Charpentier, Arthur; Ka, Ndéné; Mussard, Stéphane; Ndiaye, Oumar Hamady

Article

Gini regressions and heteroskedasticity

Econometrics

Provided in Cooperation with:
MDPI – Multidisciplinary Digital Publishing Institute, Basel

Suggested Citation: Charpentier, Arthur; Ka, Ndéné; Mussard, Stéphane; Ndiaye, Oumar Hamady (2019) : Gini regressions and heteroskedasticity, Econometrics, ISSN 2225-1146, MDPI, Basel, Vol. 7, Iss. 1, pp. 1-16,
https://doi.org/10.3390/econometrics7010004

This Version is available at:
https://hdl.handle.net/10419/247504

econometrics

MDPI

Article

Gini Regressions and Heteroskedasticity †

Arthur Charpentier [1], Ndéné Ka [2], Stéphane Mussard [3,*] and Oumar Hamady Ndiaye [3]

[1] Centre de Recherche en Economie et Management (CREM), Université de Rennes, 35000 Rennes, France; arthur.charpentier@univ-rennes1.fr

[2] Département D'économie, Université Alioune Diop de Bambey, Bambey BP 30, Senegal; ndene.ka@uadb.edu.sn

[3] CHROME, Université de Nîmes, 30000 Nîmes, France; ndiayeomarhamady@gmail.com

* Correspondence: stephane.mussard@unimes.fr; Tel.: +33-468221644

† The authors would like to thank their three reviewers. The usual disclaimer applies.

Received: 20 July 2018; Accepted: 4 January 2019; Published: 14 January 2019

Abstract: We propose an Aitken estimator for Gini regression. The suggested $\mathcal{A}$-Gini estimator is proven to be a U-statistics. Monte Carlo simulations are provided to deal with heteroskedasticity and to make some comparisons between the generalized least squares and the Gini regression. A Gini-White test is proposed and shows that a better power is obtained compared with the usual White test when outlying observations contaminate the data.

Keywords: Gini; heteroskedasticity; jackknife; U-statistics

JEL Classification: C14; C3

1. Introduction

Among ℓ_1 regressions, the Gini regression initiated by Olkin and Yitzhaki (1992) is increasingly used in econometrics. It enables traditional hypotheses to be relaxed such as the linearity of the model. Moreover, it is well suited for the study of variables contaminated by outliers or measurement errors. The reader is referred to Yitzhaki and Schechtman (2013) for a complete overview of the Gini methodology.

Shelef and Schechtman (2011) and Carcea and Serfling (2015) investigated independently the use of the Gini autocovariance functions to estimate, respectively, the parameter of AR(1) and ARMA processes in the case of heavy tailed distributions such as Pareto processes. Recently, Mussard and Ndiaye (2018) investigated the semi-parametric Gini regression for vector autoregressive models in which non-spherical disturbances occur. They showed that premultiplying the model by a matrix that neutralizes the Gini covariance of the error terms may produce non-biased Gini estimators.

In the context of semi-parametric Gini regressions, we showed that the Aitken transformation (Aitken 1935) for non-spherical disturbances based on the variance provides exactly the same estimator obtained by neutralizing the Gini covariance of the error term. However, the convergence of the former estimator requires the existence of the second moment of the error term, whereas the latter is a U-statistics. Monte Carlo simulations are addressed in order to show the superiority of the Aitken-Gini estimator compared with the traditional GLS estimator in the presence heteroskedasticity. It is also shown that the usual White test to detect heteroskedasticity should be done in the Gini sense, that is, by testing the Gini covariance of the regressors instead of their variance. In this case, more power is obtained for small samples. Finally, a feasible generalized Gini regression is provided. It consists in estimating the residuals of the regression (with the semi-parametric Gini regression) and to plug those residuals in the model to purge the heteroskedasticity (with a Gini instrumental variable regression). Monte Carlo simulations prove the superiority of this algorithm in the presence of outlying observations compared with the usual White algorithm based on generalized least squares.

The remainder of this paper is outlined as follows. We begin in Section 2 with the two versions of the Aitken-Gini estimator based, respectively, on the variance and on the Gini covariance, before showing their equivalence. Section 3 is devoted to the convergence property. Section 4 presents Monte Carlo simulations. It is shown that the combination of nonspherical disturbances and outliers (measurement errors) imply a loss of efficiency, which is not so important in the Gini case compared with GLS. In addition, the power of White's test is analyzed in the presence of outlying observations. Section 5 closes the article.

2. Aitken-Gini Estimators

It is common practice to deal with heteroskedasticity and autocorrelation with generalized least squares. However, if the data are contaminated by outliers, the loss of efficiency can drastically affect the coefficient estimates. In the following, outliers are assumed to be contaminated data, such as measurement errors, which lead to bad estimates. It is shown that the employ of the Aitken-Gini estimator may be preferred to GLS when the data are contaminated by outlying observations. The model is the following:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta}_g + \boldsymbol{\varepsilon}_g, \tag{1}$$

where $\mathbf{y}$ is the dependent variable ($n \times 1$ vector), $\mathbf{X} \equiv [x_{ik}]$ is the matrix of the regressors (of size $n \times K$ with a first column of ones), $\boldsymbol{\beta}_g$ is the $K \times 1$ vector of parameters to be estimated, and $\boldsymbol{\varepsilon}_g$ is the vector of perturbation terms (of size $n \times 1$). Following Olkin and Yitzhaki (1992), the semi-parametric Gini regression yields an estimator of $\boldsymbol{\beta}_g$,

$$\hat{\boldsymbol{\beta}}_g = (\mathbf{R}_{\mathbf{x}}'\mathbf{X})^{-1}\mathbf{R}_{\mathbf{x}}'\mathbf{y}, \tag{2}$$

where $\mathbf{R}_{\mathbf{x}}$ is the rank matrix of $\mathbf{X}$. The rank matrix $\mathbf{R}_{\mathbf{x}}$ is the matrix in which, for each regressor $\mathbf{x}_k$ ($k = 1, \ldots, K$), the observations x_{ik} ($i = 1, \ldots, n$) are replaced by their rank within $\mathbf{x}_k$ (the smallest value of x_{ik} is replaced by 1, and the highest one by n). Olkin and Yitzhaki (1992) showed that Gini estimators may be of particular relevance when outliers arise in the data. It is worth mentioning that, in the sequel, only the semi-parametric Gini regression is investigated. The parametric Gini regression is a numerical technique relying on the minimization of the Gini index of the residuals, which yields the same estimator as the semi-parametric Gini regression when the model is linear.

2.1. Mimicking the Usual Aitken Estimator

The generalized least squares (GLS) technique requires, in the case of heteroskedasticity and non serial correlation, the traditional following hypotheses $\mathbb{E}(\varepsilon_{g,i}) = 0$, $\text{cov}(\varepsilon_{g,i}, \varepsilon_{g,j}) = 0$ for all $i \neq j$, and $\mathbb{E}(\varepsilon_{g,i}^2) = \sigma_i^2$, such that $\mathbb{E}[\boldsymbol{\varepsilon}_g \boldsymbol{\varepsilon}_g'] = \sigma^2\boldsymbol{\Omega}$ with

$$\boldsymbol{\Omega} = \begin{pmatrix} a_1 & 0 & 0 & \cdots & 0 \\ 0 & a_2 & 0 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \cdots & a_n \end{pmatrix}, \ a_i > 0, \ \forall i = 1, \ldots, n.$$

Let us denote by $\text{Var}(\boldsymbol{\varepsilon}_g) = \mathbb{E}[\boldsymbol{\varepsilon}_g \boldsymbol{\varepsilon}_g']$ the variance of the error term such that $\text{Var}(\boldsymbol{\varepsilon}_g) := \sigma^2\boldsymbol{\Omega}$. Let $\mathbf{P} = \boldsymbol{\Omega}^{-\frac{1}{2}}$, then setting $\mathbf{y}^* := \mathbf{P}\mathbf{y}$, $\mathbf{X}^* := \mathbf{P}\mathbf{X}$ and $\boldsymbol{\varepsilon}_g^* := \mathbf{P}\boldsymbol{\varepsilon}_g$ yields:[1]

$$\mathbf{P}\mathbf{y} = \mathbf{P}\mathbf{X}\boldsymbol{\beta}_g + \mathbf{P}\boldsymbol{\varepsilon}_g \iff \mathbf{y}^* = \mathbf{X}^*\boldsymbol{\beta}_g + \boldsymbol{\varepsilon}_g^*. \tag{3}$$

Thereby, a first Aitken-Gini estimator may be derived.

---

[1] This technique corresponds to the weighted least squares. Mathematically, things can be extended to the case where $\boldsymbol{\Omega}$ is not diagonal using the singular value decomposition, but interpretation is much harder. Only the diagonal case is studied here.

