Université de Montréal

Mémoire de maîtrise

Identification des mesures d'inégalité dans les modèles de sélection

Rédigé par :
Désiré Kédagni

Dirigé par :
William McCausland

Département de sciences économiques
Faculté des arts et des sciences

Date de soumission
23 juillet 2015

Résumé

Dans ce mémoire, je considère un modèle de sélection standard avec sélection non aléatoire. D'abord, je discute la validité et la ''sharpness[1]'' des bornes sur l'intervalle interquantile de la distribution de la variable aléatoire latente non censurée, dérivées par Blundell et al. (2007). Ensuite, je dérive les bornes ''sharp[2]'' sur l'intervalle interquantile lorsque la distribution observée domine stochastiquement au premier ordre celle non observée. Enfin, je discute la ''sharpness'' des bornes sur la variance de la distribution de la variable latente, dérivées par Stoye (2010). Je montre que les bornes sont valides mais pas nécessairement ''sharp''. Je propose donc des bornes inférieures ''sharp'' pour la variance et le coefficient de variation de ladite distribution.

Mots clés : Sélection, Identification partielle, Mesures d'inégalité.

[1] veut dire étroitesse.

[2] c'est-à-dire les plus étroites possibles.

Contents

IDENTIFICATION OF INEQUALITY MEASURES IN SAMPLE SELECTION MODELS

DÉSIRÉ KÉDAGNI

Université de Montréal

ABSTRACT. In this master thesis, I consider a standard selection model with non-randomly censored outcome. First, I discuss validity and sharpness of bounds on the interquantile range of the distribution of the uncensored outcome, derived by Blundell et al. (2007). Second, I give sharp bounds on the interquantile range respectively under stochastic dominance of the unobserved outcome distribution by the observed one, and in presence of an exclusion variable. Third, I discuss sharpness of the variance bounds given by Stoye (2010). I show that the bounds are not necessarily sharp and I provide sharp lower bounds on the variance and the coefficient of variation.

Keywords: sample selection, partial identification, inequality measures.

JEL Classification: C14, C21, D31, D63.

Date: The present version is of August 29, 2015.

I'm deeply grateful to Marc Henry, William McCausland and Ismael Mourifié for their helpful comments and advice.

1. Introduction

The sample selection problem as discussed in Gronau (1974) and Heckman (1974, 1979) arises when the outcome of interest is only observed for a non-randomly selected subpopulation. Gronau (1974) criticizes the fact that the empirical validation of the labor economics theory is often based on the observed wage distribution whereas much of the theory concerning labor-force participation, wages, and earnings centers on the wage-offer distribution. Treating the wage offers and the observed wages as interchangeable is particularly suspect when there are substantial numbers of unemployed workers. In this case, the observed distribution represents only one part of the wage-offer distribution, while the other part is rejected by the job seekers as unacceptable.

Without additional assumptions, the wage-offer distribution is not point-identified, but only *partially identified* (see Manski, 1989). Parameters such as interquantile range, variance and coefficient of variation, as functions of the unidentified distribution, are also often only partially identified. Recognizing partial identification helps avoid selection bias at the expense of increasing uncertainty. The advantage is that the identification region contains the population parameter with probability one. The identification region is said to be *sharp* if it is the tightest set that includes the parameter of interest with probability one. Bounds on the wage-offer distribution are provided in Manski (1994, 2003), Stoye (2010) and in many other papers. Their bounds are pointwise sharp, but not functionally sharp because they do not take into account the functional property[1] mentioned in Crowder (1991), Bedford and Meilijson (1997), Vazquez-Alvarez, Melenberg and van Soest (2002), Blundell et al. (2007) and Henry et al. (2015). I explain this in detail in Section 3.

In this paper, I'm interested in identification of three well-known measures of dispersion: the variance, the coefficient of variation and the interquantile range. I explain how useful each of these measures are in Section 2. I show that the variance bounds given by Stoye (2010) are not necessarily sharp. I derive sharp lower bounds on the variance and the coefficient of variation.

Concerning the interquantile range, Vazquez-Alvarez, Melenberg and van Soest (1999) constructed bounds on quantiles in the presence of full and partial item nonresponse in the case where item nonresponse is nonrandom. Later, Vazquez-Alvarez, Melenberg and van Soest (2002), in their working paper "Selection bias and Measures of Inequality" provided bounds on the interquartile range[2], and sketched a proof of the sharpness of those bounds. Likewise, Blundell et al. (2007) derived bounds on the interquantile range, which turn out to be sharp, as I show below, but they did not show the sharpness of the bounds.

[1]Note however that this terminology is exclusively due to Henry et al. (2015).

[2]which is a spread case of the interquantile range for quartiles 0.25 and 0.75.

Considering a sample selection model, I prove under some assumptions the sharpness of bounds on the interquantile range derived by Blundell et al. (2007). Unlike the paper of Vazquez-Alvarez, Melenberg and van Soest (2002), I show conditions under which the bounds hold and are sharp. I also derive sharp bounds on the interquantile range under stochastic dominance of the unobserved outcome distribution by the observed one, and in presence of an exclusion variable.

This note is organized as follows: the first section presents the sample selection model discussed in this article, the second explains bounds on the distributions of interest; the third, bounds on the interquantile range as well as a proof of validity and sharpness and the fourth section discusses the sharpness of variance bounds.

2. The model

I consider the following censoring model $Y = Y_1 D$, where Y is a real-valued observed outcome, D is an observed selection indicator and Y_1 is a real-valued unobserved potential outcome. I denote by F_1 the distribution function of the potential outcome Y_1. For example, D could be the labor-force participation (equal to 1 if the individual is working and 0 otherwise) and Y_1 the wage-offer.

In this paper, I study three commonly used measures of dispersion for the distribution F_1: the variance, the coefficient of variation and the interquantile range. Unlike the variance, which depends on the unit of measurement, the coefficient of variation does not. For example, in comparing wage dispersion for different countries, say the U.K. (with the pound as currency) and the U.S. (with the dollar as currency), one cannot compare directly the wage variance for the two countries. But, we may compare instead their coefficients of variation. Both the variance and the coefficient of variation are easy to compute, but they are sensitive to outliers. The interquantile range is often used to avoid the comparison noise due to outliers. Although there are many other inequality measures (like the Gini index, the Theil index, etc.), these three basic measures of dispersion are the ones of interest in this article. They are defined in the following sections.

I assume, for example, that there is a positive minimum wage $y_{\min}$ and a maximum wage $y_{\max}$ that a worker cannot exceed. Therefore, I state the following assumption.

Assumption 1. *Y is bounded with compact support, i.e. $Supp(Y) \equiv [y_{\min}, y_{\max}]$.*

The following section summarizes the results in the literature about the distribution of the potential outcome Y_1.

3. Sharp Bounds on the distribution F_1

The starting point is the works of Manski (1994, 2003), which provide bounds on F_1. As I explain in the previous section, Y is a censored outcome, Y_1 is the potential outcome of interest and

D is the selection variable. We observe $Y = Y_1$ only when $D = 1$ and 0 otherwise. Let F_1 denote the c.d.f. of Y_1.

For all $y \in \mathbb{R}$, we have:

$$F_1(y) \equiv (\mathbb{P}(Y_1 \leq y) = \mathbb{P}(Y_1 \leq y, D = 1) + \mathbb{P}(Y_1 \leq y, D = 0) \tag{3.1}$$

$$F_1(y) = \mathbb{P}(Y_1 \leq y|D = 1)\mathbb{P}(D = 1) + \underbrace{\mathbb{P}(Y_1 \leq y|D = 0)}_{\textbf{counterfactual}}\mathbb{P}(D = 0) \tag{3.2}$$

Without additional assumptions, we only know that the counterfactual probability $\mathbb{P}(Y_1 \leq y|D = 0)$ lies between 0 and 1. Then, the distribution function F_1 is only partially identified. This gives the Manski pointwise bounds on the distribution F_1.

$$pF_{11}(y) \leq F_1(y) \leq pF_{11}(y) + 1 - p \tag{3.3}$$

where $F_{11}(y) = P(Y_1 \leq y|D = 1)$ and $p = \mathbb{P}(D = 1)$, $0 < p < 1$.

Since the functions $pF_{11}(y)$ and $pF_{11}(y) + 1 - p$ are not c.d.f.s [3], Stoye (2010) bounds the counterfactual probability by $F_L(y) = 1\{y \geq y_{\min}\}$ and $F_U(y) = 1\{y \geq y_{\max}\}$ instead of 0 and 1 respectively. Then we have the bounds of equation 3.4 below.

Let θ be a parameter or a function of interest. Denote $H(\theta)$ the identification region of θ. Then the identification region of the c.d.f of interest F_1 is that derived by Manski (1994) and Stoye (2010):

$$H(F_1) = \{F^* : pF_{11} + (1 - p)F_U \leq F^* \leq pF_{11} + (1 - p)F_L\} \tag{3.4}$$

But not all distributions within the identification region are observationally compatible with the data. Any distribution that is compatible with the data, in addition to be in the identification region, must satisfy equation (3.5) below (see figure 1), otherwise the identification region alone is not sharp. For all $y,\ y'$ such that $y < y'$

$$p\left(F_{11}(y') - F_{11}(y)\right) \leq F_1(y') - F_1(y) \tag{3.5}$$

Indeed, equation 3.5 follows from the following inequality.

