INE PAN Working Paper Series

Paper number 51

Well-being gaps: theory

Martyna Kobus, Radosław Kurek

Warsaw, 28.02.2022

wp@inepan.waw.pl

Well-being gaps: theory

Martyna Kobus[*], Radosław Kurek[†]

Abstract

In this paper we extend the well-known problem of estimating counterfactual effects to the case when there is not a single outcome, but possibly many outcomes (e.g. income and health). We call this well-being gaps, in line with e.g. pay or income gaps. This extension requires new definitions of well-known counterfactual effects, one that take into account the multidimensional nature of outcomes. Also, in a multidimensional setting new counterfactual effects are possible. We provide decomposition analysis in the spirit of Oaxaca-Blinder for these newly defined effects. We study the role of dependence (i.e. substitution) between outcomes in the overall counterfactual effect.

Keywords: gender gaps; well-being; counterfactual distribution, decomposition analysis
JEL classification: D30; I31; C02

1 Introduction

The interest in the gender gap, or racial gap, and similar questions of inequality derives from an interest in well-being of target populations, generally, and allocation and distributive questions generally. The choice of wages, or "income" as the basis for measurement of inequality, poverty, mobility, and the gaps thereof, reflects convenience and primacy of income as one indicator of well-being. The health status, educational attainments, or social amenities provided by the communities in which families and individuals reside, determine their well-being. The latter is a latent concept, like happiness, and the challenges in its measurement have been acknowledged for many years. Recent Nobel prize in economics was awarded to Angus Deaton for his efforts in these regards, as was the case with Amartya Sen's. The multivariate analysis of wellbeing gaps is diminished when it is conducted in one dimension, any one dimension, since the same dollar level of wages has vastly different values for a healthy or unhealthy individual, or one that has the investment and consumption benefits of high levels of educational attainment, living in a society that offers strong social nets. Multivariate distributions of attributes need to be contrasted for a more informed and realistic view of allocations and well-being.

[*]Institute of Economics, Polish Academy of Sciences, 00-330 Warsaw, Nowy Swiat 72, Poland, mkobus@inepan.waw.pl.

[†]Institute of Economics PAS, radek.kurek@inepan.waw.pl.

This poses additional challenges to the ones increasingly being addressed in one dimension, income/wages, in the literature on counterfactual distributions, quantile effects, and statistical inference therein. This new challenge comes from the dependence between dimensions of well-being and the degree of their substitution. This needs to be taken into account in how one defines the counterfactual effect. Furthermore, there is an inherently challenging aggregation problem in a multidimensional setting, reflected in the choice of summary measures. If substantial heterogeneity is present, then there is great arbitrariness in any gap measure. For example, with quantile effects, one typically has to assume rank invariance or similarity, which are often unsupported empirically. This is an assumption that women endowed counterfactually with men's skill set will occupy the same ranks in counterfactual and observed setting. Such quantile effects thus rule out the possibility that when endowed with each other's skill composition or market returns, men and women may engage differently in skills substitutions and other decisions. This confirms the reservations of Heckman et al. (1997) as to the meaningfulness of such gaps.

Indeed, typically used summary measures imply aggregation over individuals and assume homogeneous evaluation functions across individuals and even groups, whereas in the presence of heterogeneity, different weights and substitution values apply to different groups. Comparisons between these groups or evaluation of transfers between them requires knowledge of how different groups trade off well-being attributes. This is a policy relevant question, for example, thinking about optimal investment choices with respect to the change in relative "prices" of the attributes, if estimated elasticities of substitution differ between welfare levels (implicit attribute price ratios differ), then the same change in "prices" will lead to different optimizing choices at different levels of welfare. At levels at which the market price ratio differs from implicit price ratio, there is room for welfare improvement. Strong homogeneity assumptions which are embedded in most summary measures for entire populations or groups are unrealistic in many contexts and in particular inappropriate for counterfactual exercises which aim at comparing similar agents.

The aim of this paper is to propose a theory for identification and estimation of well-being gaps. These can be gaps between between men and women, or any exogenuously defined groups. This leads to some new definitions of counterfactual effects which take into account the multidimensionality of quantiles and the role of attributes' dependence in counterfactual analysis. Some developments in multidimensional aggregation in well-being inequality analysis offer potential directions which we hope to introduce to the promising paradigm of counterfactual analysis of distributions that should enhance our understanding of many new issues, and standards for defining counterfactual experiments and meanings. We think of a counterfactual distribution as either resulting from a change in the distribution of covariates (keeping women's wage structure fixed but giving them men's characteristics), or of a change in the relationship between covariates and outcome, that is, a change in the conditional distribution of outcome given a set of characteristics (keeping women's characteristics fixed but giving them men's wage structure). Here it is possible to define a new counterfactual effect such that we give women access to marginal conditional distribution of men's outcomes, but we keep a copula from the distribution of women. Informally speaking, while we provide women with men's outcome, we let them keep their preferences (substitution) between the two outcomes. We address heterogeneity explicitly by using Maasoumi and Racine (2016) parametrization which recovers substitution weights. However, here it is worth noting that in our case quantiles are sets. Parametrization helps to avoid problems related to defining measures of distance between two sets. Again, with parametrization, new counterfactual effects are possible. Firstly, it is the difference in well-being between women's real and counterfactual distribution, as measured by the parametrization function. Secondly, we keep the substitution parameters in the parametrization fixed, thus allowing women to keep their preferences in the counterfactual scenario. As to traditional quantile gaps, it is worth noting that in our case quantiles are sets. Finally, we ask a question of the role of dependence between outcome in the magnitude of the counterfactual effect. This question is difficult to answer in full generality, however, we offer ways to solve this issue. This paper is organized as follows. In Section 2 we review the related literature and place the current paper in the context of this literature. In Section 3 we develop a theory of well-being gaps and state some results for the case of two groups (e.g. gender well-being gap). In Section 4 we extend the notation of the previous Section to the case of an arbitrary number of groups. Finally, we conclude (Section **??**).

