2015). I estimate the TQTEs, which were initially proposed by Bitler, Hoynes, and Domina (2014), to reveal the potential mechanism behind these unexpected findings. Frumento, Mealli, Pacini, and Rubin (2012) identify labor market opportunities as the main reason for effect heterogeneity. I approximate labor market opportunities using expected earnings. Because the earnings distributions are unequal by gender, the existing structural gender earnings inequality can be a mechanism for effect heterogeneity by gender.

My results suggest that the award of an offer to participate in the Job Corps increases the average weekly earnings of eligible individuals by approximately 15 U.S. dollars. The absolute earnings gains are approximately 5 U.S. dollars lower for females than for males. Therefore, the random award of offers to participate in the Job Corps increases the gender earnings gap by approximately 8%. On average, structural earnings inequality accounts for 82% of the effect heterogeneity by gender, whereas only 18% of the effect heterogeneity can be explained by Job Corps trainability differences by gender. A selection rule for awarding Job Corps offers that balances the earnings opportunities of awardees by gender does not reduce average returns and has only little impact on the gender earnings gap.

References

- ABADIE, A., J. D. ANGRIST, AND G. W. IMBENS (2002): “Instrumental Variables Estimates of the Effect of Subsidized Training on the Quantiles of Trainee Earnings,” *Econometrica*, 70(1), 91–117.
- ATHEY, S., AND G. W. IMBENS (2006): “Identification and Inference in Nonlinear Difference-In-Difference Models,” *Econometrica*, 74(2), 431–497.
- BARGAIN, O., K. ORSINI, AND A. PEICHL (2013): “Comparing Labor Supply Elasticities in Europe and the US: New Results,” *Journal of Human Resources*, 49(3), 723–838.
- BERGEMANN, A., AND G. J. VAN DEN BERG (2008): “Active Labor Market Policy Effects for Women in Europe – A Survey,” *Annales d’Économie et de Statistique*, 91/92, 385–408.
- BITLER, M. P., J. B. GELBACH, AND H. W. HOYNES (2008): “Distributional impacts of the Self-Sufficiency Project,” *Journal of Public Economics*, 92(3-4), 748–765, (see also NBER Working Paper 11626).
- (2014): “Can Variation in Subgroups’ Average Treatment Effects Explain Treatment Effect Heterogeneity? Evidence from a Social Experiment,” *Review of Economics and Statistics*, forthcoming.
- BITLER, M. P., H. W. HOYNES, AND T. DOMINA (2014): “Experimental Evidence on Distributional Effects of Head Start,” *NBER Working Paper*, 20434.
- BLANCO, G., C. A. FLORES, AND A. FLORES-LAGUNES (2013): “Bounds on the Quantile Treatment Effects of Job Corps on Wages,” *Journal of Human Resources*, 48(3), 659–701.
- BLINDER, A. (1973): “Wage Discrimination: Reduced Form and Structural Estimates,” *Journal of Human Resources*, 8, 436–455.
- BURGHARDT, J., P. Z. SCHOCHET, S. MCCONNELL, T. JOHNSON, R. M. GRITZ, S. GLAZERMAN, J. HOMRIGHAUSEN, AND R. JACKSON (2001): “Does Job Corps Work? Summary of the National Job Corps Study,” *Princeton, NJ, Mathematica Policy Research, Inc.*

- CARD, D., J. KLUVE, AND A. WEBER (2015): “What Works? A Meta Analysis of Recent Active Labor Market Program Evaluation,” *NBER Working Paper*, 21431.
- CHAISEMARTIN, C., AND X. D’HAULTFEUILLE (2015): “Fuzzy Differences-in-Differences,” *Warwick Economics Research Paper Series*, 1065.
- CHEN, X., AND C. A. FLORES (2015): “Bounds on Treatment Effects in the Presence of Sample Selection and Noncompliance: The Wage Effects of Job Corps,” *Journal of Business and Economic Statistics*, 33(4), 523–540.
- CHEN, X., C. A. FLORES, AND A. FLORES-LAGUNES (2015): “Going Beyond LATE: Bounding Average Treatment Effects of Job Corps Training,” *IZA Discussion Paper*, 9511.
- CHERNOZHUKOV, V., AND I. FERNÁNDEZ-VAL (2005): “Subsampling Inference on Quantile Regression,” *Sankhya: The Indian Journal of Statistics*, 67(2), 253–276.
- CHERNOZHUKOV, V., I. FERNÁNDEZ-VAL, AND B. MELLY (2013): “Inference on Counterfactual Distributions,” *Econometrica*, 81(6), 2205–2268.
- CHERNOZHUKOV, V., AND C. HANSEN (2005): “An IV Model of Quantile Treatment Effects,” *Econometrica*, 73(1), 245–261.
- EREN, O., AND S. OZBEKLIK (2014): “Who Benefits from Job Corps? A Distributional Analysis of an Active Labor Market Program,” *Journal of Applied Econometrics*, 29(4), 586–611.
- FIRPO, S. (2007): “Efficient Semiparametric Estimation of Quantile Treatment Effects,” *Econometrica*, 75(1), 259–276.
- FORTIN, N. M., T. LEMIEUX, AND S. FIRPO (2010): “Decomposition Methods in Economics,” in *Handbook of Labor Economics, Vol. 4 Part A*, ed. by O. Ashenfelter, and D. Card, pp. 1–102. North Holland.
- FRANDSEN, B. R., AND L. J. LEFGREN (2015): “A Test for Rank Similarity and Partial Identification of the Distribution of Treatment Effects,” *Working Paper*.