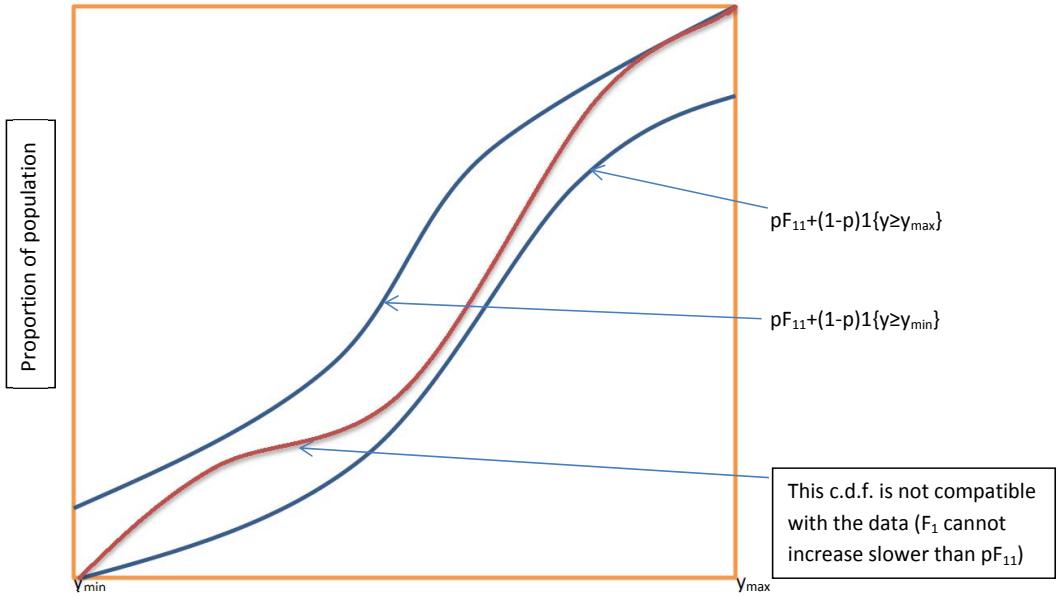
$$\mathbb{P}(y < Y_1 \leq y') \geq \mathbb{P}(y < Y_1 \leq y', D = 1)$$

The property implied by equation (3.5) is called the ***functional property*** [4] in Henry et al. (2005). This property states that the distribution $F_1(y)$ of the potential outcome Y_1, cannot increase slower than $pF_{11}(y)$ (see figure 1). I call $H^*(F_1)$ the *sharp identified set* of F_1, that is, the set of distribution functions that satisfy simutaneously equations (3.4) and (3.5).

[3] Because $pF_{11}(1) = p < 1$ and $pF_{11}(0) + 1 - p = 1 - p > 0$.

[4] This property is also derived by Crowder (1991), Bedford and Meilijson (1997),Vazquez-Alvarez, Melenberg and van Soest (2003), Blundell et al. (2007).

FIGURE 1. Bounds on the distribution F_1



Since the distribution of the potential outcome Y_1 is only partially identified, parameters that are functions of this distribution are often only partially identified too, unless additional assumptions are made. I explain through the following definitions what a partially identified parameter and sharp bounds for a scalar parameter mean precisely in this article.

Definition 1. *"A parameter is **partially identified** if the data generating process, together with assumptions a researcher is willing to make, reveals some **nontrivial information** about it but does not identify it in the conventional sense, that is, distinct parameter values may be **observationally equivalent**" (Stoye 2010).*

Definition 2. *Bounds for **a scalar parameter** are sharp if the lower and upper bounds are attainable and any convex combination of them is also attainable.*

I now present bounds on quantiles and interquantile ranges, and a proof of the sharpness of bounds on the interquantile range.

4. Sharp bounds on interquantile ranges

Definition 3. *Let $q \in [0,1]$. The quantile of order q of the c.d.f. F_1 is defined as*

$$Q(F_1;q) = \inf\{y \in supp(Y) : F_1(y) \geq q\} \ \forall\ q \in [0,1] \tag{4.1}$$

I use the convention in Stoye (2010) and set $Q(F_{11};q) = y_{\min}$ for $q \leq 0$ and $Q(F_{11};q) = y_{\max}$ for $q \geq 1$.

Definition 4. *Let $q_1, q_2 \in [0,1]$ such that $q_1 < q_2$. The interquantile range of orders q_1 and q_2 of F_1 is defined as*

$$IQR(F_1; q_1, q_2) = Q(F_1; q_2) - Q(F_1; q_1) \tag{4.2}$$

The interquantile range is the length of the interval containing $q_2 - q_1$ of the observations, leaving a fraction q_1 on its left and $1 - q_2$ on its right.

The following definition from Stoye (2010) may help understand the bounds for the quantile and the interquantile range.

Definition 5. (1) *θ is a D_1-parameter if it increases with first-order stochastic dominance:*

$$F(y) \leq G(y) \ \forall \ y \Longrightarrow \theta(F) \geq \theta(G). \tag{4.3}$$

(2) *θ is a D_2-parameter if for distributions that have equal expectation, it decreases with second-order stochastic dominance:*

$$\int y dF = \int y dG \ \& \ \int_{y_{\min}}^{k} F(y)dy \leq \int_{y_{\min}}^{k} G(y)dy \ \forall \ k \Longrightarrow \theta(G) \geq \theta(F). \tag{4.4}$$

For example, the quantile and the expectation are D_1-parameters while the variance and the coefficient of variation are D_2-parameters. But the interquantile range in neither a D_1 nor D_2-parameter.

Denote $y^l(q) = Q(F_{11}; 1 - \frac{1-q}{p})$ and $y^u(q) = Q(F_{11}; \frac{q}{p}) \ \forall \ q \in [0,1]$. Then, since the quantile is D_1-parameter, bounds on the quantile $Q(F_1;q)$ follow directly from bounds on the distribution F_1. That is,

$$y^l(q) \ \leq \ Q(F_1;q) \ \leq \ y^u(q) \tag{4.5}$$

Thus, bounds on the interquantile range $IQR(F_1; q_1, q_2) = Q(F_1; q_2) - Q(F_1; q_1), \ q_1 < q_2$, can be obtained by taking the difference of the bounds on the corresponding quantiles. Indeed,

$$y^l(q_2) - y^u(q_1) \ \leq \ IQR(F_1; q_1, q_2) \ \leq \ y^u(q_2) - y^l(q_1) \tag{4.6}$$

As the interquantile range is nonnegative, the lower bound is $IQR^l(F_1; q_1, q_2) = \max\left\{0, y^l(q_2) - y^u(q_1)\right\}$. Blundell et al. (2007) uses property (3.5) to tighten the upper bound. The idea is explained in detail in the proof of claim 1 in Appendix A.2. However, the following assumption is important for the validity of the bound.

Assumption 2. *F_{11} is continuous on $\left[y^l(q_1), y^u(q_2)\right)$ and strictly increasing on $[y^l(q_1), y^u(q_1))$.*

The following claim holds.

Claim 1. *Under assumptions 1 and 2, the following quantities given by Blundell et al. (2007)*

$$\begin{aligned} IQR^l(F_1;q_1,q_2) &= \max\left\{0, y^l(q_2) - y^u(q_1)\right\} \\ IQR^u(F_1;q_1,q_2) &= \sup_{y_0\in[y^l(q_1),y^u(q_1)]}\left\{Q\left(F_{11};F_{11}(y_0)+\frac{q_2-q_1}{p}\right)-y_0\right\} \end{aligned}$$

are sharp bounds for the interquantile range $IQR(F_1;q_1,q_2)$.

Proof. See Appendix A.2. □

Remark 1. *Assumption 2 is necessary for the validity of the bounds in claim 1 as I show in the proof. In the case where there is a jump in the distribution F_{11} on $[y^l(q_1), y^u(q_2))$, the bounds are not valid. For example, if there is only one jump in F_{11} at $\tilde{y}$ between y_0 and $\tilde{y}^u(q_2)$, then the following function*

$$F_1^{*y_0}(y) = \begin{cases} p\left(F_{11}(y) - F_{11}(y_0)\right) + q_1 \text{ if } y \in \left[y^l(q_1), \tilde{y}\right) \\ \\ p\left(F_{11}(y) - F_{11}(y_0)\right) + q_1 - p(F_{11}(\tilde{y}) - F_{11}(\tilde{y}^-)) \text{ if } y \in [\tilde{y}, y^u(q_2)) \end{cases}$$

gives the upper bound for the interquantile range. Notice that this function is continuous and satisfies the functional property.

Remark 2. *Because the interquantile range is partially identified, we may have uncertainty about its nondecreasingness as I explain next. The bounds on the interquantile range $Q(F_1;q_1,q_2)$ are such that for all $0 < q_1 < q_1' < q_2' < q_2 < 1$, $IQR^l(F_1;q_1',q_2') \leq IQR^l(F_1;q_1,q_2)$ and $IQR^u(F_1;q_1',q_2') \leq IQR^u(F_1;q_1,q_2)$. But, this does not tell us whether or not $IQR(F_1;q_1',q_2') \leq IQR(F_1;q_1,q_2)$, unless $IQR^u(F_1;q_1',q_2') \leq IQR^l(F_1;q_1,q_2)$.*

I now derive sharp bounds on the interquantile range under a commonly held assumption in the literature, the stochastic dominance assumption.