2 Relation to the literature

This paper builds a bridge between the literature on well-being measurement, which now widely adopts the view that well-being is inherently a multidimensional concept (Stiglitz et al. 2009) and the econometrics literature on group inequalities. In more detail, the paper relates to four big areas of economics research, namely, empirical evidence on gender gaps (Blau and Kahn 1997), wellbeing research (Deaton 2015), estimation of counterfactual distributions and policy effects (Chernozhukov et al. 2013) and decomposition methods (Fortin et al. 2011). The paper results will directly expand the economics literature in these areas.

The distribution of wages in the US has evolved differently for men and women over several decades. Using conventional gap measures as the mean and quantiles, economists have shown that the gender gap has decreased over time, especially in the 1980s and early 1990s. This trend has slowed down since the mid-1990s. Women are catching up with men (Blau and Kahn 1997, Goldin 2014). The quantile gaps, however, have evolved differently over the past decades (Albrecht et al. 2003). The measurement of the gap is filled with problems of heterogeneity and selection. In the presence of significant heterogeneity, all gap measures are biased. Rank invariance assumption, namely, that agents occupy a given quantile in both factual and counterfactual setting, does not hold. Men and women, when endowed with each other's skill set or market returns, may engage differently in skills substitutions, work and job decisions. Gap measures are biased also when there is selection into employment. That is, if non-working men and women systematically differ from working men and women, measures of the gap would be biased. For example, if only high-earning women enter labor market we may observe convergence in the gap with no actual progress. There are only few attempts in the gender gap literature that account for selection (Blau and Kahn 2006, Mulligan and Rubinstein 2008). Recently, Maasoumi and Wang (2017) examine the gap using General Entropy measures which overcomes the problem of heterogeneity. These are anonymous aggregative measures of the gap. They also use the novel copula approach developed by Arellano and Bonhomme (2017) to model selection. Once selection is accounted for, the gap converges slower than without selection and in fact the trend reverses in some parts of the distribution between mid-1990s and the most recent recession. In the great recession, there was a marked decline in the gap among low-skilled workers, perhaps due to a relative deterioration in wages of the low-skilled males. Convergence is much smaller amongst the least educated or black women, especially during the recent years.

Economists have for long acknowledged that improvements in well-being are not confined to economic growth (Sen 1973). Governments (for example in UK, France, Canada, Japan) and international organizations have been increasingly responding to economists' calls to go beyond GDP in measuring nations' progress. Measuring well-being is now a prominent item on the agenda of many statistical offices. There is a wide spread recognition that well-being measurement is critical for informing policy making about areas that matter to people. Specific measures taken in the last decade include the construction of the OECD's Better Life Index, the report by Stiglitz-Sen-Fitoussi commission on the measurement of economic performance and social progress, the program initiated by the British Office for National Statistics on measuring nation's well-being or the 2011 resolution of the United Nations General Assembly (No. 65/309) which explicitly "invites Member States to purse elaboration of additional measures that better capture the importance of the pursuit of happiness and well-being" than GDP-based indicators.

It thus seems natural to think about well-being differences and not just wage differences when comparing groups (men and women, black and white, and other groups). This justifies the types of problems tackled in this paper. This broad view of the gap requires developing the theory and estimation procedures for measuring it. This problem has not been considered so far in econometrics literature, which constitutes the pioneering nature of the paper.

Since gaps considered in the paper are multidimensional we briefly describe the state of the art and type of problems present in multidimensional aggregation. The key problem in multi-dimensional settings is the problem of aggregation. Since the problem of finding credible weights is difficult and thus complete rankings of welfare states are subject to arbitrariness in the functional form of the aggregator function and assumption of strong homogeneity, the researchers have sought to overcome this problem by focusing on partial rankings i.e. through stochastic dominance criteria. This approach dates back to classic work of Atkinson and Bourguignon (1982) who develop conditions to compare distributions uniformly in a way that is consistent with classes of welfare aggregators. These are, however, restrictive. The problem of heterogeneity is significant i.e. groups with different needs cannot be represented with identical utility functions. Allowing for transfers between groups demands a clarification of how different groups value one one attribute over the other, for example, income over health. If certain plausible assumptions can be made about the trade-offs between incomes and other variables, such as health or other "needs" and if groups can be ranked with respect to "needs", then partial comparability would obtain (Atkinson and Bourguignon 1987). For example, if one were to follow Sen's equity axiom which states that groups can be ranked by marginal valuation of income, namely, the least healthy groups has the highest marginal valuation of income, the nest healthiest groups has the second highest marginal valuation of income, and so on, then one can rank welfare states using appropriate dominance conditions. However, Maasoumi and Racine (2016) example based on Indonesian data suggests that heterogeneity is so high that the above assumptions seem implausible to be true across the board. Consequently, they do not find many findings of statistically significant multidimensional rankings. In other words, whenever heterogeneous populations are involved interpersonal comparisons of well-being are inevitable, and if heterogeneity is significant this may lead to impossibility of rankings. This has major implications for all summary measures of inequality, poverty, or equivalence scales, which all make assumptions about relatively homogeneous evaluation functions across individuals and even groups. Robust analysis may have to avoid summary measures as the assumptions they employ seem to be empirically unsupported. Maasoumi and Racine (2016) offer a way to recover weights which stems from the following insight. Non-parametrically estimated joint distribution contains information in the purest form, free from arbitrary assumptions. Non-parametrically estimated quantile sets are iso-quants of well-being aggregators i.e. they represent individuals with the same level of well-being. Well-being here is a latent concept, but the authors argue that there are strong parametric candidates for well-being functions, namely, generalized entropy measures that minimize the divergence between the joint distribution and the distribution of its constituent dimensions. Mapping the non-parametric joint distribution, possibly conditional on other characteristics, to the chosen aggregator functions provides estimates for weights and substitution parameters. Such index based evaluation function provide then a complete ranking of welfare states.