Proposition 1. *Let* $\mathbf{R}_{\mathbf{X}^*}$ *be the rank matrix of* $\mathbf{X}^*$*, then applying the usual semi-parametric Gini regression in Equation (2) to the model in Equation (3) yields:*

$$\hat{\boldsymbol{\beta}}_g = (\mathbf{R}'_{\mathbf{X}^*}\mathbf{X}^*)^{-1}\mathbf{R}'_{\mathbf{X}^*}\mathbf{y}^* \text{ such that } \mathbb{E}[\boldsymbol{\varepsilon}^*_g\boldsymbol{\varepsilon}^{*'}_g] = \sigma^2\mathbb{I}_n.$$

Proof. The application of the semi-parametric Gini regression to the model in Equation (3) is obvious. Note that:

$$\mathbb{E}[\boldsymbol{\varepsilon}^*_g\boldsymbol{\varepsilon}^{*'}_g] = \mathbf{P}\sigma^2\boldsymbol{\Omega}\mathbf{P}' = \sigma^2\mathbb{I}_n.$$

□

We obtain a result quite close to Equation (2), which has the form of estimators by instrumental variables (IV) (see Yitzhaki and Schechtman (2004) for the link between Gini regressions and IV). Indeed, setting $\mathbf{Z}' := \mathbf{R}'_{\mathbf{X}^*}\mathbf{P}$, we get an IV estimator:

$$\hat{\boldsymbol{\beta}}_g = (\mathbf{R}'_{\mathbf{X}^*}\mathbf{X}^*)^{-1}\mathbf{R}'_{\mathbf{X}^*}\mathbf{y}^* = (\mathbf{R}'_{\mathbf{X}^*}\mathbf{P}\mathbf{X})^{-1}\mathbf{R}'_{\mathbf{X}^*}\mathbf{P}\mathbf{y} = (\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{y}. \tag{4}$$

The Gini estimator in Equation (4) is derived by mimicking the usual Aitken estimator, which may be used if, and only if, the variability of the error term is defined with respect to the variance, i.e., $\mathbb{E}[\boldsymbol{\varepsilon}^*_g\boldsymbol{\varepsilon}^{*'}_g] = \sigma^2\mathbb{I}_n$. However, the Gini methodology is employed whenever the underlying variability is the covariance-Gini defined by Schechtman and Yitzhaki (1987), the co-Gini from now on, which is examined in the next subsection.

2.2. The Aitken-Gini Estimator

The usual Aitken estimator described in the previous subsection is valid whenever the second moments of $\boldsymbol{\varepsilon}_g$ are known and when no outliers occur in $\mathbf{X}$, for which the first moments exist. The Gini estimator may be one solution to overcome this difficulty without invoking the existence of the second moment of $\boldsymbol{\varepsilon}_g$. For that purpose, we must define the transformed model,

$$\mathbf{P}\mathbf{y} = \mathbf{P}\mathbf{X}\boldsymbol{\beta}_{ag} + \mathbf{P}\boldsymbol{\varepsilon}_g \iff \mathbf{y}^* = \mathbf{X}^*\boldsymbol{\beta}_{ag} + \boldsymbol{\varepsilon}^*_g, \tag{5}$$

such that there is no heteroskedasticity in the Gini sense, that is, the co-Gini of $\varepsilon_{g,i}$ remains constant for all $i = 1, \ldots, n$. Let the co-Gini operator be defined such that:

$$\operatorname{cog}(\varepsilon_{g,i}, \varepsilon_{g,i}) := \operatorname{cov}(\varepsilon_{g,i}, F_{\boldsymbol{\varepsilon}}(\varepsilon_{g,i})),$$

where $F_{\boldsymbol{\varepsilon}}(\varepsilon_{g,i})$ is the cumulative distribution function of $\varepsilon_{g,i}$. In this respect, we have $\mathbb{E}[\boldsymbol{\varepsilon}_g F'_{\boldsymbol{\varepsilon}}(\boldsymbol{\varepsilon}_g)] = g\boldsymbol{\Omega}^G$ with $g \geq 0$ such that,

$$\boldsymbol{\Omega}^G = \begin{pmatrix} b_1 & 0 & 0 & \cdots & 0 \\ 0 & b_2 & 0 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \cdots & b_n \end{pmatrix}, \; b_i > 0, \; \forall i = 1, \ldots, n.$$

The Aitken-Gini estimator must be defined according to the **P**-rank idempotent property of the transformation matrix **P**.

Definition 1. *A squared matrix* **P** *is said to be* **P***-rank idempotent if, for any given real random variable* X*, such that* $Z = \mathbf{P}X$*,*

$$F_{\mathbf{x}}(X) = F_{\mathbf{z}}(\mathbf{P}X),$$

where $F_{\mathbf{x}}$ *and* $F_{\mathbf{z}}$ *stand for the cumulative distribution functions of* X *and* Z*, respectively.*

For estimation purposes, this assumption implies that the rank vector of X remains invariant after any given transformation $\mathbf{P}$. This assumption is necessary to obtain spherical disturbances.

Proposition 2. *If there exists a* $\mathbf{P}$*-rank idempotent matrix such that* $F_{\boldsymbol{\varepsilon}^*}(\mathbf{P}\boldsymbol{\varepsilon}_g) = F_{\boldsymbol{\varepsilon}}(\boldsymbol{\varepsilon}_g)$ *then applying the usual semi-parametric Gini regression in Equation (2) to the model in Equation (5) yields the Aitken-Gini estimator:*

$$\hat{\boldsymbol{\beta}}_{ag} = (\mathbf{R}'_{\mathbf{x}^*}\mathbf{X}^*)^{-1}\mathbf{R}'_{\mathbf{x}^*}\mathbf{y}^* \text{ such that } \mathbb{E}[\boldsymbol{\varepsilon}^*_g F'_{\boldsymbol{\varepsilon}^*}(\boldsymbol{\varepsilon}^*_g)] = g\mathbb{I}_n.$$

Proof. If $\mathbf{P}$ is supposed to be $\mathbf{P}$-rank idempotent, then $F'_{\boldsymbol{\varepsilon}^*}(\mathbf{P}\boldsymbol{\varepsilon}_g) = F'_{\boldsymbol{\varepsilon}}(\boldsymbol{\varepsilon}_g)$. From the transformed model in Equation (5):

$$\mathbb{E}[\boldsymbol{\varepsilon}^*_g F'_{\boldsymbol{\varepsilon}^*}(\boldsymbol{\varepsilon}^*_g)] = \mathbb{E}[\mathbf{P}\boldsymbol{\varepsilon}_g F'_{\boldsymbol{\varepsilon}^*}(\mathbf{P}\boldsymbol{\varepsilon}_g)] = \mathbf{P}\mathbb{E}[\boldsymbol{\varepsilon}_g F'_{\boldsymbol{\varepsilon}^*}(\mathbf{P}\boldsymbol{\varepsilon}_g)] = g\mathbb{I}_n.$$

Since $F'_{\boldsymbol{\varepsilon}^*}(\mathbf{P}\boldsymbol{\varepsilon}_g) = F'_{\boldsymbol{\varepsilon}}(\boldsymbol{\varepsilon}_g)$, then

$$\mathbf{P}\mathbb{E}[\boldsymbol{\varepsilon}_g F'_{\boldsymbol{\varepsilon}}(\boldsymbol{\varepsilon}_g)] = \mathbf{P}g\boldsymbol{\Omega}^G = g\mathbb{I}_n \implies \mathbf{P} = [\boldsymbol{\Omega}^G]^{-1}.$$

Note that the invertibility of $\boldsymbol{\Omega}^G$ is ensured since it is positive semi-definite. □

2.3. A Reconciliation

In the previous subsections, two Aitkien-Gini estimators have been derived. We can actually show that the Gini estimator $\hat{\boldsymbol{\beta}}_g$ that mimics the GLS is equivalent to the Aitken-Gini estimator $\hat{\boldsymbol{\beta}}_{ag}$.

Proposition 3. *Let* ξ_i *be an i.i.d. process such that* $\mathrm{Var}(\xi_i) = \sigma^2$ *and* $\mathrm{cog}(\xi_i, \xi_i) = g > 0$*. Let* $\varepsilon_{gi} = \xi_i\sqrt{h(i)}$ *such that* $i = 1, \ldots, n$ *for some real-valued function* $h : \mathbb{N}_+ \to \mathbb{R}_{++}$ *and assume that* $\boldsymbol{\varepsilon}^*_g := \boldsymbol{\Omega}^{-\frac{1}{2}}\boldsymbol{\varepsilon}_g$ *and* $\widetilde{\boldsymbol{\varepsilon}}_g := [\boldsymbol{\Omega}^G]^{-1}\boldsymbol{\varepsilon}_g$*. Then, the following assertions hold:*

(i) $\mathrm{Var}(\varepsilon^*_{g,i}) = \sigma^2$.

(ii) $\mathrm{cog}(\varepsilon^*_{g,i}, \varepsilon^*_{g,i}) = g$.

(iii) $\hat{\boldsymbol{\beta}}_g = \hat{\boldsymbol{\beta}}_{ag}$ *and* $\widetilde{\boldsymbol{\varepsilon}}_g = \boldsymbol{\varepsilon}^*_g$.