Assumption 3 (Stochastic dominance)**.** *F_{11} first order stochastically dominates F_{10}, that is:* $F_{11}(y) \leq F_{10}(y)\ \forall\ y$.

Denote $F(y) = \mathbb{P}(Y \leq y)$, $y^l(q) = Q(F_{11}; 1 - \frac{1-q}{p})$, $y_{11}(q) = Q(F_{11};q)$ and $y(q) = Q(F;q)$.

Assumption 4. *F_{11} is continuous on $\left[y^l(q_1), y_{11}(q_2)\right)$ and strictly increasing on $[y^l(q_1), y_{11}(q_1))$.*

Assumption 4 is technical and is the analog of assumption 2 under stochastic dominance of F_{10} by F_{11}. Assumption 3 however, is justified motivated by economic reasons. For example, in the case where the selection variable is the labor force participation, assumption 3 means that

the distribution of wages of workers first-order stochastically dominates that of nonworkers. This expresses a positive selection into the labor market[5].

Proposition 1. *Under assumptions 1, 4, and 3, sharp bounds for the interquantile range $IQR(F_1; q_1, q_2)$ are given by:*

$$\begin{aligned} IQR^{lD}(F_1; q_1, q_2) &= \max\left\{0, y^l(q_2) - y_{11}(q_1)\right\} \\ IQR^{uD}(F_1; q_1, q_2) &= \sup_{y_0 \in [y^l(q_2), y_{11}(q_2)]} \left\{ y_0 - Q\left(F_{11}; F_{11}(y_0) - \frac{q_2 - q_1}{p}\right)\right\} \end{aligned}$$

Proof. By a suitable adaptation of the proof of claim 1, the proof of this proposition 1 is straightforward. □

4.1. **Sharp bounds under exclusion restriction.** Having an instrument (exclusion variable) Z could help tighten the bounds. I consider here the cases where the exclusion variable has finite or compact support. In the case where its support is unbounded, the potential outcome distribution could be identified at infinity as I explain below. Although I consider a real-valued exclusion variable in this paper, I conjecture that the results generalize to the multidimensional exlcusion variable. I state the following assumption for the exclusion variable.

Assumption 5 (Exclusion Restriction)**.** *There is a variable Z such that Z is statistically independent of Y_1 i.e. $Z \perp Y_1$.*

Note that the exclusion variable Z affects the selection variable D, but not the potential outcome Y_1. Then for all $y \in \mathbb{R}$, we have

$$\begin{aligned} \mathbb{P}(Y_1 \leq y) &= \mathbb{P}(Y_1 \leq y | Z = z) \ \forall \ z \in \ Supp(Z) \\ &= \mathbb{P}(Y_1 \leq y, D = 1 | Z = z) + \mathbb{P}(Y_1 \leq y, D = 0 | Z = z) \ \forall \ z \in \ Supp(Z) \end{aligned}$$

If the support of Z is rich enough, we could find $z_\infty \in Supp(Z)$ such that $D(z_\infty) = 1$ a.s., that is, $P(Y_1 \leq y) = \mathbb{P}(Y_1 \leq y, D = 1 | Z = z_\infty) = \mathbb{P}(Y \leq y, D = 1 | Z = z_\infty) = \mathbb{P}(Y \leq y | Z = z_\infty)$. In this case, we say that the distribution of Y_1 is identified at infinity.

Notation 1. $F_{11}(y|z) = \mathbb{P}(Y_1 \leq y | D = 1, Z = z)$, $Q(F_{11}; q|z) = \inf\{y \in Supp(Y) : F_{11}(y|z) \geq q\}$, $p(z) = \mathbb{P}(D = 1 | Z = z)$, $y^l(q|z) = Q(F_{11}; 1 - \frac{1-q}{p(z)}|z)$ *and* $y^u(q|z) = Q(F_{11}; \frac{q}{p(z)}|z)$.

I use the following assumption throughout this subsection.

[5]See Blundell et al. (2007). This assumption stems from the fact that individuals with higher potential wages will be more likely to work unless the difference between wages and reservation wages is negatively associated with wages. Individuals with higher preference for work and low reservation wages can be expected to have invested more in human capital in the past and thus to end up with higher wages.

Assumption 6. $0 < p(z) < 1 \ \forall \ z \in Supp(Z)$.

4.1.1. *Discrete instrument with finite support.* Suppose that the support of Z is finite. Then I derive sharp bounds for the interquantile range under the following assumption.

Assumption 7. *$F_{11}(y|z)$ is continuous in y on $\left[y^l(q_1|z), y^u(q_2|z)\right)$ and strictly increasing in y on $[y^l(q_1|z), y^u(q_1|z))$ for every $z \in Supp(Z)$.*

Proposition 2. *Under assumptions 1, 5, 6 and 7, sharp bounds on $IQR(F_1; q_1, q_2)$ are*

$$\begin{aligned} IQR^{lE}(F_1; q_1, q_2) &= \max_z \left\{\max\left\{0, y^l(q_2|z) - y^u(q_1|z)\right\}\right\}, \\ IQR^{uE}(F_1; q_1, q_2) &= \min_z \left\{ \sup_{y_0 \in [y^l(q_1|z), y^u(q_1|z)]} \left\{ Q\left(F_{11}; F_{11}(y_0|z) + \frac{q_2 - q_1}{p(z)} | z \right) - y_0 \right\} \right\} \end{aligned}$$

.

Proof. See Appendix B.1. □

4.1.2. *Continuous instrument with compact support.* Suppose $Supp(Z) = \left[z^l, z^u\right]$. I add the following assumption to derive sharp bounds for the interquantile range.

Assumption 8. *$p(z)$ is continuous in z and $F_{11}(y|z)$ is continuous in z for all y.*

Proposition 3. *Under assumptions 1, 5, 6,7 and 8, sharp bounds on $IQR(F; q_1, q_2)$ are*

$$\begin{aligned} IQR^{lE}(F_1; q_1, q_2) &= \sup_{z \in [z^l, z^u]} \left\{\max\left\{0, y^l(q_2|z) - y^u(q_1|z)\right\}\right\}, \\ IQR^{uE}(F_1; q_1, q_2) &= \inf_{z \in [z^l, z^u]} \left\{ \sup_{y_0 \in [y^l(q_1|z), y^u(q_1|z)]} \left\{ Q\left(F_{11}; F_{11}(y_0|z) + \frac{q_2 - q_1}{p(z)} | z \right) - y_0 \right\} \right\} \end{aligned}$$

.

Proof. See Appendix B.2. □

5. Bounds on the variance and the coefficient of variation

In this section, I discuss the sharpness of bounds on the variance, derived by Stoye (2010) and I propose sharp lower bounds on the variance and the coefficient of variation. The definition of the variance is helpful to understand the derivation of the bounds.

Definition 6. *The variance measures the average absolute dispersion around the mean. The variance $V(F_1)$ is defined as*

$$V(F_1) = \mathbb{E}\left[Y_1 - \mu_1\right]^2 = \int_{y_{\min}}^{y_{\max}} (y - \mu_1)^2 dF_1(y) \tag{5.1}$$

where $\mu_1 = \mathbb{E}\left[Y_1\right]$.

5.1. **The Stoye(2010) joint bounds idea.** The joint bounds idea in Stoye (2010) is to provide the identified set $H(\mu_1)$ for the mean μ_1 and then for each value $\mu \in H(\mu_1)$, provide the identification region of the variance of all distributions that have μ as mean. Corollary 4 in Stoye (2010) states that given a fixed value of μ_1, the identification region for the variance[6] (when $Supp(Y) = [0,1]$) is:

$$H\left(V(F_1)\right) = \left[p\mu_{F_{11}^2} + (1-p)\mu_{10}^2 - \mu_1^2 \ , \ pV(F_{11}) + \mu_1 - \mu_1^2\right] \tag{5.2}$$

where $\mu_{11} \equiv \mathbb{E}\left[Y_1|D=1\right]$, $\mu_{10} \equiv \mathbb{E}\left[Y_1|D=0\right] = (\mu_1 - p\mu_{11})/(1-p) \ \forall \ p \in (0,1)$ and $\mu_{F_{11}^2} \equiv \mathbb{E}\left[Y_1^2|D=1\right]$.

I acknowledge this joint bounds idea is helpful as I show in the following lemma.