3 Theory of well-being gaps for two populations

3.1 Benchmark model

Let us suppose that we have two populations, of men (denoted 0) and women (denoted 1). For each population $k = 0, 1$, there is a random d_x-vector X_k of covariates and a random outcome d_y-vector Y_k, where for now we assume $d_y = 2$. We denote the support of X_k by $\mathcal{X}_k \subseteq \mathbb{R}^{d_x}$ and denote the region of interest for outcomes Y_k by $\mathcal{Y}_k \subseteq \mathbb{R}^2$. The conditional distribution functions $F_{Y_0|X_0}(y|x)$ and $F_{Y_1|X_1}(y|x)$ describe the stochastic assignment of wages to workers with characteristics x, for men and women, respectively. Let $F_{<0|0>}$ and $F_{<1|1>}$ represent the observed distribution function of wages for men and women, and let $F_{<0|1>}$ represent the counterfactual distribution of wages that would have prevailed for women had they faced the men's wage schedule $F_{Y_0|X_0}$:

$$F_{<0|1>}(y) := \int_{\mathcal{X}_1} F_{Y_0|X_0}(y|x)\mathrm{d}F_{X_1}(x),\ y \in \mathcal{Y}_0 \tag{1}$$

Such integral is well-defined if the support condition holds if $\mathcal{X}_0$, the support of characteristics for men, includes $\mathcal{X}_1$, the support of women's charactertistics, namely, $\mathcal{X}_1 \subseteq \mathcal{X}_0$. This distribution is called counterfactual because it does not arise as a distribution from any observable population. Rather, it is constructed by integrating the conditional distribution of wages for men with respect to the distribution of characteristics for women.

The difference in the observed wage distributions between men and women can be decomposed in the spirit of Oaxaca (1973) and Blinder (1973) as

$$F_{<1|1>} - F_{<0|0>} = \underbrace{\left(F_{<1|1>} - F_{<0|1>}\right)}_{\text{Structural effect}} + \underbrace{\left(F_{<0|1>} - F_{<0|0>}\right)}_{\text{Composition effect}}. \tag{2}$$

These are called counterfactual effects (CE). The first term in brackets is due to differences in the wage structure and the second term is a composition effect due to differences in characteristics. We can decompose similarly any functional of the observed wage distributions such as the quantile function or Lorenz curve into wage structure and composition effects. These counterfactual effects are well defined econometric parameters and are widely used in empirical analysis, for example, the first term of the decomposition is a measure of gender discrimination. It is important to note that these effects do not necessarily have a causal interpretation without additional conditions.

3.2 Counterfactual effects: further definitions

We may as well define counterfactual effects in terms of related quantiles. With distribution treatment effects, we say that the probability that women earn x income is higher by $\mathrm{d}\tau$ had they had the distribution of men's characteristics. With quantile treatment effects, we say that with probability τ women earn x more income had they had the distribution of men's characteristics. Let

$$Q_{<0|1>}(\tau) = F^{-1}_{Y_0|X_1}(\tau) = \{y \in \overline{\mathcal{Y}_0} : F_{Y_0|X_1}(y) < \tau\},\ \tau \in (0,1) \tag{3}$$

knowing that conditional quantile models imply conditional distribution models through the relation

$$F_{Y_j|X_j} = \int_{(0,1)} 1_{\{Q_{Y_j|X_j}(u|x) \leq y\}} \mathrm{d}u,\ j \in \{0,1\} \tag{4}$$

The counterfactual effect is then $Q_{<1|1>}(\tau) - Q_{<0|1>}(\tau)$, but given that in $\mathbb{R}^2$ quantiles are not numbers but sets, this difference is not very meaningful. This is one of the reasons why we resort to parametrisation of quantile sets in further sections.