Proof. (i) Let us remark that $\mathrm{Var}(\varepsilon_{gi}) \overset{iid}{=} \mathrm{Var}(\xi\sqrt{h(i)}) = \sigma^2 h(i)$. Consequently,

$$\sigma^2\boldsymbol{\Omega} = \sigma^2 \begin{pmatrix} h(1) & 0\cdots & 0 \\ 0 & h(i) & \vdots \\ \vdots & \cdots & h(n) \end{pmatrix}.$$

Thus, by Proposition 1 we get that $\mathbf{P} = \boldsymbol{\Omega}^{-\frac{1}{2}}$, consequently the transformed model provides $\boldsymbol{\varepsilon}^*_g = \mathbf{P}\boldsymbol{\varepsilon}_g = \boldsymbol{\Omega}^{-\frac{1}{2}}\boldsymbol{\varepsilon}_g = (\xi_1, \ldots, \xi_n)'$. Hence, $\mathrm{Var}(\varepsilon^*_{g,i}) = \mathrm{Var}(\xi_i) = \sigma^2$.

(ii) We have $\mathrm{cog}(\varepsilon_{g,i}, \varepsilon_{g,i}) \overset{iid}{=} \mathrm{cog}(\xi\sqrt{h(i)}, F(\xi\sqrt{h(i)})) = \sqrt{h(i)}\,\mathrm{cog}(\xi, F(\xi))$. Thereby, $\mathrm{cog}(\varepsilon_{g,i}, \varepsilon_{g,i}) = g\sqrt{h(i)}$ and so:

$$g\boldsymbol{\Omega}^G = g \begin{pmatrix} \sqrt{h(1)} & 0\cdots & 0 \\ 0 & \sqrt{h(i)} & \vdots \\ \vdots & \cdots & \sqrt{h(n)} \end{pmatrix}.$$

By Proposition 2, we get that $\mathbf{P} = [\boldsymbol{\Omega}^G]^{-1}$. The transformed model yields $\widetilde{\boldsymbol{\varepsilon}}_g = \mathbf{P}\boldsymbol{\varepsilon}_g = [\boldsymbol{\Omega}^G]^{-1}\boldsymbol{\varepsilon}_g = (\xi_1, \ldots, \xi_n)'$. Hence, $\mathrm{cog}(\widetilde{\varepsilon}_{g,i}, \widetilde{\varepsilon}_{g,i}) = \mathrm{cog}(\xi_i, \xi_i) = g$.

(iii) Considering Assertions (i) and (ii), it follows that $\boldsymbol{\varepsilon}_g = \widetilde{\boldsymbol{\varepsilon}}_g$. Note that both estimators $\hat{\boldsymbol{\beta}}_g$ and $\hat{\boldsymbol{\beta}}_{ag}$ are issued from $(\mathbf{R}'_{\mathbf{x}^*}\mathbf{X}^*)^{-1}\mathbf{R}'_{\mathbf{x}^*}\mathbf{y}^*$ with $\mathbf{X}^* = \mathbf{P}\mathbf{X}$, $\mathbf{y}^* = \mathbf{P}\mathbf{y}$ and $\mathbf{R}_{\mathbf{x}^*}$ the rank matrix of $\mathbf{X}^*$. Note that

$\mathbf{P} = \mathbf{\Omega}^{-\frac{1}{2}}$ is employed in the first case and $\mathbf{P} = [\mathbf{\Omega}^G]^{-1}$ in the second one. Since $\mathbf{\Omega}^{-\frac{1}{2}} = [\mathbf{\Omega}^G]^{-1}$, then $\hat{\boldsymbol{\beta}}_g = \hat{\boldsymbol{\beta}}_{ag}$, which concludes the proof. □

The previous results are all based on the rank matrix of $\mathbf{X}^*$; consequently, as shown by Olkin and Yitzhaki (1992), the semi-parametric Gini regression is robust to outliers. Although the previous proposition indicates that an equivalence exists between the two Aitken-Gini estimators $\hat{\boldsymbol{\beta}}_g$ and $\hat{\boldsymbol{\beta}}_{ag}$; it is noteworthy that $\hat{\boldsymbol{\beta}}_{ag}$ requires fewer assumptions, since the first moment of $\boldsymbol{\varepsilon}_g$ has to be known only, whereas $\hat{\boldsymbol{\beta}}_g$ is based on the existence of the two first moments of $\boldsymbol{\varepsilon}_g$.

3. Sampling Properties

The aim of this section is to show, as above, that two strategies are available to get the sampling variance of the Aitken-Gini estimator. The first one is to consider that the second moment of $\varepsilon_{g,i}$ exists, as usual in the case of least squares regression, to derive the asymptotic variance of $\hat{\boldsymbol{\beta}}_g$. The second one assumes that the second moment of $\varepsilon_{g,i}$ does not exist, thus the variance of $\hat{\boldsymbol{\beta}}_{ag}$ is derived by jackknife. For this purpose, one needs additional assumptions.

We start again with the transformed model:

$$\mathbf{Py} = \mathbf{PX}\boldsymbol{\beta}_{ag} + \mathbf{P}\boldsymbol{\varepsilon}_g \iff \mathbf{y}^* = \mathbf{X}^*\boldsymbol{\beta}_{ag} + \boldsymbol{\varepsilon}_g^*. \tag{6}$$

The hypotheses of the model are the following:

H1: $\mathbb{E}[\boldsymbol{\varepsilon}_g^*] = \mathbf{0}$.

H2: Whenever $\mathbf{P}$ is a non null matrix, the perturbation term $\boldsymbol{\varepsilon}_g^*$ is linearly approximated as follows:[2]

$$\boldsymbol{\varepsilon}_g^* = \mathbf{y}^* - \mathbf{X}^*\boldsymbol{\beta}_{ag}.$$

H3: The perturbation term $\boldsymbol{\varepsilon}_g^*$ is independent of $\mathbf{R}'_{\mathbf{x}^*}$.

H4: $\mathbf{R}'_{\mathbf{x}^*}\mathbf{X}^*$ is a positive definite matrix.

H5: The matrix $\mathbf{\Omega}^G$ is diagonal and it contains finite elements $\omega_i > 0$.

H6: The second moment of $\varepsilon_{g,i}^*$ exists for all $i = 1, \ldots, n$ such that $\mathbb{E}[(\boldsymbol{\varepsilon}_g^*)^2] = \sigma^2\mathbb{I}_n$.

Hypothesis **H2** is necessary because the semi-parametric Gini regression does not rely on the usual linearity assumption of the regressors. First, we prove that the estimators $\hat{\boldsymbol{\beta}}_g$ and $\hat{\boldsymbol{\beta}}_{ag}$ are unbiased. The proof is made for $\hat{\boldsymbol{\beta}}_{ag}$ only, since it is similar in both cases.

Proposition 4. *Under Hypotheses* **H1**–**H5**, $\hat{\boldsymbol{\beta}}_{ag}$ *is an unbiased estimator of* $\boldsymbol{\beta}_{ag}$.

Proof. From Equation (6), $\hat{\boldsymbol{\beta}}_{ag} = (\mathbf{R}'_{\mathbf{x}^*}\mathbf{X}^*)^{-1}\mathbf{R}'_{\mathbf{x}^*}\mathbf{y}^*$ and $\mathbf{X}^* = \mathbf{PX} = [\mathbf{\Omega}^G]^{-1}\mathbf{X}$, with $\mathbf{R}'_{\mathbf{x}^*}\mathbf{X}^*$ being invertible by Hypothesis **H4**, and with $\mathbf{\Omega}$ being invertible by Hypothesis **H5**. Let $\mathbf{Z}' := \mathbf{R}'_{\mathbf{x}^*}[\mathbf{\Omega}^G]^{-1}$, then we have by Hypothesis **H2**:

$$\hat{\boldsymbol{\beta}}_{ag} = (\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{y} = (\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'(\mathbf{X}\boldsymbol{\beta}_{ag} + \boldsymbol{\varepsilon}_g) = \boldsymbol{\beta}_{ag} + (\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\boldsymbol{\varepsilon}_g. \tag{7}$$

Since $\mathbf{Z}'\boldsymbol{\varepsilon}_g = \mathbf{R}'_{\mathbf{x}^*}[\mathbf{\Omega}^G]^{-1}\boldsymbol{\varepsilon}_g = \mathbf{R}'_{\mathbf{x}^*}\boldsymbol{\varepsilon}_g^*$, therefore, by Hypotheses **H1** and **H3**:

$$\mathbb{E}[\hat{\boldsymbol{\beta}}_{ag}] = \mathbb{E}[\boldsymbol{\beta}_{ag} + (\mathbf{Z}'\mathbf{X})^{-1}\mathbf{R}'_{\mathbf{x}^*}\boldsymbol{\varepsilon}_g^*] = \boldsymbol{\beta}_{ag}.$$

Note that Hypothesis **H3** is respected whenever outliers do not contaminate the sample. □

[2] The regression curve of the Gini regression does not require any linear assumption of the model. Only a linear approximation is necessary to estimate the error term.

3.1. Convergence

We suppose that the second moments of $\varepsilon^*_{g,i}$ exists [Hypothesis **H6**] in order to derive the asymptotic variance of $\hat{\boldsymbol{\beta}}_g$ and to check for its convergence. By the result of Proposition 3, we set $\mathbf{Z}' := \mathbf{R}'_{\mathbf{x}^*}[\boldsymbol{\Omega}]^{-\frac{1}{2}} = \mathbf{R}'_{\mathbf{x}^*}[\boldsymbol{\Omega}^G]^{-1}$.