- FRUMENTO, F., F. MEALLI, B. PACINI, AND D. RUBIN (2012): “Evaluating the Effect of Training on Wages in the Presence of Noncompliance and Missing Outcome Data,” *Journal of the American Statistical Association*, 107(498), 450–466.
- GAY, R. S., AND M. E. BORUS (1980): “Validating Performance Indicators for Employment and Training Programs,” *Journal of Human Resources*, 15(1), 29–48.
- HECKMAN, J. J., J. A. SMITH, AND N. CLEMENTS (1997): “Making the Most out of Programme Evaluations and Social Experiments: Accounting for Heterogeneity in Programme Impacts,” *Review of Economic Studies*, 64(4), 487–535.
- JOB CORPS ANNUAL REPORT (2008): *US Department of Labor Publications, Washington, DC*.
- JUHN, C., K. M. MURPHY, AND B. PIERCE (1991): “Accounting for the Slowdown in Black-White Wage Convergence,” in *Workers and their Wages: Changing Patterns in the United States*, ed. by M. H. Koster, pp. 107–143. American Enterprise Institute, Washington.
- KILLINGSWORTH, M., AND J. HECKMAN (1986): “Female labor supply: A survey,” in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and R. Layard, pp. 3144–3259. North-Holland, Amsterdam.
- KOENKER, R., AND G. BASSETT (1978): “Regression Quantiles,” *Econometrica*, 46(1), 33–50.
- LEE, D. S. (2009): “Training, Wages, and Sample Selection: Estimate Sharp Bounds on Treatment Effects,” *Review of Economic Studies*, 76(3), 1071–1102.
- MACHADO, J. A. F., AND J. MATA (2005): “Counterfactual Decomposition of Changes in Wage Distributions using Quantile Regression,” *Journal of Applied Econometrics*, 20(4), 445–465.
- MALLAR, C., S. KERACHSKY, C. THORNTON, AND D. LONG (1982): “Evaluation of the Economic Impact of the Job Corps Program,” *Princeton, NJ, Mathematica Policy Research, Inc.*

- MELLY, B., AND G. SANTANGELO (2015): “The Changes-in-Changes Model with Covariates,” *mimeo*.
- MORRIS, M., A. D. BERNHARDT, AND M. S. HANDCOCK (1994): “Economic Inequality: New Methods for New Trends,” *American Sociological Review*, 59(2), 205–219.
- OAXACA, R. (1973): “Male-Female Wage Differentials in Urban Labour Markets,” *International Economic Review*, 14(3), 693–709.
- POWELL, D. (2016): “Quantile Treatment Effects in the Presence of Covariates,” *Working Paper*.
- RINNE, U., M. SCHNEIDER, AND A. UHLENDORFF (2011): “To Bad to Benefit? Effect Heterogeneity of Public Training Programs,” *Applied Economics*, 43(25), 3465–3494.
- ROSENBAUM, P. R., AND D. B. RUBIN (1983): “The Central Role of Propensity Score in Observational Studies for Causal Effects,” *Biometrika*, 70(1), 41–55.
- RUBIN, D. B. (1974): “Estimating the Causal Effect of Treatments in Randomized and Non-Randomized Studies,” *Journal of Educational Psychology*, 66(5), 688–701.
- SCHOCHET, P. Z. (2001): “National Job Corps Study: Methodological Appendixes on the Impact Analysis,” *Princeton, NJ, Mathematica Policy Research, Inc.*
- SCHOCHET, P. Z., J. BURGHARDT, AND S. MCCONNELL (2008): “Does Job Corps Work? Impact Findings from the National Job Corps Study,” *American Economic Review*, 98(5), 1864–1886.