Now we come back to distribution effects, because in the multidimensional framework we can have new counterfactual effects based on the distribution. Using Sklar's Theorem, we write

$$F_{<0|1>}(y) = C_{<0|1>}(F_{Y^1<0|1>}(y^1), F_{Y^2<0|1>}(y^2)) \tag{5}$$

We can now re-define the counterfactual effect (2). We start by re-defining the first counterfactual distribution

$$\tilde{F}_{<0|1>}(y) := \int_{\mathcal{X}_1} C_{Y_1|X_1}(F_{Y_0^1|X_0}(y^1|x), F_{Y_0^2|X_0}(y^2|x)|x)\mathrm{d}F_{X_1}(x),\ y \in \mathcal{Y}_0, \tag{6}$$

and the second counterfactual distribution

$$F_{<0|1>}(y) := \int_{\mathcal{X}_1} C_{Y_0|X_0}(F_{Y_0^1|X_0}(y^1|x), F_{Y_0^2|X_0}(y^2|x)|x)\mathrm{d}F_{X_1}(x),\ y \in \mathcal{Y}_0. \tag{7}$$

Let us now focus on the structural counterfactual effect $F_{<1|1>} - F_{<0|1>}$. Given (6) it can be further decomposed into

$$F_{<1|1>} - F_{<0|1>} = \left(F_{<1|1>} - \tilde{F}_{<0|1>}\right) + \left(\tilde{F}_{<0|1>} - F_{<0|1>}\right) \tag{8}$$

In $(F_{<1|1>} - \tilde{F}_{<0|1>})$ from the distribution of women we subtract a distribution where the characteristics X and preferences (copula) are that of women, but marginal distributions are that of men. That is, we give women access to men's conditional distribution except that we keep as fixed their preferences towards the substitution of two outcomes. On the other hand, in $(\tilde{F}_{<0|1>} - F_{<0|1>})$ we subtract two counterfactual distributions that differ by a copula and marginal distribution of outcomes, but have the same women's characteristics. Comparing to the previous effect, it is as if we exchanged the preferences of women towards two outcomes for men's preferences.

3.3 Multidimensional vs. marginal counterfactual effects

Given the multidimensionality of the problem, we are interested in how the main feature of multidimensionality, namely dependence, impacts on counterfactual effects. We have the following result.

Lemma 1. $F_{<0|1>1} = F_{Y_1<0|1>}$ *and* $F_{<0|1>2} = F_{Y_2<0|1>}$

Proof.

$$
\begin{aligned}
F_{<0|1>1}(y_1) = \int_{\mathrm{supp}Y_{<0|0>2}} F_{<0|1>}(y_1, y_2)\mathrm{d}y_2 = \int_{\mathrm{supp}Y_{<0|0>2}} \int_{\mathrm{supp}X_1} F_{Y_0|X_0}(y_1, y_2|x)\mathrm{d}x\mathrm{d}y_2 = \\
= \int_{\mathrm{supp}X_1} \int_{\mathrm{supp}Y_{<0|0>2}} F_{Y_0|X_0}(y_1, y_2|x)\mathrm{d}y_2\mathrm{d}x = \int_{\mathrm{supp}X_1} F_{Y_{01}|X_0}(y_1|x)\mathrm{d}x = \\
= F_{Y_1<0|1>}
\end{aligned}
$$

□

Lemma 1 states that the cumulative counterfactual distribution of Y_1 (and of Y_2) is the marginal cdf of the joint counterfactual distribution of Y_1, Y_2. That is, the construction of counterfactual marginal distribution can be recovered from counterfactual joint distribution by taking its marginals.

Therefore, we can bound the joint counterfactual effect.

Lemma 2.

$$
\max(F_{<1|1>1} + F_{<1|1>2} - 1, 0) - \min(F_{<0|1>1}, F_{<0|1>2}) \leq \Delta^1 = F_{<1|1>}(y) - F_{<0|1>}(y) \leq
$$

$$
\min(F_{<1|1>1}, F_{<1|1>2}) - \max(F_{<0|1>1} + F_{<0|1>2} - 1, 0)
$$

Utilizing Fréchet-Hoeffding bounds for copulas (Nelsen, 2006) we construct a lower and upper bound for the counterfactual effect in the joint distribution depending on its marginal distributions. In the lower bound we take joint distribution of $F_{<1|1>}$ with the highest possible dependence and subtract the joint distribution of $F_{<0|1>}$ with the lowest possible dependence; the reverse holds for the upper bound.

So we know that the overall counterfactual effect is somewhere between these two bounds, but how is it affected by changing dependence? Or, how much of this overall counterfactual effect is due to differences in dependence of two outcomes (e.g. due to differences in income-health gradient) between men and women? While these are very interesting questions and they appear naturally in a multidimensional framework as they point exactly to a distinctive feature of multidimensionality which is dependence, they are also very difficult to answer in full generality. We can say the following.

Lemma 3. *Constructing counterfactual distribution for* (Y_1, Y_2) *preserves the sign of correlation between* Y_1 *and* Y_2.

Lemma 3 states that in (7) the copula the direction of dependence between Y_1 and Y_2 in $C_{Y_0|X_0}$ is preserved by the copula of $F_{<0|1>}$. However, how do these two copulas differ (or maybe they are the same) – this is the question for further research. We suspect, however, that to answer quantitatively the question of what is the impact of dependence in the joint counterfactual effect is probably impossible in full generality and requires a choice of the particular family of copula functions. Another solution to this problem is to use a modified definition of the joint counterfactual effect, namely,

$$\Delta^2 = (F_{<1|1>} - F_{<1|1>1}F_{<1|1>2}) - (F_{<0|1>} - F_{<0|1>1}F_{<0|1>2}), \tag{9}$$

that is, distributions are compared to the independent distribution. Thus, we eliminate the effect of whatever is independent, and leave out only the difference in dependence. We can write

$$\Delta^2 = (F_{<1|1>} - F_{<0|1>}) - (F_{<1|1>1}F_{<1|1>2} - F_{<0|1>1}F_{<0|1>2}),$$

where the second difference is a comparison of two distributions as if they had copulas for independent distribution. Therefore, even if the issue of quantitative importance of dependence in counterfactual effects cannot be answered explicitly, we can try to solve it from a different angle by using alternative definitions that shed light on this issue as well.