Proposition 5. *Under Hypotheses* **H1**–**H6**, *the following assertions hold.*

(i) $\text{Var}(\hat{\boldsymbol{\beta}}_g|\mathbf{X}) = \sigma^2(\mathbf{R}'_{\mathbf{x}^*}[\boldsymbol{\Omega}^G]^{-1}\mathbf{X})^{-1}\mathbf{R}'_{\mathbf{x}^*}\mathbf{R}_{\mathbf{x}^*}[(\mathbf{R}'_{\mathbf{x}^*}[\boldsymbol{\Omega}^G]^{-1}\mathbf{X})^{-1}]'$.

(ii) $\hat{\boldsymbol{\beta}}_g$ *is convergent.*

Proof. (i) From Equation (7), we deduce that:

$$\begin{aligned}\text{Var}(\hat{\boldsymbol{\beta}}_g|\mathbf{X}) &= \mathbb{E}\left[(\hat{\boldsymbol{\beta}}_g - \boldsymbol{\beta}_g)(\hat{\boldsymbol{\beta}}_g - \boldsymbol{\beta}_g)'\right] \\ &= \mathbb{E}\left[(\mathbf{Z}'\mathbf{X})^{-1}\mathbf{R}'_{\mathbf{x}^*}\boldsymbol{\varepsilon}^*_g\boldsymbol{\varepsilon}^{*'}_g\mathbf{R}_{\mathbf{x}^*}[(\mathbf{Z}'\mathbf{X})^{-1}]'\right] \\ &= \sigma^2(\mathbf{R}'_{\mathbf{x}^*}[\boldsymbol{\Omega}^G]^{-1}\mathbf{X})^{-1}\mathbf{R}'_{\mathbf{x}^*}\mathbf{R}_{\mathbf{x}^*}[(\mathbf{R}'_{\mathbf{x}^*}[\boldsymbol{\Omega}^G]^{-1}\mathbf{X})^{-1}]'. \qquad (8)\end{aligned}$$

(ii) From Proposition 4, $\hat{\boldsymbol{\beta}}_g$ is an unbiased estimate of $\boldsymbol{\beta}_g$. We have $\mathbf{Z}'\mathbf{X} = \mathbf{R}'_{\mathbf{x}^*}[\boldsymbol{\Omega}^G]^{-1}\mathbf{X} = \mathbf{R}'_{\mathbf{x}^*}\mathbf{X}^*$. The matrix $\mathbf{Z}'\mathbf{X}$ exists since $\boldsymbol{\Omega}^G$ is invertible by Hypothesis **H5**. Thus by Hypothesis **H4**,

$$\text{plim}\frac{1}{n}\mathbf{Z}'\mathbf{X} = \text{plim}\frac{1}{n}\mathbf{R}'_{\mathbf{x}^*}[\boldsymbol{\Omega}^G]^{-1}\mathbf{X} = \text{plim}\frac{1}{n}\sum_{i=1}^{n}\omega_i^{-1}\mathbf{r}_i\mathbf{x}'_i$$

is a positive definite matrix (with $\mathbf{r}_i$ and $\mathbf{x}_i$ being rows of $\mathbf{R}'_{\mathbf{x}^*}$ and $\mathbf{X}$, respectively). Then, the asymptotic variance covariance matrix exists and amounts to:

$$as.\text{Var}(\hat{\boldsymbol{\beta}}_g|\mathbf{X}) = \frac{\sigma^2}{n}\left(\text{plim}\,\frac{1}{n}\mathbf{Z}'\mathbf{X}\right)^{-1}\left(\frac{1}{n}\mathbf{R}'_{\mathbf{x}^*}\mathbf{R}_{\mathbf{x}^*}\right)\left(\text{plim}\,\frac{1}{n}(\mathbf{Z}'\mathbf{X})'\right)^{-1}.$$

Letting n tend towards infinity, we get that $\lim_{n\to+\infty} as.\text{Var}(\hat{\boldsymbol{\beta}}_g|\mathbf{X}) = \mathbf{0}$. □

As mentioned by Yitzhaki and Schechtman (2013), the inference on the regressors of the semi-parametric Gini regression has to be performed with U-statistics. In this case, the convergence is ensured without invoking Hypothesis **H6**.

3.2. Convergence with U-Statistics

As shown by Yitzhaki and Schechtman (2013), Gini estimators are U-statistics. The main advantage of dealing with the class of U-statistics, based on the generalized notion of average, is to find unbiased estimators and to derive their asymptotic property. The reader is referred to Serfling (1980, chp. 5) for more details. A brief review of this chapter is provided below.

Let $X_1, X_2, \ldots, X_m$ be independent observations from a population on a distribution F. The parameter $\vartheta(F)$ of the population is a parametric function for which an unbiased estimator exists. It is expressed as:

$$\vartheta(F) = \mathbb{E}[\phi(X_1, X_2, \ldots, X_m)] = \int\cdots\int \phi(x_1, \ldots, x_m)dF(x_1)\cdots dF(x_m),$$

where $\phi(x_1, x_2, \ldots, x_m)$, called the kernel, is a symmetric function[3]. The U-statistics of $\vartheta(F)$ is an estimator based on a sample of size n, $X_1, \ldots, X_n$, such that $n \geq m$. Averaging the kernel ϕ, the U-statistics is written as:

$$U_n := U(X_1, X_2, \ldots, X_n) = \binom{n}{m}^{-1} \sum_c \phi(X_{i_1}, X_{i_2}, \ldots, X_{i_m}),$$

where $\sum_c$ denotes the sum over all combinations of m elements $\{i_1, \ldots, i_m\}$ from $\{1, \ldots, n\}$. The U-statistic for the parameter ϑ is an unbiased estimator of ϑ and the distribution of $\sqrt{n}(U - \vartheta)$ tends to a normal distribution as $n \to \infty$ under the condition that $\mathbb{E}[\phi^2(X_1, X_2, \ldots, X_m)] < \infty$. The variance of a U-statistic also relies on the existence of second moments. Let the sets $A := \{a_1, \ldots, a_m\}$ and $B := \{b_1, \ldots, b_m\}$ be composed of m distinct integers among the set $\{1, \ldots, n\}$, with c the number of common integers of sets A and B. Let $\tilde{\phi} := \phi - \vartheta$, then, by symmetry of $\tilde{\phi}$ as well as the independence of the observations $X_1, \ldots, X_n$ of the sample:

$$\xi_c := \mathbb{E}\{\tilde{\phi}(X_{a_1}, \ldots, X_{a_m})\tilde{\phi}(X_{b_1}, \ldots, X_{b_m})\}$$

Defining

$$U_n - \vartheta = \binom{n}{m}^{-1} \sum_c \tilde{\phi}(X_{i_1}, X_{i_2}, \ldots, X_{i_m}),$$

the variance of a U-statistic is given by:

$$\mathrm{Var}(U_n) = \mathbb{E}\{(U_n - \vartheta)^2\} = \binom{n}{m}^{-2} \sum_{c=0}^{m} \binom{n}{m}\binom{m}{c}\binom{n-m}{m-c}\xi_c,$$

with $\binom{n}{m}\binom{m}{c}\binom{n-m}{m-c}$ the number of possibilities for sets A and B to get c elements in common, and with the hypothesis $\mathbb{E}\{\phi^2(X_1, \ldots, X_m)\} < \infty$. The estimation by jackknife of the variance of U_n does not necessitate such an assumption:

$$\mathrm{Var}(U_n) = \frac{n-1}{n} \sum_{i=1}^{n} \left[U_{-i} - \frac{1}{n} \sum_{i=1}^{n} U_{-i} \right]^2,$$

where U_{-i} is the estimator based on a sample of size $n - 1$, without the ith observation.

Proposition 6. *Each element $\hat{\beta}_{ag,k}$ of the Aitken-Gini estimator $\hat{\boldsymbol{\beta}}_{ag} = (\hat{\beta}_{ag,1}, \ldots, \hat{\beta}_{ag,K})$ is a function of U-statistics, thus estimating the variance of $\hat{\beta}_{ag,k}$ by jackknife for all $k = 1, \ldots, K$ implies that $\hat{\beta}_{ag,k}$ is a consistent estimator of $\beta_{ag,k}$ such that $\hat{\beta}_{ag,k} \overset{a}{\sim} \mathcal{N}$ for all $k = 1, \ldots, K$, neither invoking the existence of the second moments of $\mathbf{X}$ nor those of $\boldsymbol{\varepsilon}_g$* **[H6]**.

Proof. See Appendix A. □

4. Tests and Simulations

In this section, it is shown that the semi-parametric Aitken-Gini estimator $\hat{\boldsymbol{\beta}}_{ag}$ is more robust than the usual GLS one when the data are contaminated by outliers with $\boldsymbol{\Omega}$ being known. Furthermore, a feasible generalized Gini regression is proposed to deal with the case where $\boldsymbol{\Omega}$ is unknown.