Lemma 1. *Let $H(\mu_1) = [p\mu_{11} + (1-p)y_{\min}, p\mu_{11} + (1-p)y_{\max}]$ be the identified set of the mean μ_1.*

If $H\left(V(\mu)\right) = \left[V^L(\mu), V^U(\mu)\right]$ is sharp for all $\mu \in H(\mu_1)$, then

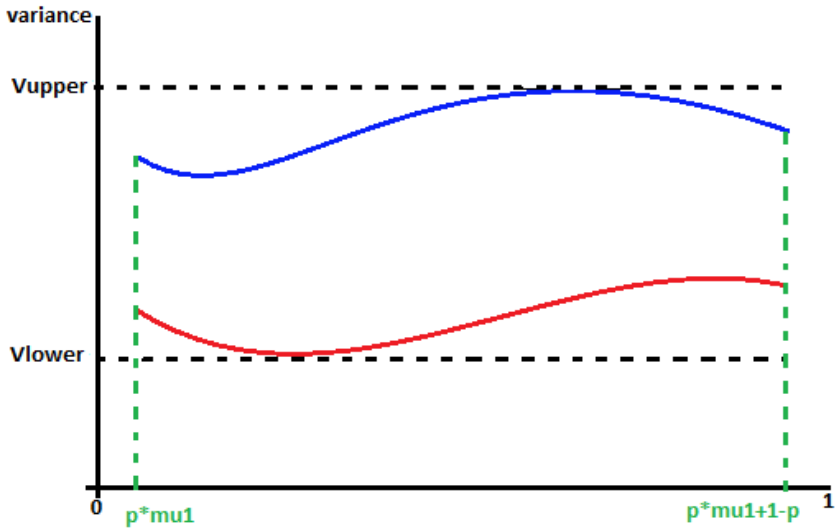
$$H\left(V(F_1)\right) = \bigcup_{\mu \in H(\mu_1)} H\left(V(\mu)\right) = \left[\inf_{\mu \in H(\mu_1)} V^L(\mu), \sup_{\mu \in H(\mu_1)} V^U(\mu)\right]$$

is also sharp.

Proof. See Appendix C.1. □

The following figure illustrates more the result of the lemma.

FIGURE 2. Illustration of the joint bounds idea by Stoye (2010) with $Supp(Y) = [0,1]$



[6]Stoye (2010) takes advantage on the fact that the variance is a D_2-parameter to derive its identification region. See Stoye (2010) for more details.

Remark 3. *The result of lemma 1 does not hold only for variance. It does hold for every function that is continuous in a translation in μ (e.g. the coefficient of variation when the identification region of μ does not contain 0).*

The idea of joint bounds is nice. However, the bounds on $V(F_1)$ in Stoye (2010) seem too large, even if everybody is observed working ($D = 1$). In fact, if bounds for a parameter are sharp, they should be identical to the true parameter whenever the probability p goes to 1, that is, if everybody is observed working. Obviously, the upper bound of Stoye for the variance goes to $V(F_1) + \mu_1 - \mu_1^2$ whenever p goes to 1, which is different from the true variance $V(F_1)$, unless μ_1 equals 0 or 1. So, the bounds for the variance $V(F_1)$ seem not to be sharp.

5.2. **Tighter bounds on the variance.** This subsection discusses the sharpness of bounds on variance in Stoye (2010) and provides tighter bounds.

Proposition 4. *Given μ_1, the following bounds are valid for the variance of the distribution F_1.*

$$V^L(F_1) \leq V(F_1) \leq V^U(F_1) \tag{5.3}$$

where

$$\begin{aligned} V^L(F_1) &= \mathbb{E}\left[(Y_1 - \mu_1)^2 | D = 1\right] + (1-p)(\mu_{10} - \mu_1)^2 \\ V^U(F_1) &= \min\left\{\mathbb{E}\left[(Y_1 - \mu_1)^2 | D = 1\right] + (1-p)\max([y_{\max} - \mu_1]^2, [y_{\min} - \mu_1]^2), p\mu_{F_{11}^2} + (1-p)y_{\max}^2 - \mu_1^2\right\} \end{aligned}$$

Proof. See Appendix C.2. □

Claim 2. *The lower bound $V^L(F_1)$ in proposition 4 is equal to the lower bound in Stoye (2010).*

Proof. See Appendix C.3. □

Claim 3. *The upper bound $V^U(F_1)$ in proposition 4 could be less than that in Stoye (2010) for some values of μ_1.*

For example, if $Supp(Y) = [0, 1]$ and $\mu_1 = p\mu_{11} + 1 - p$, the lower bound $V^L(F_1)$ is equal to the upper bound $V^U(F_1)$, which means that the variance is point-identified. But, the lower bound of Stoye (2010) is not equal to his upper bound. Then, the upper bound $V^U(F_1)$ is less than that of Stoye (2010).

Moreover, in the the following example, the upper bound $V^U(F_1)$ is less than that of Stoye.

Example 1. *Assume $F_1 \sim \mathcal{U}_{[0,1]}$, but this is unknown and $F_{11} \sim \mathcal{U}_{[0,p]}$.*

$$\mu_{F_{11}}^2 = \int_0^1 y^2 dF_{11}(y) = \int_0^p y^2/p dy = p^2/3; \quad \mu_1 = 1/2; \quad \mu_{11} = p/2; \quad \mu_{10} = \frac{1/2 - p/2 * p}{1-p} = (1+p)/2$$

In this cases, the lower and upper bounds of Stoye(2010) are respectively

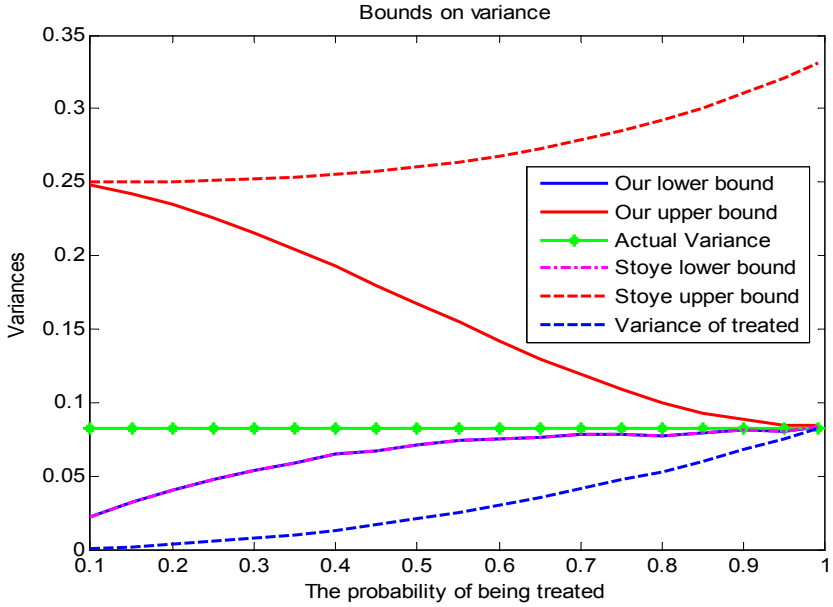
$$\begin{aligned} V^{SL} &= p*p^2/3+(1-p)*(\frac{1+p}{2})^2-(1/2)^2 = p^3/12-p^2/4+p/4 \\ V^{SU} &= p*p^2/12+1/2-(1/2)^2 = p^3/12+1/4 \end{aligned}$$

The upper bound in proposition 4 is

$$\begin{aligned} V^U &= \min\left\{p(p^2/12)+p(p/2-1/2)^2+(1-p)\max[(1-1/2)^2,(0-1/2)^2], p^3/3+(1-p)-1/4\right\} \\ &= \min\left\{1/3p^3-1/2p^2+1/4p+(1-p)*1/4, p^3/3+(1-p)-1/4\right\} \\ &= 1/3p^3-1/2p^2+1/4p+(1-p)/4 \end{aligned}$$

Cleary, $V^{SU}-V^U = p^2(2-p)/4 > 0 \ \forall \ p \in (0,1)$. *Then, in this example, the upper bound in proposition 4 is much tighter than that of Stoye for all* $p \in (0,1)$. *Figure 3 shows the gap between the bounds for* $p \in [0.1, 0.99]$.

FIGURE 3. Stoye's bounds and my bounds on variance for $F_1 \sim \mathcal{U}_{[0,1]}$ and $F_{11} \sim \mathcal{U}_{[0,p]}$.



Then, the following proposition holds.

Proposition 5. *Given* $\mu_d \in H(\mu_d)$, *tighter bounds on variance are given by:*

$$H^*\left(V(F_1)\right) = \left[V^{SL}(F_1), \min\left\{V^{SU}(F_1), V^U(F_1)\right\}\right] \tag{5.4}$$

where V^{SL} *and* V^{SU} *are respectively the lower and upper bounds in Stoye (2010),* V^U *is the upper bound in proposition 4.*

In example 1, this proposition yields $H^*\left(V(F_1)\right) = \left[V^{SL}(F_1), V^U(F_1)\right] \ \forall \ \mu_1 \in H(\mu_1)$.

5.3. **The sharp lower bound on the variance.** I propose here the sharp lower bound for the variance of F_1 when there is no information about the mean μ_1.