Another approach that would help to make quantitative statements is to para-metrize quantile sets. This is the approach of Maasoumi and Racine (2016) which we follow.

3.4 Parametrization of quantile sets

Since we are interested in a multidimensional analysis of well-being, the obtained quantiles are sets e.g. for $d_y = 2$ this is the locus of (y_1^k, y_2^k) for which the distribution function admits a particular value from $(0, 1)$, as in an iso-quant. The fundamental problem is then to obtain credible weights of the aggregation functions that represent a latent concept such as well-being. We follow Maasoumi and Racine (2016) who fit a desired parametric form to non-parametric estimates of such iso-quants and discover the unknown substitution parameters. Multidimensional quantiles are equi-probable surfaces that correspond to the values of aggregation functions, parameters of which we seek to discover. Their proposed aggregation functions were first derived in Maasoumi (1986) as ideal aggregators in the sense of minimising the divergence between its distribution and the distribution of its constituent components. Without loss of generality, let $d_y = 2$. The CES aggregation function is given by

$$S(y_1, y_2) = A(\alpha y_1^{-\beta} + (1-\alpha) y_2^{-\beta})^{\frac{-1}{\beta}}$$

where $A > 0$ and $0 < \alpha < 1$. The partial derivatives are as follows,

$$S_{y_1} = \frac{\partial S(y_1, y_2)}{\partial y_1} = A\alpha(\alpha y_1^{-\beta} + (1-\alpha) y_2^{-\beta})^{(\frac{-1}{\beta})-1} y_1^{-\beta-1}$$

$$S_{y_2} = \frac{\partial S(y_1,y_2)}{\partial y_2} = A(1-\alpha)(\alpha y_1^{-\beta} + (1-\alpha)y_2^{-\beta})^{(\frac{-1}{\beta})-1} y_2^{-\beta-1}$$

Along an iso-well-being quantile we have $\Delta S = 0$ (i.e. $S_{y_1}\partial y_1 + S_{y_2}\partial y_2 = 0$), therefore

$$-\frac{S_{y_1}}{S_{y_2}} = \frac{\partial y_2}{\partial y_1} = \frac{\alpha}{1-\alpha}\left(\frac{y_2}{y_1}\right)^{\beta+1} \tag{10}$$

We exploit the fact that, for $y = (y_1, y_2)$, conditional on x, we can obtain estimates of $\frac{\partial y_2}{\partial y_1}$ directly from the estimated quantile $\widehat{Q}_{Y_j|X_k}(\tau)$ (i.e. for a given value of τ we can compute $\frac{\partial y_2}{\partial y_1}$ since the level of multidimensional well-being is constant). The estimates of α and β are then obtained via (nonlinear) regression of our non-parametrically estimated $\frac{\partial y_2}{\partial y_1}$ on $\frac{y_2}{y_1}$ using (10). The values of α and β may vary with the quantile. This means that different dimensions of well-being contribute differently to well-being at each level of well-being. This accounts for heterogeneity in preferences. With three dimensions of well-being we add conditions on higher order derivatives, however this becomes uninterpretable with more than three dimensions. Then, natural solution appears to be using simultaneous quantile equations models.

Coming back to counterfactual effects, the S parametrization is related to the distribution via quantile sets

$$Q_{<0|1>}(\tau) = \{y \in \mathbb{R}^2 : F_{<0|1>}(y) = \tau\}$$

$$Q_{<1|1>}(\tau) = \{y' \in \mathbb{R}^2 : F_{<1|1>}(y') = \tau\}.$$

Thus, we use the following notation: $y = \left(y_1^{<0|1>}(\tau), y_2^{<0|1>}(\tau)\right)$ and $y' = \left(y_1^{<1|1>}(\tau), y_2^{<1|1>}(\tau)\right)$. For the S function we put $S^{<0|1>}\left(y_1^{<0|1>}(\tau), y_2^{<0|1>}(\tau)\right)$ and $S^{<1|1>}\left(y_1^{<1|1>}(\tau), y_2^{<1|1>}(\tau)\right)$. We can now define the counterfactual effect using S parametrization.

$$\Delta^3 = S^{<1|1>}\left(y_1^{<1|1>}(\tau), y_2^{<1|1>}(\tau)\right) - S^{<0|1>}\left(y_1^{<0|1>}(\tau), y_2^{<0|1>}(\tau)\right). \tag{11}$$

That is, at a given point τ, (11) is the effect of the difference in wellbeing (as measured by S) between women and hypothetical women had they have men's conditional distribution of wellbeing dimensions. Please note that this difference is a difference of two numbers, because both S functions are defined exactly so that they are constant along the isoquant determined by τ in their respective distributions. We can further decompose (11) into two more detailed effects.