---

[3] If ϕ is not symmetric in its arguments, we can also average over the $m!$ permutations

$$\phi^*(X_1, X_2, \ldots, X_m) = (m!)^{-1} \sum_p \phi(x_{i_1}, x_{i_2}, \ldots, x_{i_m}),$$

with $\sum_p$ the sum over all permutations of $(1, \ldots, m)$.

4.1. Monte Carlo Simulations

We performed Monte Carlo simulations to assess the robustness of the semi-parametric Gini regression with outlying observations in $\mathbf{X}$ and the presence of heteroskedasticity. In this section, we assume that the heteroskedasticity shape is known. The steps of the Monte Carlo simulation were as follows.

Step 1: Generate three independent normal distributions $\mathbf{x}_j \sim \mathcal{N}(0,1)$ of size $n = 100$ such that $j = 2,3,4$, with $\mathbf{x}_1 = (1,\ldots,1)$ and $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4)$.

Step 2: Sort the matrix $\mathbf{X}$ by ascending order according to the vector $\mathbf{x}_2$ (the second column of $\mathbf{X}$). Multiply the last row of $\mathbf{X}$ by $\theta = 100,\ldots,10,000$ (with increments of 100) in order to inflate the most important value of $\mathbf{x}_2$: $\mathbf{X}_n^o := \theta \mathbf{X}_n$.

Step 3: For each outlier θ, perform $B = 1000$ simulations, i.e., generate $B \times 3$ independent normal distributions $\mathbf{x}_j \sim \mathcal{N}(0,1)$ for all $j = 2,3,4$ with one outlier valued to be $\mathbf{X}_n^o := \theta \mathbf{X}_n$.

Step 4: Generate heteroskedasticity as follows $\mathbf{\Omega}^G = diag(\sqrt{100i})$ and fix a vector $\boldsymbol{\beta}_{ag} = (10,3,-10,58)$ to compute $\mathbf{y} = \mathbf{X}\boldsymbol{\beta}_{ag} + \widetilde{\varepsilon}_g$ with $\varepsilon_{g,i} \sim \mathcal{N}(0,1)$ independent of $\mathbf{x}_j$ and with $\widetilde{\varepsilon}_g = [\mathbf{\Omega}^G]^{-1}\varepsilon_g$.

Step 5: Regress $\mathbf{y}$ on $\mathbf{X}^o$ with the semi-parametric Gini regression and with GLS in order to estimate $\boldsymbol{\beta}_{ag}$. Compute the standard deviation of $\hat{\boldsymbol{\beta}}_{ag}$ by jackknife in the first case, and the standard deviation of the GLS estimators in the second case (for each value of θ). Measure the mean squared error of the coefficient estimates $\hat{\boldsymbol{\beta}}_{ag}$ over B replications (for each θ) for both techniques: Gini and GLS.

The jackknife standard deviations of the estimators $\hat{\beta}_{ag,k}$ for $k = 1,\ldots,4$ are reported in Figure 1 for each value of the outlier θ. As depicted in Figure 1, jackknife standard deviations of the Aitken-Gini estimator are lower than those of the GLS estimator, which are drastically affected by the introduction of one outlier in a sample of $n = 100$ observations. This corresponds to a contamination of the sample of only 1%. Since the outlying observation corresponds to the last row of $\mathbf{X}$ ($\mathbf{X}_n^o = \theta \mathbf{X}_n$), in which there is the most important value of $\mathbf{x}_2$, the vector $\mathbf{x}_2$ is the most contaminated regressor. Therefore, as shown in Figure 1 (top right), important variations of the standard deviation of $\hat{\beta}_{ag,2}$ are recorded, especially for the GLS estimator (red curve) compared with the Gini one (black curve).

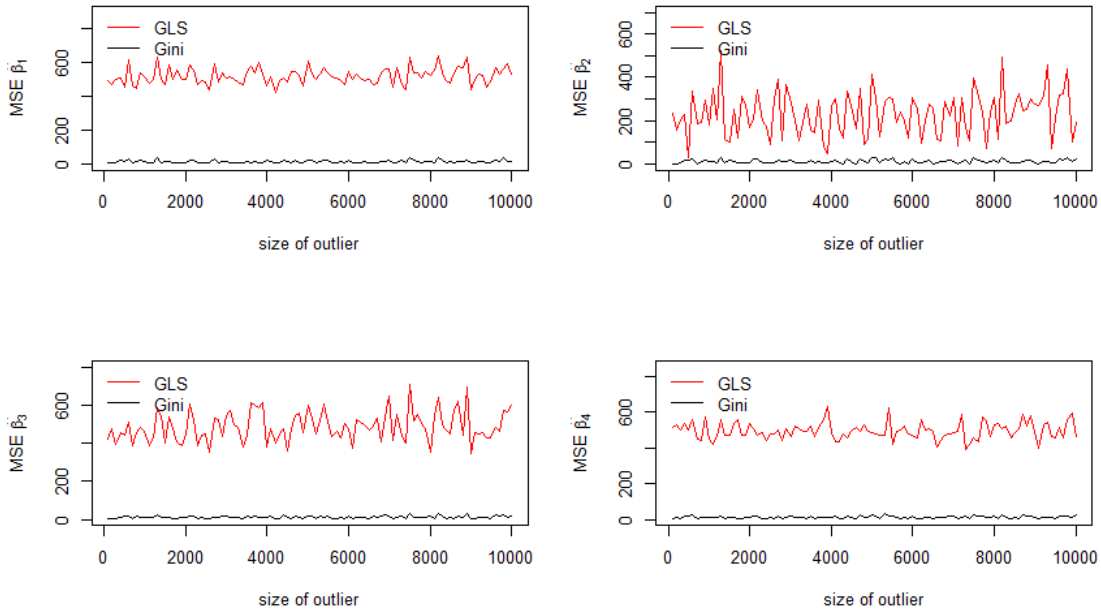

Figure 1. Standard deviations of the coefficients.

In addition, it is possible to compute the mean squared errors of the GLS estimator and the Aitken-Gini one. The contamination process is the same as before.

The Aitken-Gini estimator is better than the usual GLS estimator for the constant of the model (Figure 2, top left) and for the second regressor $\mathbf{x}_2$ (Figure 2, top right), as depicted in Figure 2. This is due to the fact that the outlier is generated by multiplying $\mathbf{X}_n$ by θ, which corresponds to inflating the most important value of $\mathbf{x}_2$ (Step 2). Consequently, the Aitken-Gini estimator $\hat{\beta}_{ag,2}$ yields a robust estimation compared with GLS. For the other cases, the MSE of the generalized least squares estimator are less important.

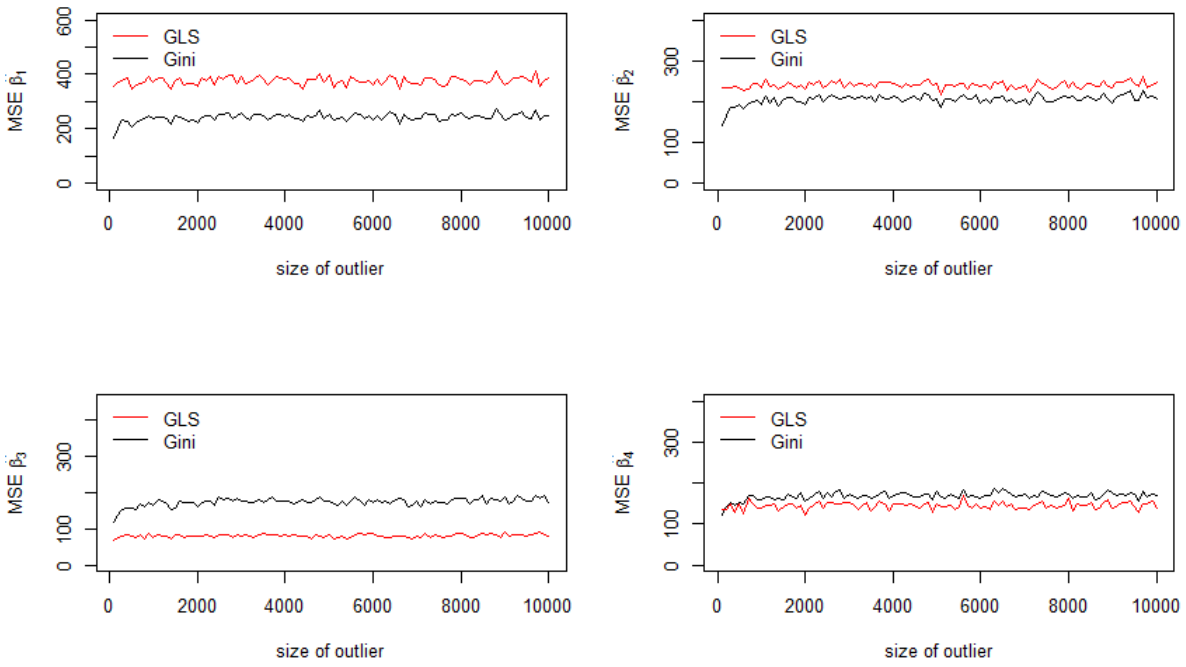

Figure 2. Mean squared errors of the coefficients.