Proposition 6. *The sharp lower bound on variance is given by:*

$$V^{L^*}(F_1) = pV(F_{11}) \tag{5.5}$$

Proof. See Appendix C.4. □

Remark 4. *I don't provide sharp upper bound for the variance because I don't find for each μ_1, a c.d.f. that achieves the bound. However, given μ_1, the upper bound $V^U(F_1)$ may be attainable. For example, for $\mu_1 = p\mu_{11} + (1-p)y_{\max}$, the upper bound $V^U(F_1) = p\mu_{F_{11}^2} + (1-p)y_{\max}^2 - \mu_1^2$ is attained by $F_1(y) = pF_{11}(y) + (1-p)1\{y \geq y_{\max}\}$.*

5.4. **The sharp lower bound on the coefficient of variation.**

Definition 7. *The coefficient of variation measures the average relative dispersion around the mean. It's defined as*

$$CV(F_1) = \frac{\sqrt{V(F_1)}}{\mu_1} = \frac{1}{\mu_1}\sqrt{\int_{y_{\min}}^{y_{\max}} (y-\mu_1)^2 dF_1(y)} \tag{5.6}$$

Proposition 7. *The sharp lower bound on the coefficient of variation is given by:*

$$CV^{L^*}(F_1) = \min\left(CV(pF_{11} + (1-p)1\{Y \geq y_{\max}\}), CV(pF_{11} + (1-p)1\{Y \geq y_{\min}\})\right) \tag{5.7}$$

Proof. See Appendix C.5. □

6. Conclusion

This master thesis discusses the sharpness of bounds on basic inequality measures: the interquantile range, the variance and the coefficient of variation. I show that the bounds derived by Blundell et al. (2007) are sharp. I provide, in the same way, sharp bounds on interquantile range under stochastic dominance and exclusion restriction assumptions. However, variance bounds provided by Stoye (2010) are not sharp. I explain this through an example in which the bounds fail to be sharp. Accordingly, I provide the sharp lower bounds on the variance and the coefficient of variation.

Appendix A. Proof of sharpness of bounds on the interquantile range

To prove claim 1, I use throughout the following lemma.

Lemma 2. *Let $\alpha \in (0,1)$. Then, the following holds.*

$$\forall\ \epsilon > 0,\ F(y-\epsilon) < \alpha \text{ and } F(y) \geq \alpha \Leftrightarrow Q(F;\alpha) = y$$

A.1. **Proof of lemma 2.**

Proof. "⇒", By way of contradiction

$$\begin{aligned}
\exists\ y_0\ s.t.\ y_0 < y \text{ and } F(y_0) \geq \alpha \quad &\Rightarrow \quad F(y_0) \geq \alpha > F(y-\epsilon)\ \forall\ \epsilon > 0 \\
&\Rightarrow \quad y_0 > y - \epsilon \text{ As } F \text{ is nondecreasing} \\
&\Rightarrow \quad y - \epsilon < y_0 < y\ \forall\ \epsilon > 0 \\
&\Rightarrow \quad y \leq y_0 < y \\
&\Rightarrow \quad y < y \text{ (absurd).}
\end{aligned}$$

"⇐", obvious from the definition of a quantile. □

A.2. **Proof of claim 1.**

Proof. - Validity: The validity of the lower bound $IQR^l(F_1; q_1, q_2) = \max\left\{0, y^l(q_2) - y^u(q_1)\right\}$ is explained in the main text. I focus on the upper bound. Take $y_0 \in \left[y^l(q_1), y^u(q_1)\right]$. According to the functional property, any c.d.f. $F_1^{y_0}$ that pass trough y_0 and that could have generated the data should have at least the same slope as the following function $\tilde{F}_1^{y_0} = p(F_{11}(y) - F_{11}(y_0)) + q_1$ over $\left[y^l(q_1), y^u(q_2)\right)$. This implies that $\tilde{F}_1^{y_0}$ first-order stochastically dominates $F_1^{y_0}$ over $\left[y^l(q_1), y^u(q_2)\right)$. Therefore, as the quantile is a D_1-parameter, we have $Q(F_1^{y_0}; q_2) \leq Q(\tilde{F}_1^{y_0}; q_2)$. Then $IQR(F_1^{y_0}; q_1, q_2) = Q(F_1^{y_0}; q_2) - Q(F_1^{y_0}; q_1) \leq Q(\tilde{F}_1^{y_0}; q_2) - y_0$. Notice that strict increasingness of F_{11} on $\left[y^l(q_1), y^u(q_1)\right]$ ensures that $Q(\tilde{F}_1^{y_0}; q_1) = y_0$. Thus, for any potential distribution F_1, we have $IQR(F_1; q_1, q_2) \leq \sup_{y_0 \in [y^l(q_1), y^u(q_1)]} \left\{Q(\tilde{F}_1^{y_0}; q_2) - y_0\right\} = IQR^u(F_1; q_1, q_2) = \sup_{y_0 \in [y^l(q_1), y^u(q_1)]} \left\{Q\left(F_{11}; F_{11}(y_0) + \frac{q_2 - q_1}{p}\right) - y_0\right\}$.

Now, assume that there is a jump in F_{11} at $\tilde{y}$ between y_0 and $\tilde{y}^u(q_2)$. Then the following function

$$\overline{F}_1^{y_0}(y) = \begin{cases} p\left(F_{11}(y) - F_{11}(y_0)\right) + q_1 \text{ if } y \in \left[y^l(q_1), \tilde{y}\right) \\ \\ p\left(F_{11}(y) - F_{11}(y_0)\right) + q_1 - \epsilon \text{ if } y \in [\tilde{y}, y^u(q_2)) \end{cases}$$

such that $p\left(F_{11}(\tilde{y}) - F_{11}(\tilde{y}^-)\right) > \epsilon > 0$ could be completed to get a c.d.f. that may have generated the data and first-order stochastically dominates $\tilde{F}_1$. It's easy to see that $IQR(\overline{F}_1^{y_0}; q_1, q_2) \geq IQR(\tilde{F}_1^{y_0}; q_1, q_2)$.

- Sharpness: I have to show that the bounds are attainable and any element within $H(IQR(F_1; q_1, q_2))$ is also attainable, that is, for each element of $H(IQR(F_1; q_1, q_2))$, I have to show a distribution that attains it.

Let's consider the case where F_{11} is continuous and strictly increasing on $[y^l(q_1), y^u(q_2))$. Then, the quantile $Q\left(F_{11}; F_{11}(y_0) + (q_2 - q_1)/p\right)$ is the ordinary inverse function $F_{11}^{-1}(F_{11}(y_0) + (q_2 - q_1)/p)$.

Step 1: The upper bound is attainable

$$\tilde{F}_1(y) = \begin{cases} 0 \text{ if } y < y_{\min} \\ p\left(F_{11}(y) - F_{11}(y_0^*)\right) + q_1 \text{ if } y \in [y_{\min}, y_{\max}) \\ 1 \text{ if } y \geq y_{\max} \end{cases} \tag{A.1}$$

where $y_0^* = argmax_{y_0 \in [y^l(q_1), y^u(q_1)]} \left\{F_{11}^{-1}(F_{11}(y_0) + (q_2 - q_1)/p) - y_0\right\}$, ($y_0^*$ exists by Weierstrass maximum Theorem). By construction, y_0^* is the q_1th quantile of $\tilde{F}_1$ and $F_{11}^{-1}(F_{11}(y_0^*) + (q_2 - q_1)/p)$ is the q_2th quantile of $\tilde{F}_1$. Indeed,

$$\begin{aligned} \tilde{F}_1(y_0^*) &= q_1 \\ \text{Let's take } \epsilon > 0. & \\ \tilde{F}_1(y_0^* - \epsilon) &= p\left(F_{11}(y_0^* - \epsilon) - F_{11}(y_0^*)\right) + q_1 \\ &< q_1 \text{ as } F_{11} \text{ is strictly increasing.} \end{aligned}$$

Then by lemma 2, y_0^* is the q_1th quantile of $\tilde{F}_1$.

$$\begin{aligned} \tilde{F}_1(F_{11}^{-1}(F_{11}(y_0^*) + (q_2 - q_1)/p)) &= p\left(F_{11}(F_{11}^{-1}(F_{11}(y_0^*) + (q_2 - q_1)/p)) - F_{11}(y_0^*)\right) + q_1 \\ &= p\left(F_{11}(y_0^*) + (q_2 - q_1)/p - F_{11}(y_0^*)\right) + q_1 \\ &= q_2 \\ \text{Let's take } \epsilon > 0. & \\ \tilde{F}_1(F_{11}^{-1}(F_{11}(y_0^*) + (q_2 - q_1)/p) - \epsilon) &= p\left(F_{11}(F_{11}^{-1}(F_{11}(y_0^*) + (q_2 - q_1)/p) - \epsilon) - F_{11}(y_0^*)\right) + q_1 \\ &< p\left(F_{11}(F_{11}^{-1}(F_{11}(y_0^*) + (q_2 - q_1)/p)) - F_{11}(y_0^*)\right) + q_1 \\ &\text{as} \quad F_{11} \text{ is strictly increasing.} \\ &= p\left(F_{11}(y_0^*) + (q_2 - q_1)/p - F_{11}(y_0^*)\right) + q_1 \\ &= q_2 \end{aligned}$$

Then by lemma 2, $F_{11}^{-1}(F_{11}(y_0^*) + (q_2 - q_1)/p)$ is the q_2th quantile of $\tilde{F}_1$.