$$\left(S^{<1|1>}\left(y_1^{<1|1>}(\tau), y_2^{<1|1>}(\tau)\right) - S^{<1|1>}\left(y_1^{<0|1>}(\tau), y_2^{<0|1>}(\tau)\right)\right) + \\ -\left(S^{<1|1>}\left(y_1^{<0|1>}(\tau), y_2^{<0|1>}(\tau)\right) - S^{<0|1>}\left(y_1^{<0|1>}(\tau), y_2^{<0|1>}(\tau)\right)\right) \tag{12}$$

The first difference in (12) concerns the comparison of wellbeing of women with wellbeing of hypothetical women who have access to men's conditional distribution, but who also preserve their own preferences i.e. substitution parameters α, β from $S^{<1|1>}$. They still 'remain' women and their 'deep parameters' (preferences) towards two goods do not change, but as a counterfactual exercise they are given access to men's wage structure and other good structure. Please note that $S^{<1|1>}\left(y_1^{<0|1>}(\tau), y_2^{<0|1>}(\tau)\right)$ is not constant along the iso-quant determined by τ, because the parameters are not suited for it to be constant. Therefore, $\left(S^{<1|1>}\left(y_1^{<1|1>}(\tau), y_2^{<1|1>}(\tau)\right) - S^{<1|1>}\left(y_1^{<0|1>}(\tau), y_2^{<0|1>}(\tau)\right)\right)$ is a vector of numbers, where the first expression is constant, but a second expression changes. In order to get a number we may take the average of these differences. The second difference in (12) concerns the change of substitution of parameters and thus the impact of this change on the overall effect.

Other interesting considerations and effects are possible. For example, we may fix women's wellbeing and ask at what point in the counterfactual distribution this level of wellbeing is achieved. This is the question of how parameters τ or α and β change in order to keep women at the same level of wellbeing. Clearly, parametrization opens possibilites for new interesting counterfactual effects.

4 Extensions to several populations

4.1 Benchmark model and results

We generalise the previous case to the case of several populations. We follow the notation and model presented in Chernozhukov et al. (2013). We suppose that the populations are labeled by $k \in \mathcal{K}$, and that for each population k, there is a random d_x-vector X_k of covariates and a random outcome d_y-vector Y_k. The covariate vector is observable in all populations, but the outcome is only observable in populations $j \in \mathcal{J} \subseteq \mathcal{K}$. We denote the support of X_k by $\mathcal{X}_k \subseteq \mathbb{R}^{d_x}$ and denote the region of interest for Y_j by $\mathcal{Y}_j \subseteq \mathbb{R}^{d_y}$. We assume for simplicity that the number of populations, $|K|$, is finite. Further, we define $\mathcal{Y}_j\mathcal{X}_j = \{(y,x) : y \in \mathcal{Y}_j, x \in \mathcal{X}_j\}$ and $\mathcal{YXJ} = \bigcup_{j\in\mathcal{J}} \mathcal{Y}_j\mathcal{X}_j \times \{j\}$, and generate other index sets by taking Cartesian products, for example, $\mathcal{JK} = \{(j,k) : j \in \mathcal{J}, k \in \mathcal{K}\}$. The counterfactual distribution and quantile functions are constructed by integrating the conditional distribution $F_{Y_j|X_j}$ in population j with respect to the covariate distribution F_{X_k} in population k, namely

$$F_{Y<j|k>}(y) := \int_{\mathcal{X}_k} F_{Y_j|X_j}(y|x)\mathrm{d}F_{X_k}(x),\ y \in \mathcal{Y}_j \tag{13}$$

$$Q_{Y<j|k>}(\tau) := \widehat{F}^{-1}_{Y_j|X_k}(\tau) = \{y \in \overline{\mathcal{Y}_j} : \widehat{F}_{Y_j|X_k}(y) < \tau\},\ \tau \in (0,1) \tag{14}$$

Such integral is well-defined if the support condition holds i.e. $\mathcal{X}_k \subseteq \mathcal{X}_j$ for all $(j,k) \in \mathcal{JK}$. Conditional quantile models imply conditional distribution models through the relation

$$F_{Y_j|X_j} = \int_{(0,1)} 1_{\{Q_{Y_j|X_j}(u|x)\leq y\}}\mathrm{d}u \tag{15}$$

We are interested in non-parametric estimates of $F_{Y<j|k>}$.

In what follows, we define a counterfactual effect as the result of a shift from one counterfactual distribution $F_{Y<l|m>}$ to another $F_{Y<j|k>}$ for some $j, l \in \mathcal{J}$ and $m, k \in \mathcal{K}$. We are interested in estimating and performing inference on the distribution effects

$$\Delta(y) = F_{Y<j|k>}(y) - F_{Y<l|m>}(y) \text{ and}$$

as well as other functionals of the counterfactual distribution.

We can now re-write several results stated in Section 3 for the case of an arbitrary number of groups which are compared. Firstly, Lemmas 1 and 2 can be stated as follows.