4.2. Tests

The aim of this subsection is to prove that the usual White test for heteroskedasticity has a low power whenever outlying observations arise in the sample, even if the contamination rate is around 1%. Another test is proposed based on the co-Gini operator, and it is shown that a good power may be obtained compared with the standard White test.

Although GLS estimators may be affected by outliers, it is worth mentioning that Aitken-Gini estimators and GLS estimators are based on two different notions of heteroskedasticity. The Aitken-Gini one captures another type of variability, the co-Gini based on ranks, compared with GLS based on the variance. In the following, focus is put on White's test since it is commonly employed in the literature.

White's model and its Gini counterpart are given by,

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \implies \hat{\varepsilon}_i^2 = \delta_0 + \sum_{k=1}^{K} \delta_k x_{ik} + \sum_{k=1}^{K} \gamma_k x_{ik}^2 + u_i \qquad \text{(White-OLS)}$$

and,

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta}_g + \boldsymbol{\varepsilon}_g \implies \underbrace{\hat{\varepsilon}_{g,i}\mathbf{r}_{\varepsilon_{g,i}}}_{\widetilde{y}_i} = \delta_0 + \sum_{k=1}^{K} \delta_k x_{ik} + \sum_{k=1}^{K} \gamma_k x_{ik}\mathbf{r}_{ik} + u_{g,i}, \qquad \text{(White-Gini)}$$

where $\mathbf{r}_{\varepsilon_{g,i}}$ is the rank of $\varepsilon_{g,i}$ and $\mathbf{r}_{ik}$ is the rank of x_{ik} (within the vector $\mathbf{x}_k$). The intuition of the White-Gini test is to exhibit the variables $\mathbf{x}_k$ that depend on the rank of the individuals. This is the case for example when we regress incomes on age. We have the same intuition for White's test performed with OLS. However, the squared residuals and the squared covariates may be inflated because of the outliers. In this respect, it is possible to use Eq.(White-Gini) to test for heteroskedasticity.

It is noteworthy that this equation cannot be estimated by the semi-parametric Gini regression since the rank vector of $\mathbf{x}_k$ and the rank vector of $\mathbf{x}_k \otimes \mathbf{r}_k$ are collinear ($\otimes$ being the Hadamard product). Consequently, both equations are estimated by OLS. The advantage of dealing with Eq.(White-Gini) is to capture the shape of heteroskedasticity in the presence of outliers. The standard White-OLS equation aims at capturing quadratic shapes in the covariates. However, the model fails to achieve this goal in the presence of outliers because outliers are squared. In the White-Gini equation, the product $x_{ik}\mathbf{r}_{ik}$ allows the quadratic shape to be detected, while the intensity of the outliers are attenuated by the role of the rank vector $\mathbf{r}_k$.

In the following tables, we provide the mean R^2 of each model over the number of Monte Carlo experiments $B = 500, 1000, 5000$. We provide in parenthesis the power of the Fisher test related to the significance of the R^2 in each model. The Monte Carlo simulations with contamination were based on the same simulation process described in Algorithm 1.

In Table 1, one observation is contaminated for a sample size $n = 30$, that is 3.33% of the sample. The same generating process was used as in the Monte Carlo simulations performed in the previous section. The outlying observation consists in multiplying the most important value of $\mathbf{x}_2$ by θ. Without outlier, the White-Gini model provides an R^2 of 0.17 (in mean over B) with a very low test power around 2%, whereas the White-OLS model yields an R^2 of 0.26 with a power of 14%. However, when θ is valued to be 100, the White-Gini model performs quite well with an R^2 of 0.45 and a power of 51%, whereas the power decreases slightly in the White-OLS model. In the White-Gini model, thanks to the rank vector of $\mathbf{x}$ as a regressor, the regression curve stays distant from the outlying observation and becomes closer to the other points. Then, the variability of the model explained by the regression curve increases (and then R^2 increases). On the contrary, for the standard model (White-OLS), the regression curve moves toward the outlying observation so that the variability of the residuals increases, and in this case R^2 decreases.

Table 1. Power of the White-Gini test: small sample $n = 30$.

	Without Outliers: $\theta = 1$		**With Outliers: $\theta = 100$**	
$B =$	$\overline{R^2}$ **(White-Gini)**	$\overline{R^2}$ **(White-OLS)**	$\overline{R^2}$ **(White-Gini)**	$\overline{R^2}$ **(White-OLS)**
500	0.17 (0.02)	0.25 (0.11)	0.45 (0.51)	0.24 (0.09)
1000	0.17 (0.03)	0.26 (0.14)	0.45 (0.51)	0.24 (0.09)
5000	0.17 (0.02)	0.26 (0.14)	0.45 (0.50)	0.24 (0.09)

Table 1: Contamination 3.33% of the sample; (), power of the test.

In Table 2, the sample size is $n = 100$ so that the contamination represents 1% of the sample. Without outlier, the White-Gini model provides a very low test power around 7%, compared with 64–70% for White-OLS model. The test power increases to reach 70% in the first case against 59% in the second case.

Table 2. Power of the White-Gini test: $n = 100$.

	Without Outliers: $\theta = 1$		**With Outliers: $\theta = 100$**	
$B =$	$\overline{R^2}$ **(White-Gini)**	$\overline{R^2}$ **(White-OLS)**	$\overline{R^2}$ **(White-Gini)**	$\overline{R^2}$ **(White-OLS)**
500	0.08 (0.07)	0.15 (0.64)	0.32 (0.69)	0.14 (0.59)
1000	0.08 (0.07)	0.15 (0.7)	0.32 (0.7)	0.14 (0.59)
5000	0.08 (0.08)	0.20 (0.65)	0.32 (0.7)	0.14 (0.58)

Table 2: Contamination 1% of the sample; (), power of the test.

Finally, in Table 3, R^2 and test power remain quite equivalent in both models. As mentioned in the literature, for large samples, White-OLS provides an excellent power. When the outliers are dilute in the sample, for instance when the contamination of the sample is only concerned with $0,1\%$ of the sample, because the sample size is large and the number of outliers very low, both tests produce the same power.

Table 3. Power of the White-Gini test: $n = 1000$.

	Without Outliers: $\theta = 1$		**With Outliers: $\theta = 100$**	
$B =$	$\overline{R^2}$ **(White-Gini)**	$\overline{R^2}$ **(White-OLS)**	$\overline{R^2}$ **(White-Gini)**	$\overline{R^2}$ **(White-OLS)**
500	0.05 (1)	0.11 (1)	0.16 (1)	0.11 (1)
1,000	0.05 (1)	0.11 (1)	0.16 (1)	0.11 (1)
5,000	0.05 (1)	0.11 (1)	0.15 (1)	0.11 (1)

Table 3: Contamination 0.1% of the sample; (), power of the test.

As shown in Tables 1–3, when outlying observations affect the sample, the power of the White-Gini test is higher than that of the usual White test.

4.3. The Feasible Generalized Gini Regression

After testing the presence of heteroskedasticity with outlying observations, a new procedure is proposed to estimate $\mathbf{\Omega}$. The so-called feasible generalized least squares (FGLS), introduced by Zellner (1962), was adapted to the Gini regression, the feasible generalized Gini regression (FGGR). Aitken's theorem no longer applies if $\mathbf{\Omega}$ is unknown and must be estimated. The feasible generalized least squares estimator is not the best linear unbiased estimator, nevertheless Kakwani (1967) proved that it is still unbiased under general conditions, and Schmidt (1976) discussed the fact that most of the properties of generalized least squares estimation remain intact in large samples, when plugging in an estimator of $\mathbf{\Omega}$. The form of the heteroskedasticity is unknown, but it can be approximated with a flexible model as,

$$\ln(\mathbf{e}^2) = \mathbf{X}\mathbf{b}_g + \mathbf{u}, \tag{9}$$

in a "Breusch–Pagan" version (in the sense that we consider a linear form of heteroskedasticity), such that $\mathbb{E}(\mathbf{u}) = \mathbf{1}$ and $\mathbf{u}$ independent of $\mathbf{X}$. Instead of using the least squared estimator, the semi-parametric Gini estimator in Equation (2) is employed to deal with contaminated data in $\mathbf{X}$:

$$\hat{\mathbf{b}}_g = (\mathbf{R}'_{\mathbf{x}}\mathbf{X})^{-1}\mathbf{R}'_{\mathbf{x}}\ln(\mathbf{e}^2).$$

Then,

$$\hat{h}_i := \exp(\ln(\hat{\mathbf{e}}_i^2)) = \exp(\mathbf{x}'_i\hat{\mathbf{b}_g}),$$

such that,

$$\hat{\mathbf{\Omega}} := \begin{pmatrix} \hat{h}_1 & 0 & \dots & 0 \\ 0 & \hat{h}_2 & \dots & 0 \\ \vdots & & \ddots & 0 \\ 0 & \dots & \dots & \hat{h}_n \end{pmatrix}.$$