Step 2: The lower bound is attainable

Case 1: $y^l(q_2) \leq y^u(q_1)$

Then $IQR^l\left(F_1; q_1, q_2\right) = 0$

$$\tilde{F}_1(y) = pF_{11}(y) + (1 - p)1\left\{y \geq y^u(q_1)\right\} \tag{A.2}$$

$$\begin{aligned}\tilde{F}_1(y^u(q_1)) &= pF_{11}(y^u(q_1)) + (1-p) \\ &= \left(pF_{11} + (1-p)1\left\{Y_1 \geq y_{\min}\right\}\right)(y^u(q_1)) \\ &> \left(pF_{11} + (1-p)1\left\{Y_1 \geq y_{\min}\right\}\right)(y^l(q_2)) \\ &\geq q_2 > q_1\end{aligned}$$

Let's take $\epsilon > 0$.

$$\begin{aligned}\tilde{F}_1(y^u(q_1) - \epsilon) &= pF_{11}(y^u(q_1) - \epsilon) \\ &= \left(pF_{11}(y) + (1-p)1\left\{Y_1 \geq y_{\max}\right\}\right)(y^u(q_1) - \epsilon) \\ &< q_1 < q_2\end{aligned}$$

Then by lemma 2, $Q\left(\tilde{F}_1; q_2\right) = y^u(q_1) = Q\left(\tilde{F}_1; q_1\right)$. From where, $IQR\left(\tilde{F}_1; q_1, q_2\right) = 0$.

Case 2: $y^l(q_2) > y^u(q_1)$

Then $IQR^l\left(F_1; q_1, q_2\right) = y^l(q_2) - y^u(q_1)$

$$\tilde{F}_1(y) = pF_{11}(y) + (1-p)1\left\{y \geq y^l(q_2)\right\} \tag{A.3}$$

$$\begin{aligned}\tilde{F}_1(y^u(q_1)) &= pF_{11}(y^u(q_1)) \\ &= \left(pF_{11} + (1-p)1\left\{Y_1 \geq y_{\max}\right\}\right)(y^u(q_1)) \\ &\geq q_1\end{aligned}$$

Let's take $\epsilon > 0$.

$$\begin{aligned}\tilde{F}_1(y^u(q_1) - \epsilon) &= pF_{11}(y^u(q_1) - \epsilon) \\ &= \left(pF_{11} + (1-p)1\left\{Y_1 \geq y_{\max}\right\}\right)(y^u(q_1) - \epsilon) \\ &< q_1\end{aligned}$$

Then by lemma 2, $Q\left(\tilde{F}_1; q_1\right) = y^u(q_1)$.

$$\begin{aligned}\tilde{F}_1(y^l(q_2)) &= \left(pF_{11} + (1-p)1\left\{Y_1 \geq y^l(q_2)\right\}\right)(y^l(q_2)) \\ &\geq q_2\end{aligned}$$

Let's take $\epsilon > 0$.

$$\begin{aligned}\tilde{F}_1(y^l(q_2) - \epsilon) &= pF_{11}(y^l(q_2) - \epsilon) \\ &= \left(pF_{11} + (1-p)1\left\{Y_1 \geq y_{\max}\right\}\right)(y^l(q_2) - \epsilon) \\ &< q_2\end{aligned}$$

Then by lemma 2, $Q\left(\tilde{F}_1; q_2\right) = y^l(q_2)$.

Step 3: Any element between the lower and upper bounds is attainable

Let $\beta \in H(IQR(F_1; q_1, q_2))$. Keeping in mind that $y_0^* + \beta \leq F_{11}^{-1}(F_{11}(y_0^*) + (q_2 - q_1)/p)$, I distinguish two cases:

case 1: $y_0^* + \beta \geq y^l(q_2)$

$$\tilde{F}_1(y) = \begin{cases} 0 \text{ if } y < y_{\min} \\ p\left(F_{11}(y) - F_{11}(y_0^*)\right) + q_1 \text{ if } y \in [y_{\min}, y_0^* + \beta) \\ pF_{11}(y) + 1 - p \text{ if } y \geq y_0^* + \beta \end{cases} \tag{A.4}$$

Then $\tilde{F}_1(y_0^*) = q_1$ and $\forall\ \epsilon > 0,\ \tilde{F}_1(y_0^* - \epsilon) < q_1 \Rightarrow Q(\tilde{F}_1; q_1) = y_0^*$. And, $\tilde{F}_1(y_0^* + \beta) \geq q_2$ and $\forall\ \epsilon > 0,\ \tilde{F}_1(y_0^* + \beta - \epsilon) < q_2 \Rightarrow Q(\tilde{F}_1; q_2) = y_0^* + \beta$.

case 2: $y_0^* + \beta < y^l(q_2)$

Then, choose $\tilde{y} \in \left[y_0^*, y^l(q_2)\right]$ s.t. $\tilde{y} + \beta \geq y^l(q_2)$ and $p\left(F_{11}(\tilde{y} + \beta) - F_{11}(\tilde{y})\right) + q_1 \leq q_2$[7], and define:

$$\tilde{F}_1(y) = \begin{cases} 0 \text{ if } y < y_{\min} \\ p\left(F_{11}(y) - F_{11}(\tilde{y})\right) + q_1 \text{ if } y \in [y_{\min}, \tilde{y} + \beta) \\ pF_{11}(y) + 1 - p \text{ if } y \geq \tilde{y} + \beta \end{cases} \tag{A.5}$$

Then $\tilde{F}_1(\tilde{y}) = q_1$ and $\forall\ \epsilon > 0,\ \tilde{F}_1(\tilde{y} - \epsilon) < q_1 \Rightarrow Q(\tilde{F}_1; q_1) = \tilde{y}$. And, $\tilde{F}_1(\tilde{y} + \beta) \geq q_2$ and $\forall\ \epsilon > 0,\ \tilde{F}_1(\tilde{y} + \beta - \epsilon) < q_2 \Rightarrow Q(\tilde{F}_1; q_2) = \tilde{y} + \beta$. □

Appendix B. Proofs for sharp bounds on IQR under exclusion restriction assumption

B.1. **Sharp bounds with discrete instrument.**

Proof. - Validity: Straightforward

- Sharpness: Since Z is discrete with finite support, then there exist $\underline{z}_0$ and $\overline{z}_0$ that achieve respectively the lower and upper bounds. That is,

$$\begin{aligned} IQR^{lE}(F_1; q_1, q_2) &= \max\left\{0, y^l(q_2|\underline{z}_0) - y^u(q_1|\underline{z}_0)\right\}, \\ IQR^{uE}(F_1; q_1, q_2) &= \sup_{y_0 \in [y^l(q_1|\overline{z}_0), y^u(q_1|\overline{z}_0)]} \left\{ Q\left(F_{11}; F_{11}(y_0|\overline{z}_0) + \frac{q_2 - q_1}{p(\overline{z}_0)}|\overline{z}_0\right) - y_0 \right\} \end{aligned}$$

Therefore, considering the case where $F_{11}(y|z)$ is continuous and strictly increasing in y on $[y^l(q_1|z), y^u(q_2|z))$, the same distributions used to prove claim 1 conditionned on $\underline{z}_0$ and $\overline{z}_0$ respectively for the lower and the upper bounds achieve the bounds. □

[7]Note that $\tilde{y}$ exists, since $p\left(F_{11}(y) - F_{11}(\tilde{y})\right) + q_1$ is a horizontal translation of $p\left(F_{11}(y) - F_{11}(y_0^*)\right) + q_1$, its value at $\tilde{y} + \beta$ is at most q_2.

B.2. **Sharp bounds with continuous instrument.** I use the following lemma to come up with the proof of sharpness of the bounds.

Lemma 3. *Let $F(y|z)$ be a conditional cumulative distribution function of a real-valued random variable Y with compact support. Assume that $F(y|z)$ is continuous in y for all z and continuous in z for all y. Then, the quantile function $Q(F;q|z)$ defined for every $q \in (0,1)$ by*

$$Q(F;q|z) = \inf\{y \in Supp(Y) : F(y|z) \geq q\} \tag{B.1}$$

is also continuous in z for all q.

Proof. Notice that $Q(F;q|z)$ is the unique solution of the following optimization problem:

$$\min f(z,y) = y \ s.t. \ y \in \Gamma(z) = \{y \in Supp(Y) : F(y|z) \geq q\}$$

I use the Theorem of the Maximum (Theorem 3.6 of Stokey and Lucas p.62). Like in Stokey and Lucas, I define

$$h(z) = \min_{y\in\Gamma(z)} y = -\max_{y\in\Gamma(z)} -y \tag{B.2}$$

$$G(z) = \{y \in \Gamma(z) : f(z,y) = h(z)\} \tag{B.3}$$

The function f is continuous. I'm going to show that the correspondence Γ is compact-valued and continuous.

Compactness: $\Gamma(z) \subset Supp(Y)$ compact. Then, $\Gamma(z)$ is bounded. Now, let $y_n \in \Gamma(z)$ s.t. $y_n \to y$. Let's show that $y \in \Gamma(z)$.

$$\begin{aligned} y_n \in \Gamma(z) &\Rightarrow F(y_n|z) \geq q \\ &\Rightarrow \lim_{n\to\infty} F(y_n|z) \geq q \\ &\Rightarrow F(\lim_{n\to\infty} y_n|z) \geq q \text{ by continuity of } F(y|z) \text{ in y} \\ &\Rightarrow F(y|z) \geq q \\ &\Rightarrow y \in \Gamma(z) \end{aligned}$$

Then, $\Gamma(z)$ is closed. Thus, $\Gamma(z)$ is compact.