Lemma 4. *$F_{Y<j|k>1} = F_{Y_1<j|k>}$ and $F_{Y<j|k>2} = F_{Y_2<j|k>}$*

Proof.

$$F_{Y<j|k>1}(y_1) = \int_{\sup Y_{j2}} F_{Y<j|k>}(y_1, y_2)\mathrm{d}y_2 = \int_{\sup Y_{j2}<j|k>} \int_{\sup X_k} F_{Y_j|X_j}(y_1, y_2|x)\mathrm{d}x\mathrm{d}y_2 =$$
$$= \int_{\sup X_k} \int_{\sup Y_{j2}<j|k>} F_{Y_j|X_j}(y_1, y_2|x)\mathrm{d}y_2\mathrm{d}x = \int_{\sup X_k} F_{Y_{j1}|X_j}(y_1|x)\mathrm{d}x =$$
$$= F_{Y_1<j|k>}$$

□

Lemma 5.

$$\max(F_{Y<j|k>1}+F_{Y<j|k>2}-1, 0)-\min(F_{Y<l|m>1}, F_{Y<l|m>2}) \leq \Delta^{DE} = F_{Y<j|k>}(y)-F_{Y<l|m>}(y) \leq$$
$$\min(F_{Y<j|k>1}, F_{Y<j|k>2}) - \max(F_{Y<l|m>1} + F_{Y<l|m>2} - 1, 0)$$

A counterfactual effect (8) can be re-written as

$$\Delta^2 \quad = \quad \left(F_{Y<j|k>} - F_{Y<j|k>1}F_{Y<j|k>2}\right) \quad - \quad \left(F_{Y<l|m>} - F_{Y<l|m>1}F_{Y<l|m>2}\right).$$

As with S parametrization, we have the following quantile sets

$$Q_{<j|k>}(\tau) = \{y \in \mathbb{R}^2 : F_{<j|k>}(y) = \tau\}$$

and a related function $S^{<j|k>}\left(y_1^{<j|k>}(\tau), y_2^{<j|k>}(\tau)\right)$. Using this we define the following counterfactual effect

$$\Delta^4 = S^{<j|j>}\left(y_1^{<j|j>}(\tau), y_2^{<j|j>}(\tau)\right) - S^{<j|k>}\left(y_1^{<j|k>}(\tau), y_2^{<j|k>}(\tau)\right) \tag{16}$$

and its decomposition

$$\left(S^{<j|j>}\left(y_1^{<j|j>}(\tau), y_2^{<j|j>}(\tau)\right) - S^{<j|j>}\left(y_1^{<j|k>}(\tau), y_2^{<j|k>}(\tau)\right)\right) +$$
$$-\left(S^{<j|j>}\left(y_1^{<j|k>}(\tau), y_2^{<j|k>}(\tau)\right) - S^{<j|k>}\left(y_1^{<j|k>}(\tau), y_2^{<j|k>}(\tau)\right)\right). \tag{17}$$

4.2 When counterfactual effects are causal

Let J denote the random variable that describes a policy variable; $\mathcal{J}$ is a set of values and $j \in \mathcal{J}$ denotes a policy. Let $Y^* = (Y_j^*)$ denote a vector of potential outcome variables for policy j, whereas $Y = Y_J^*$ denote the realized outcome variable. When the policy J is not randomly assigned, it is well known that the distribution of the observed outcome Y conditional on $J = j$, that is, the distribution of $Y|J = j$, may differ from the distribution of Y_j^*. However, if J is randomly assigned conditional on the covariates X, that is, if the conditional exogeneity assumption holds, then the distributions of $Y|X, J = j$ and $Y_j^*|X$ agree. In this case, the observable conditional distributions have a causal interpretation and so do the counterfactual distributions generated from these conditionals by integrating out X. Formally, let $F_{Y_j^*|J}(y|k)$ denote the distribution of the potential outcome Y_j^* in the population with $J = k \in \mathcal{J}$. The causal effect of exogenously changing the policy from l to j on the distribution of the potential outcome in the population with realized policy $J = k$ is $F_{Y_j^*|J}(y|k) - F_{Y_l^*|J}(y|k)$. The policy J corresponds to an indicator for the population labels $j \in \mathcal{J}$, and the observed outcome and covariates are generated as $Y_j = (Y|J = j)$ and $X_k = (X|J = k)$.

Lemma 6 below shows that under conditional exogeneity, for any $j, k \in \mathcal{J}$, the counterfactual distribution $F_{Y<j|k>}(y)$ exactly corresponds to $F_{Y_j^*|J}(y|k)$ and, hence, the causal effect of exogenously changing the policy from l to j in the population with $J = k$ corresponds to the CE of changing the conditional distribution from l to j, that is

$$F_{Y_j^*|J}(y|k) - F_{Y_l^*|J}(y|k) = F_{Y<j|k>}(y) - F_{Y<l|k>}(y)$$

Lemma 6. *Suppose that Y^* is independent of $J|X$ almost surely. Assuming $\mathcal{X}_k \subseteq \mathcal{X}_j$ we get*

$$F_{Y<j|k>}(\cdot) = F_{Y_j^*|J}(\cdot|k) \text{ for } j, k \in \mathcal{J}.$$

Now let $j \in \mathcal{J} = \{0, 1\}$. Formally, let $F_{Y_j^*|J}(y|1)$ denote the distribution of the potential outcome Y_j^* in the women's population. The causal effect of exogenously changing the policy from 1 to 0 on the distribution of the potential outcome in the women's population is $F_{Y_1^*|J}(y|1) - F_{Y_0^*|J}(y|1)$. Lemma 6 is then

$$F_{Y<1|1>}(\cdot) = F_{Y_1^*|J}(\cdot|1) \text{ and } F_{Y<0|1>}(\cdot) = F_{Y_0^*|J}(\cdot|1).$$

Causal interpretation of $F_{<1|1>} - F_{<0|1>}$ implies causal interpretation of $(F_{<1|1>} - \tilde{F}_{<0|1>})$, because using Sklar's theorem if Y is independent on J conditional on X, then the copula of Y is also independent.