From $\hat{\mathbf{\Omega}}$, we deduce an estimation of $\mathbf{P}$ denoted $\hat{\mathbf{P}} = \hat{\mathbf{\Omega}}^{-\frac{1}{2}}$. Thus, we get that $\hat{\mathbf{X}} := \hat{\mathbf{\Omega}}^{-\frac{1}{2}}\mathbf{X}$. Let $\mathbf{R}_{\hat{\mathbf{x}}}$ be the rank matrix of $\hat{\mathbf{X}}$, hence the FGGR estimator is given by:

$$\hat{\boldsymbol{\beta}}_{FGGR} = (\mathbf{R}'_{\hat{\mathbf{x}}}\hat{\mathbf{X}})^{-1}\mathbf{R}'_{\hat{\mathbf{x}}}\mathbf{y}. \tag{10}$$

On the contrary, the usual FGLS estimator is given by,

$$\hat{\boldsymbol{\beta}}_{FGLS} = (\hat{\mathbf{X}}'\hat{\mathbf{X}})^{-1}\hat{\mathbf{X}}'\mathbf{y},$$

with $\mathbf{b}_g$ estimated by generalized least squares in the first step [Equation (9)]. However, in Proposition 2, it is shown that the Aitken-Gini estimator is based on the $\mathbf{P}$-rank idempotent hypothesis. This assumption states that the rank vector of the residuals must remain invariant after the transformation of the model with respect to matrix $\mathbf{P}$, thereby an Aitken estimator is obtained. However, whenever one outlier occurs in the sample, e.g., the ith row of $\mathbf{X}$ is such that $\mathbf{X}_i^o \to \pm\infty$),

then it comes that $\varepsilon_{g,i} \to \pm\infty$ so that the respect of the **P**-rank idempotent hypothesis is not necessarily ensured for the outlying observation. Consequently, the FGGR estimator is biased. Indeed, the semi-parametric Gini estimator in Equation (10) is based on the rank matrix $\mathbf{R}_{\hat{\mathbf{x}}}$ of $\mathbf{PX}$ which contains errors since $\mathbf{P}$ is computed on the basis of contaminated data. Replacing $\mathbf{R}_{\hat{\mathbf{x}}}$ by $\mathbf{R}_{\mathbf{x}}$, being the rank matrix of $\mathbf{X}$, avoids such a contamination. It is worth mentioning that replacing the rank matrix of the covariates by another rank matrix of some data correlated to the covariates corresponds to the Gini instrumental variable estimator introduced by Yitzhaki and Schechtman (2004). In this respect, the FGGR estimator becomes a feasible generalized Gini regression by instrumental variable (FGGR-IV):

$$\hat{\boldsymbol{\beta}}_{FGGR-IV} = (\mathbf{R}'_{\mathbf{x}}\hat{\mathbf{X}})^{-1}\mathbf{R}'_{\mathbf{x}}\mathbf{y}. \tag{11}$$

Because $\mathbf{R}_{\mathbf{x}}$ is the rank matrix of the initial contaminated data, it comes that the residuals of the transformed model issued from the FGGR-IV estimator are more likely to respect the **P**-rank idempotent hypothesis compared with FGGR based on $\mathbf{PX}$ because both matrices are contaminated.

We performed some Monte Carlo simulations to compare the mean squared errors of the FGGR-IV and FGLS estimators.[4]

Step 1: Generate three independent normal distributions $\mathbf{x}_j \sim \mathcal{N}(0,1)$ of size $n = 100$ for all $j = 2,3,4$ with $\mathbf{x}_1 = (1,\ldots,1)$.

Step 2: As in the previous simulation, sort the matrix $\mathbf{X} = (\mathbf{x}_1,\ldots,\mathbf{x}_4)$ by ascending order according to $\mathbf{x}_2$, except multiply the last row of $\mathbf{X}$ by $\theta = 1,\ldots,100$ (with increments of 1 to avoid problems of matrix invertibility) to inflate the most important value of $\mathbf{x}_2$.

Step 3: For each outlier θ, perform $B = 1000$ simulations, i.e., generate $B \times 3$ independent normal distributions $\mathbf{x}_j \sim \mathcal{N}(0,1)$ for all $j = 2,3,4$ with one outlier valued to be $\mathbf{X}_n^o := \theta\mathbf{X}_n$.

Step 4: Fix a vector $\boldsymbol{\beta}_{ag} = (10,3,-10,58)$ to compute $\mathbf{y} = \beta_{ag,1}\mathbf{x}_1 + \beta_{ag,2}\mathbf{x}_2 + \beta_{ag,3}\mathbf{x}_3 * \mathbf{x}_3 + \beta_{ag,4}\mathbf{x}_4 + \widetilde{\varepsilon}_g$ with $\varepsilon_{g,i} \sim \mathcal{N}(0,1)$, that is, suppose that the heteroskedasticity comes from $\mathbf{x}_3$.

Step 5: Compute the coefficients estimated based on FGGR-IV and FGLS with their MSEs over $B = 1000$ replications.

Figure 3 depicts an interesting correction of heteroskedasticity performed by the FGGR-IV estimator when outliers contaminate only 1% of the sample. The FGGR-IV estimator provides MSEs close to 0 except for the constant and for $\hat{\beta}_{ag,3}$. Because the model has been specified such that $\beta_{ag,3}\mathbf{x}_3 * \mathbf{x}_3$, then the outlier is even more inflated in this case (bottom left in Figure 3).

To show that the FGGR-IV estimator is relevant with other forms of heteroskedasticity, Step 4 was replaced by Step 4', in which an ARCH(1) is modeled (see Figure 4):

Step 4': Fix a vector $\boldsymbol{\beta}_{ag} = (10,3,-10,58)$ to compute $\mathbf{y} = \beta_{ag,1}\mathbf{x}_1 + \beta_{ag,2}\mathbf{x}_2 + \beta_{ag,3}\mathbf{x}_3 + \beta_{ag,4}\mathbf{x}_4 + \widetilde{\varepsilon}_g * \sqrt{\mathbf{1} + 2\epsilon^2}$ with $\varepsilon_{g,i} \sim \mathcal{N}(0,1)$ and $\epsilon_i \sim \mathcal{N}(0,1)$.

The results depicted in Figure 4 are even more clear: the MSEs of all FGGR-IV estimators tend toward 0, consequently the bias of those estimators also tend toward 0.

---

[4] The results of FGGR are not reported because of their bad results compared with FGLS.

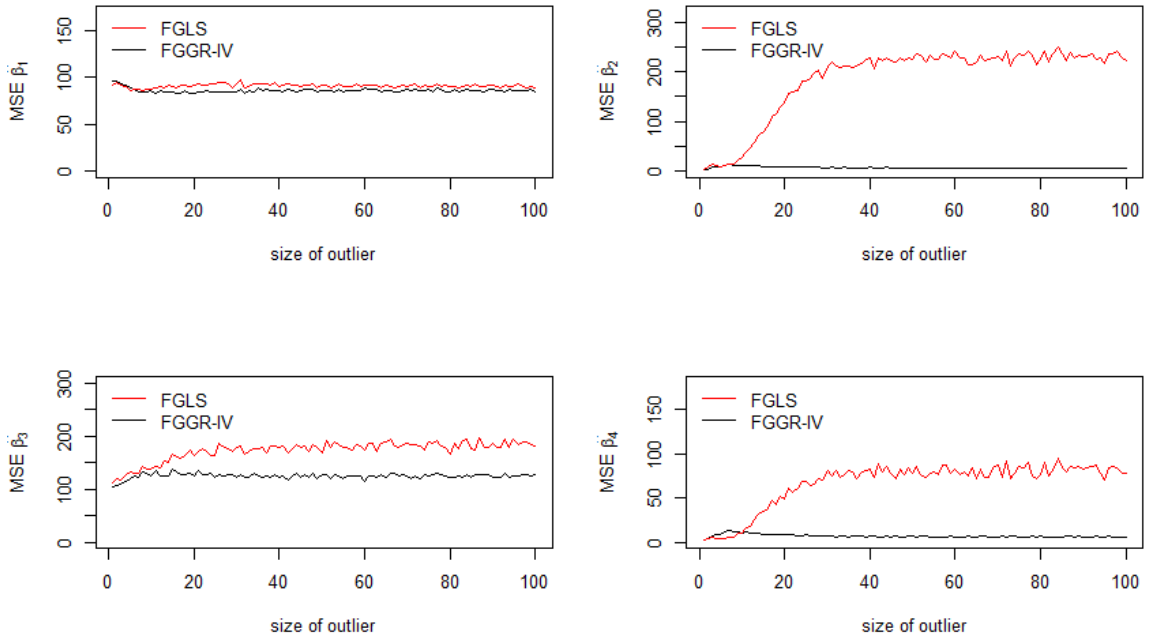

Figure 3. Mean squared errors of the coefficients: FGGR-IV and FGLS.