Continuity: I show that $\Gamma(z)$ is lower hemicontinuous (l.h.c.) and upper hemicontinuous (u.h.c.). I use the definitions in Stokey and Lucas p.56. $\Gamma(z)$ is nonempty for all z.

l.h.c.: Take $y \in \Gamma(z)$. Then, $F(y|z) \geq q$. Let z_n be a sequence s.t. $z_n \to z$. By continuity of $F(y|z)$ in z, $F(y|z_n) \to F(y|z)$. That is, $\forall\ \epsilon > 0,\ \exists\ n_\epsilon : \forall\ n > n_\epsilon,\ |F(y|z_n) - F(y|z)| < \epsilon$, which implies that $F(y|z_n) > F(y|z) - \epsilon \geq q - \epsilon$. Hence, for $\epsilon \to 0,\ \exists\ n_0 :\ \forall\ n > n_0,\ F(y|z_n) \geq q$. Then, considering the sequence $\{y_n = y\}_{n=n_0}^{\infty}$, we have $y_n \to y$ and $y_n \in \Gamma(z_n)$. This shows that $\Gamma(z)$ is l.h.c..

u.h.c.: Take $z_n \to z$ and $y_n \in \Gamma(z_n)$. I'm going to show that there exists a subsequence $\{y_k\} \to y \in \Gamma(z)$. We have $F(y_n|z_n) \geq q \ \forall \ n$. Moreover, $y_n \in \Gamma(z_n)$, which is real-valued and bounded. Then, by Bolzano-Weierstrass theorem, there exists a subsequence $\{y_k\}$ s.t. $y_k \to y$. Now, it remains to show that $y \in \Gamma(z)$. Indeed, we have the following implications.

$$\begin{aligned} F(y_k|z_n) \geq q \ \forall \ n, \ \forall \ k \quad &\Rightarrow \quad \lim_{n\to\infty} F(y_k|z_n) \geq q \\ &\Rightarrow \quad F(y_k|\lim_{n\to\infty} z_n) = F(y_k|z) \geq q \text{ by continuity in z} \\ &\Rightarrow \quad \lim_{k\to\infty} F(y_k|z) \geq q \\ &\Rightarrow \quad F(\lim_{k\to\infty} y_k|z) = F(y|z) \geq q \text{ by continuity in y} \\ &\Rightarrow \quad y \in \Gamma(z) \end{aligned}$$

Then, $\Gamma(z)$ is u.h.c..

Therefore, by the Theorem of Maximum, the function $h(z)$ is continuous and the correspondence $G(z)$ is nonempty, compact-valued and u.h.c.. Since, the quantile function $Q(F;q|z)$ is the unique solution of the problem, $Q(F;q|z)$ is continuous in z. □

The following is the proof of proposition 3.

Proof. - Validity: Straightforward

- Sharpness: Under assumptions 6, 7 and 8, by lemma 3, the quantities $y^l(q|z) = Q(F_{11}; 1 - \frac{1-q}{p(z)}|z)$ and $y^u(q|z) = Q(F_{11}; \frac{q}{p(z)}|z)$ are continuous in z. Since the function max is continuous, $\max\left\{0, y^l(q_2|z) - y^u(q_1|z)\right\}$ is continuous in z. Then, by Weierstrass Theorem, there exists z_0 such that

$$IQR^{lE}(F_1; q_1, q_2) = \max\left\{0, y^l(q_2|\underline{z}_0) - y^u(q_1|\underline{z}_0)\right\}.$$

Therefore, considering the case where $F_{11}(y|z)$ is continuous and strictly increasing in y on $[y^l(q_1|z), y^u(q_2|z))$, the same distributions used to prove claim 1 conditionned on $\underline{z}_0$ attain the lower bound.

Moreover, the quantile $Q(F_{11}; q|z) = F_{11}^{-1}(q|z)$ is continuous in $q \in (q_1, q_2)$ for all z and in z for all $q \in (q_1, q_2)$. Then, under assumptions, $F_{11}(y_0|z) + \frac{q_2-q_1}{p(z)}$ is continuous in y_0 for all z (resp. in z for all y_0) and $Q(F_{11}; F_{11}(y_0|z) + \frac{q_2-q_1}{p(z)}) - y_0$ is continuous in y_0 for all z (resp. in z for all y_0). The correspondence $\Gamma(z) = \left[y^l(q_1|z), y^u(q_1|z)\right]$ is continuous in z[8], then by the Theorem of Maximum, the function

$$\sup_{y_0 \in [y^l(q_1|z), y^u(q_1|z)]} \left\{ Q\left(F_{11}; F_{11}(y_0|z) + \frac{q_2 - q_1}{p(z)}|z\right) - y_0 \right\}$$

is continuous in z. Therefore, by Weierstrass, there exist $\overline{z}_0$ and $y_0^*(\overline{z}_0)$ such that

[8]Under assumptions, the proof of the continuity of $\Gamma(z)$ is very similar to that of lemma 3.

$$IQR^{uE}(F_1; q_1, q_2) = Q\left(F_{11}; F_{11}(y_0^*(\overline{z}_0)|\overline{z}_0) + \frac{q_2 - q_1}{p(\overline{z}_0)}|\overline{z}_0\right) - y_0^*(\overline{z}_0).$$

Thus, the same distributions used in the proof of claim 1 conditionned on $\overline{z}_0$ achieve the upper bound. □

Appendix C. Proofs for bounds on the variance and the coefficient of variation

C.1. **Proof of lemma 1.**

Proof. $V^i(\mu)$ $(i = L,\ U)$ is continuous (in μ) over $H(\mu_1)$, which is compact. Then , by Weirstrass theorem, $\exists\ \underline{\mu}$ and $\overline{\mu}$ in $H(\mu_1)$ such that

$$\inf_{\mu \in H(\mu_1)} V^L(\mu) = V(\underline{\mu}) \text{ and } \sup_{\mu \in H(\mu_1)} V^U(\mu) = V(\overline{\mu})$$

Because $\underline{\mu}$ and $\overline{\mu}$ belong to $H(\mu_1)$ and $H(\mu_1)$ is sharp then, there exist $\underline{F}$ and $\overline{F}$ in $H^*(F_1)$ s.t. $\mu(\underline{F}) = \underline{\mu}$ and $\mu(\overline{F}) = \overline{\mu}$. We know that $V\left(\mu(\underline{F})\right)$ (resp. $V(\mu(\overline{F})) \in H\left(V(\mu(\underline{F}))\right)$ (resp. $H\left(V(\mu(\overline{F}))\right)$). As $H\left(V(\mu(\underline{F}))\right)$ (resp. $H\left(V(\mu(\overline{F}))\right)$) is sharp, there exists a distribution $\underline{F}^*$ (resp. $\overline{F}^*$) in $H^*(F_1)$ s.t. $V(\underline{F}^*) = V\left(\mu(\underline{F})\right)$ (resp. $V(\overline{F}^*) = V\left(\mu(\overline{F})\right)$). Q.E.D. □

C.2. **Proof of proposition 4.**

Proof. By the Law of Iterated Expectations (L.I.E.), the following equality holds:

$$\mathbb{E}\left[(Y_1 - \mu_1)^2\right] = p\mathbb{E}\left[(Y_1 - \mu_1)^2|D = 1\right] + (1-p)\mathbb{E}\left[(Y_1 - \mu_1)^2|D = 0\right]$$

The unidentified term is $\mathbb{E}\left[(Y_1 - \mu_1)^2|D = 0\right]$. Obviously, by monotonicity of (positive) integral,

$$0 \leq \mathbb{E}\left[(Y_1 - \mu_1)^2|D = 0\right] \leq \max\left\{\left[y_{\max} - \mu_1\right]^2, \left[y_{\min} - \mu_1\right]^2\right\} \tag{C.1}$$

But, taking an attentive look at $\mathbb{E}\left[(Y_1 - \mu_1)^2|D = 0\right]$ yields:

$$\begin{aligned}\mathbb{E}\left[(Y_1 - \mu_1)^2|D = 0\right] &= \mathbb{E}\left[((Y_1 - \mu_{10}) + (\mu_{10} - \mu_1))^2|D = 0\right] \\ &= \mathbb{E}\left[(Y_1 - \mu_{10})^2|D = 0\right] + (\mu_{10} - \mu_1)^2\end{aligned}$$

Since $0 \leq \mathbb{E}\left[(Y_1 - \mu_{10})^2|D = 0\right] \leq \max\left\{\left[y_{\max} - \mu_{10}\right]^2, \left[y_{\min} - \mu_{10}\right]^2\right\}$, then

$$(\mu_{10} - \mu_1)^2 \leq \mathbb{E}\left[(Y_1 - \mu_1)^2|D = 0\right] \leq \max\left\{\left[y_{\max} - \mu_{10}\right]^2, \left[y_{\min} - \mu_{10}\right]^2\right\} + (\mu_{10} - \mu_1)^2 \tag{C.2}$$

From equations C.1 and C.2, the followings holds.

$$\begin{aligned}&(\mu_{10} - \mu_d)^2 \leq \mathbb{E}\left[(Y_1 - \mu_1)^2|D = 0\right] \leq \\ &\min\left\{\max\left\{\left[y_{\max} - \mu_1\right]^2, \left[y_{\min} - \mu_1\right]^2\right\}, \max\left\{\left[y_{\max} - \mu_{10}\right]^2, \left[y_{\min} - \mu_{10}\right]^2\right\} + (\mu_{10} - \mu_1)^2\right\}\end{aligned}$$

Because $(y - \mu_1)^2 \leq (y - \mu_{10})^2 + (\mu_{10} - \mu_1)^2\ \forall\ y$ (triangle inequality), the second upper bound is higher than the first.