5 Conclusions

In this paper we tackle the problem of estimating counterfactual effects, but when the outcome is multidimensional. This concerns for example estimation of well-known gender gaps, but in the case when we compare not only men's and women's wages, but also for example their life expectancy, or any other outcome of interest. Therefore, we call these effects well-being gaps.

The multidimensional framework requires new definitions of well-known counterfactual effects. For distribution effects, the definitions are the same, but new counterfactual effects can be proposed given the multidimensional nature of outcomes. In particular, we utilize Sklar's representation of a multidimensional distribution in terms of a copula and marginal distributions. Therefore, in the definition of counterfactual effect we can separately vary copula and marginals, which gives rise to new definitions. They also have natural interpretation, for example, keeping copula fixed can be interpreted as keeping the preferences towards well-being outcomes fixed for a given group.

The problem that appears in a multidimensional context that is absent from a unidimensional framework concerns quantile effects. In our context, quantiles are sets, and there appears a problem of how to define the difference between two sets. While this problem can be solved in theory (e.g. Hausdorf measure of distance), it is a priori far from obvious whether such measures will be easy tools for estimation and inference. Therefore, we solve this problem by adhering to the approach of Maasoumi and Racine (2016) who parametrize quantiles via well-known entropy well-being measures and then estimate substitution parameters of these measures. This approach allows for a high degree of heterogeneity i.e. substitution parameters can be different depending on the quantile. With this tool, we can again define new counterfactual effects.

Finally, we study the role of dependence (i.e. substitution) between outcomes in the overall counterfactual effect. Utilizing bounds on copula function, we can bound the joint counterfactual effect. We define new effects that allow us to single out dependence from the overall effect. This part of the theory, however, requires more work.

Acknowledgement

This research was funded by National Science Centre, Poland, grant number 2017/26/M/HS4/00905.

References

Albrecht J, Bjorklund A, Vroman S. 2003. In there a glass ceiling in Sweden?, Journal of Labor Economics 21: 145-177.

Arellano M, Bonhomme S. 2017. Quantile Selection Models: with an application to understanding changes in wage inequality, Econometrica 85: 1-28.

Atkinson A, Bourguignon F. 1982. The comparison of multi-dimensioned distributions of economic status, Review of Economic Studies 49: 183-201.

Atkinson A, Bourguignon F. 1987. Income distribution and differences in needs. In: Feiwel

G. (Ed.), Arrow and the Foundations of the Theory of Economic Policy, Macmillan.
Blau F D, Kahn L M. 1997. Swimming Upstream: Trends in the Gender Wage Differential in the 1980s, Journal of Labor Economics 15: 1-42.
Blau F D, Kahn L M. 2006. The U.S. Gender Pay Gap in the 1990s: slowing convergence, Industrial and Labor Relations Review 60: 45-66.
Blau F D, Kahn L M. 2017. The Gender Wage Gap: Extent, Trends, and Explanations, Journal of Economic Literature (forthcoming).
Chernozhukow V, Fernández-Val I, Melly B. 2013. Inference on counterfactual distributions, Econometrica 81(6): 2205-2268.
Chernozhukov V, Imbens G, Newey W K. 2007. Instrumental variable estimation of non-separable models, Journal of Econometrics 139(1): 4-14.
Chernozhukov V, Hansen C. 2005. An IV model of quantile treatment effects, Econometrica 73(1): 245-261.
Chesher A. 2005. Nonparametric identification under discrete variation, Econometrica 73(5): 1525-1550.
Deaton A. 2015. The Great Escape: Health, Wealth, and the Origins of Inequality, Princeton University Press.
Fortin N, Lemieux T, Firpo S. 2011. Decomposition Methods in Economics, Handbook of Labor Economics 4: 1-102.
Frandsen B, Lefgren L J. 2017. Testing rank similarity, Review of Economics and Statistics (forthcoming).
Goldin C. 2014. A Grand Gender Convergence: Its Last Chapter, American Economic Review 104: 1091-1119.
Heckman J J, Smith J, Clements N. 1997. Making the most out of programme evaluations and social experiments: accounting for heterogeneity in program impacts, Review of Economic Studies 64: 487-535.
Maasoumi E. 1986. The Measurement and Decomposition of Multi-dimensional Inequality, Econometrica, 54(4):991-997.
Maasoumi E, Wang L. 2017. The Gender Gap between Earnings Distributions, Maasoumi E, Racine J. 2016. A solution to aggregation and an application to multidimensional well-being frontiers, Journal of Econometrics 191: 374-383.
Mulligan C B, Rubinstein Y. 2008. Selection, investment and women's relative wages over time, Quarterly Journal of Economics 123: 1061-1110.
Nelsen R B. 2006. An Introduction to Copulas. 2nd edition, New York: Springer-Verlag.
Sen A K. 1973. On economic inequality. Oxford, UK: Clarendon Press.
Stiglitz J, Sen, A, Fitoussi, JP. 2009. Report by the Commission on the Measurement of Economic Performance and Social Progress, www.stiglitz-sen-fitoussi.fr