Now, let's rewrite $V(F_1)$.

$$\begin{aligned} V(F_1) &= \mathbb{E}\left[Y_1^2\right] - \mu_1^2 \\ &= p\mathbb{E}\left[Y_1^2|D=1\right] + (1-p)\mathbb{E}\left[Y_1^2|D=0\right] - \mu_1^2 \\ &\leq p\mu_{F_{11}^2} + (1-p)y_{\max}^2 - \mu_1^2 \end{aligned}$$

From where the result holds. □

C.3. **Proof of claim 2.**

Proof.

$$\begin{aligned} V^L(F_1) &= p\mathbb{E}\left[\left((y-\mu_{11}) + (\mu_{11}-\mu_1)\right)^2 |D=1\right] + (1-p)(\mu_{10}-\mu_1)^2 \\ &= p\mathbb{E}\left[(y-\mu_{11})^2 |D=1\right] + p(\mu_{11}-\mu_1)^2 + (1-p)(\mu_{10}-\mu_1)^2 \\ &= pV(F_{11}) + p\mu_{11}^2 + (1-p)\mu_{10}^2 - \mu_d^2 \\ &= p(V(F_{11}) + \mu_{11}^2) + (1-p)\mu_{10}^2 - \mu_1^2 \\ &= p\mu_{F_{11}^2} + (1-p)\mu_{10}^2 - \mu_1^2 \end{aligned}$$

□

C.4. **Proof of proposition 6.**

Proof. $\forall\ \mu_1 \in H(\mu_1)$, the c.d.f. $\tilde{F}_1(y) = pF_{11}(y) + (1-p)1\{y \geq \mu_{10}\}$ attains the lower bound $V^L(F_1) = \inf_{\mu_1 \in H(\mu_1)} \left\{p\mathbb{E}\left[(y-\mu_1)^2|D=1\right] + (1-p)(\mu_{10}-\mu_1)^2\right\} = \inf_{\mu_1 \in H(\mu_1)} \left\{p\mu_{F_{11}^2} + (1-p)\mu_{10}^2 - \mu_1^2\right\}$. Then from lemma 1, the lower bound $\inf_{\mu_1 \in H(\mu_1)} V^L(F_1)$ is attainable. The function

$$\mu_1 \longmapsto \inf_{\mu_1 \in H(\mu_1)} \left\{p\mu_{F_{11}^2} + (1-p)\mu_{10}^2 - \mu_1^2\right\}$$

is strictly convex. Then the first order condition (f.o.c.) is sufficient to get a global minimum. The f.o.c. implies

$$\begin{aligned} & 2(1-p)(\frac{1}{1-p})\mu_{10} - 2\mu_1 = 0 \\ \Rightarrow\ & 2(\mu_{10}-\mu_1) = 0 \\ \Rightarrow\ & \mu_1 = \mu_{10} \\ \Rightarrow\ & \mu_1 = \mu_{11} \end{aligned}$$

The second derivative w.r.t. μ_1 is equal to $2p/(1-p)$, which is greater than 0 as $p \in (0,1)$. Replacing μ_1 by μ_{11} in the lower bound yields $V^L(F_1) = pV(F_{11})$. □

C.5. **Proof of proposition 7.**

Proof. By definition, $CV(F_1) = \frac{\sqrt{V(F_1)}}{\mu_1}$. Then, as $V(F_1) \geq V^L(F_1)\ \forall\ \mu_1 \in H(\mu_1)$, I have $\frac{\sqrt{V(F_1)}}{\mu_1} \geq \frac{\sqrt{V^L(F_1)}}{\mu_1}\ \forall\ \mu_1 \in H(\mu_1)$, that is, $CV(F_1) \geq \frac{\sqrt{V^L(F_1)}}{\mu_1}\ \forall\ \mu_1 \in H(\mu_1)$. This lower bound $\frac{\sqrt{V^L(F_1)}}{\mu_1}$ is attained by the distribution $\tilde{F}_1 = pF_{11}(y) + (1-p)1\{y \geq \mu_{10}\}\ \forall\ \mu_1 \in H(\mu_1)$. And, because the coefficient of variation $CV(F_1)$ is continuous (in μ_1) over $H(\mu_1)$, lemma 1 applies, that is,

$$CV^{L*}(F_1) = \inf_{\mu_1 \in H(\mu_1)} \left\{ \frac{\sqrt{p\mu_{F_{11}^2} + (1-p)\mu_{10}^2 - \mu_1^2}}{\mu_1} \right\}$$

As the function

$$\mu_1 \longmapsto \frac{\sqrt{p\mu_{F_{11}^2} + (1-p)\mu_{10}^2 - \mu_1^2}}{\mu_1}$$

is continuous on the compact set $H(\mu_1) = [p\mu_{11} + (1-p)y_{\min}, p\mu_{11} + (1-p)y_{\max}]$, the minimum is attained (Weirstrass).

First, I show that there is no interior solution to this minimization problem. The f.o.c. for an interior solution is

$$\begin{aligned}
&\frac{2(\mu_{10} - \mu_1)\mu_1 - V^L(F_1)}{\mu_1^2 \left[V^L(F_1)\right]^{1/2}} = 0 \\
\Rightarrow\ & V^L(F_1) = 2(\mu_{10} - \mu_1)\mu_1 \\
\Rightarrow\ & \mu_{10} - \mu_1 \geq 0 \text{ and } V^L(F_1) = 2(\mu_{10d} - \mu_1)\mu_1 \\
\Rightarrow\ & \mu_1 \geq \mu_{11} \text{ and } V^L(F_1) = 2(\mu_{10} - \mu_1)\mu_1 \\
\Rightarrow\ & CV^{L*}(F_1) = \inf_{\mu_1 \in [\mu_{11}, p\mu_{11} + (1-p)y_{\max}]} \frac{\left[2(\mu_{10} - \mu_1)\mu_1\right]^{1/2}}{\mu_1} \\
\Rightarrow\ & CV^{L*}(F_1) = \inf_{\mu_1 \in [\mu_{11}, p\mu_{11} + (1-p)y_{\max}]} \left[2p/(1-p) * (1 - \mu_{11}/\mu_1)\right]^{1/2} \\
\Rightarrow\ & \mu_1^* = \mu_{11},\ V^L(F_1) = 2p/(1-p)(\mu_1 - \mu_{11})\mu_1 = 0 \text{ and } V^L(F_1) = pV(F_{11}) > 0 \text{ (contradiction)}
\end{aligned}$$

Then, $CV^L(F_1)$ is minimized at $\mu_1 = p\mu_{11} + (1-p)y_{\min}$ or $\mu_1 = p\mu_{11} + (1-p)y_{\max}$, that is, $\mu_{10} = y_{\min}$ or $\mu_{10} = y_{\max}$, and the corresponding distribution is $pF_{11}(y) + (1-p)1\{y \geq y_{\min}\}$ or $pF_{11}(y) + (1-p)1\{y \geq y_{\max}\}$. This completes the proof. □

References

Bedford, T., and I. Meilijson (1997): "A Characterisation Of Marginal Distributions Of (Possibly Dependent) Lifetime Variables Which Right Censor Each Other," *The Annals of Statistics*, 25(4), 1622–1645.

Blundell, R., A. Gosling, H. Ichimura, and C. Meghir (2007): "Changes In The DIstribution Of Male And Female Wages Accounting For Employment Composition Using Bounds," *Econometrica*, 75(2), 323–363.

Crowder, M. (1991): "On the Identifiability Crisis in Competing Risks Analysis," *Scandinavian Journal of Statistics*, 18(3), 223–233.

Gronau, R. (1974): "Wage Comparisons - A Selectivity Bias," *Journal of Political Economy*, 82(6), 1119–1143.

Heckman, J. (1974): "Shadow Prices, Market Wages, and Labor Supply," *Econometrica*, 42(4), 679–694.

Heckman, J. (1979): "Sample Selection Bias as a Specification Error," *Econometrica*, 47(1), 153–161.

Henry, M., R. Méango, and I. Mourifié (2015): "Sharp bounds for the Roy model," *Unpublished manuscript*.

Manski, C. F. (1989): "Anatomy of the selection problem," *Journal of Human resources*, 24, 343–360.

Manski, C. F. (1994): "The Selection Problem," *in Advances Economics, Sixth World Congress, C. Sims (ed.), Cambridge University Press*, 1, 143–170.

Manski, C. F. (2003): "Partial Identification of Probability Distributions," *Springer-Verlag*.

Stokey, N. L., and J. R. E. Lucas (1989): *Recursive Methods in Economic Dynamics, with E. C. Prescott*. Havard university press edn.

Stoye, J. (2010): "Partial Identification of Spread Parameters," *Quantitative Economics*, 1, 323–357.

Vazquez-Alvarez, R., B. Melenberg, and A. van Soest (1999): "Bounds on Quantiles in the Presence of Full and Partial Item Nonresponse," *Unpublished manuscript*.

Vazquez-Alvarez, R., B. Melenberg, and A. van Soest (2002): "Selection bias and measures of inequality," *ISSC Discussion Paper Series, WP2003/04, University College Dublin. Institute for the Study of Social Change*.