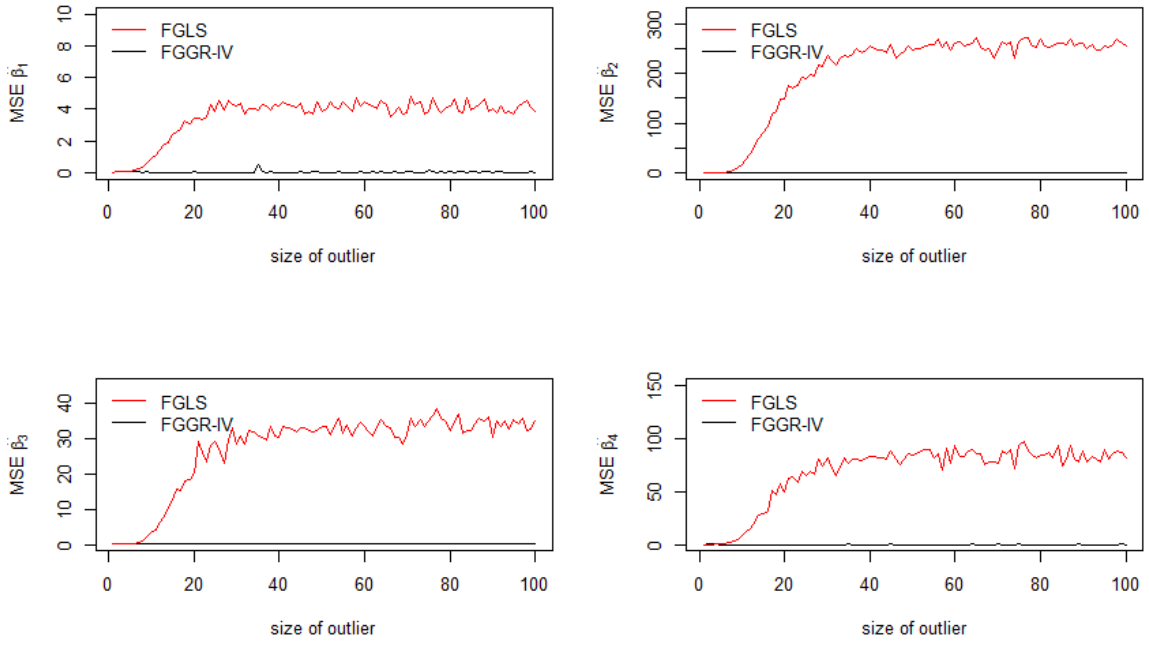

Figure 4. Mean squared errors of the coefficients: FGGR-IV and FGLS with ARCH(1).

5. Concluding Remarks

In this paper, we have demonstrated that two equivalent Gini estimators may be proposed to deal with heteroskedasticity: the former deals with heteroskedasticity in the variance sense and the latter with heteroskedasticity in the Gini sense. The jackknife variance of these estimators are shown to be robust in the presence of outlying observations compared with the usual GLS technique, i.e., the loss of efficiency is less important in the Gini case. The simulations presented in Tables 1–3 show that a contamination of 1% of the sample may drastically affect the power of the White-OLS test, so that the White-Gini test may be preferred to detect the presence of heteroskedasticity when outlying observations occur in the sample.

Author Contributions: Conceptualization, S.M. and N.K.; methodology, A.C.; software, A.C. and S.M.; validation, O.H.N.; writing–original draft preparation, all authors; writing–review and editing, all authors.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Proof of Proposition 6: We follow the proof obtained by Ka and Mussard (2016) in the case of fixed effects panel Gini regressions. Let $\mathbf{r}_k$ be the kth column of $\mathbf{R}_{\mathbf{x}^*}$ and $\mathbf{x}_k^*$ the kth column of $\mathbf{X}^*$, for all $k = 1, \ldots, K$. Let $\hat{\boldsymbol{\beta}}_{ag} =: (\hat{\beta}_{ag1}, \ldots, \hat{\beta}_{agK})$. From Equation (3), we get that:

$$\mathbf{y}^* = \hat{\beta}_{ag1}\mathbf{x}_1^* + \cdots + \hat{\beta}_{agK}\mathbf{x}_K^* + \boldsymbol{\varepsilon}_g.$$

Hence, the following identities hold:[5]

$$\begin{aligned}
\text{cov}(\mathbf{y}^*, \mathbf{r}_1^*) &= \hat{\beta}_{ag1}\text{cov}(\mathbf{x}_1^*, \mathbf{r}_1^*) + \cdots + \hat{\beta}_{agK}\text{cov}(\mathbf{x}_K^*, \mathbf{r}_1^*) + \text{cov}(\boldsymbol{\varepsilon}_g, \mathbf{r}_1^*) \\
&\vdots \\
\text{cov}(\mathbf{y}^*, \mathbf{r}_k^*) &= \hat{\beta}_{ag1}\text{cov}(\mathbf{x}_1^*, \mathbf{r}_k^*) + \cdots + \hat{\beta}_{agK}\text{cov}(\mathbf{x}_K^*, \mathbf{r}_k^*) + \text{cov}(\boldsymbol{\varepsilon}_g, \mathbf{r}_k^*) \\
&\vdots \\
\text{cov}(\mathbf{y}^*, \mathbf{r}_K^*) &= \hat{\beta}_{ag1}\text{cov}(\mathbf{x}_1^*, \mathbf{r}_K^*) + \cdots + \hat{\beta}_{agK}\text{cov}(\mathbf{x}_K^*, \mathbf{r}_K^*) + \text{cov}(\boldsymbol{\varepsilon}_g, \mathbf{r}_K^*).
\end{aligned}$$

Setting $\hat{\beta}_{\varepsilon j}^* := \dfrac{\text{cov}(\boldsymbol{\varepsilon}_g, \mathbf{r}_j^*)}{\text{cov}(\mathbf{x}_j^*, \mathbf{r}_j^*)}$, $\hat{\beta}_{0j}^* := \dfrac{\text{cov}(\mathbf{y}^*, \mathbf{r}_j^*)}{\text{cov}(\mathbf{x}_j^*, \mathbf{r}_j^*)}$ and $\hat{\beta}_{kj}^* := \dfrac{\text{cov}(\mathbf{x}_k^*, \mathbf{r}_j^*)}{\text{cov}(\mathbf{x}_j^*, \mathbf{r}_j^*)}$, and dividing the three last equations by, respectively, $\text{cov}(\mathbf{x}_1^*, \mathbf{r}_1^*)$, $\text{cov}(\mathbf{x}_k^*, \mathbf{r}_k^*)$ and $\text{cov}(\mathbf{x}_K^*, \mathbf{r}_K^*)$ yields:

$$\begin{aligned}
\hat{\beta}_{01}^* &= \hat{\beta}_{ag1} + \cdots + \hat{\beta}_{agK}\hat{\beta}_{K1}^* + \hat{\beta}_{\varepsilon 1}^* \\
&\vdots \\
\hat{\beta}_{0k}^* &= \hat{\beta}_{ag1}\hat{\beta}_{1k} + \cdots + \hat{\beta}_{agK}\hat{\beta}_{Kk}^* + \hat{\beta}_{\varepsilon k}^* \\
&\vdots \\
\hat{\beta}_{0K}^* &= \hat{\beta}_{ag1}\hat{\beta}_{1K} + \cdots + \hat{\beta}_{agK} + \hat{\beta}_{\varepsilon K}^*.
\end{aligned}$$

Now, we define the two following columns vectors $\hat{\mathbf{b}}_0^* := (\hat{\beta}_{01}^*, \ldots, \hat{\beta}_{0K}^*)$ and $\hat{\mathbf{b}}_\varepsilon^* := (\hat{\beta}_{\varepsilon 1}^*, \ldots, \hat{\beta}_{\varepsilon K}^*)$. Then, it comes:

$$\hat{\boldsymbol{\beta}}_{ag} = \begin{pmatrix} 1 & \hat{\beta}_{21}^* & \cdots & \hat{\beta}_{K1}^* \\ \vdots & \vdots & \cdots & \vdots \\ \hat{\beta}_{1K}^* & \hat{\beta}_{2K}^* & \cdots & 1 \end{pmatrix}^{-1} \begin{pmatrix} \hat{\beta}_{01}^* - \hat{\beta}_{\varepsilon 1}^* \\ \vdots \\ \hat{\beta}_{0K}^* - \hat{\beta}_{\varepsilon K}^* \end{pmatrix} =: \hat{\mathbf{B}}^{*-1}\left[\hat{\mathbf{b}}_0^* - \hat{\mathbf{b}}_\varepsilon^*\right] \tag{A1}$$

The previous expression shows that the Aitken-Gini estimator $\hat{\boldsymbol{\beta}}_{ag}$ is a function of slope coefficients of semi-parametric simple Gini regressions $\hat{\beta}_{..}^*$. Consequently, it is a semi-parametric Gini estimator.

[5] This technique was introduced by Yitzhaki and Schechtman (2013, chp. 8) in the case of the standard Gini regression.

Note that $\hat{\beta}^*_{\varepsilon j}$, $\hat{\beta}^*_{0j}$ and $\hat{\beta}^*_{kj}$ are all ratios of U-statistics. Indeed, consider (X, Y) a continuous bivariate distribution with F and G the marginal cumulative distribution functions of X and Y, respectively. By Proposition 9.2 of Yitzhaki and Schechtman (2013), there exists an unbiased and consistent estimator U_a of $4\text{cov}(Y, F(X))$:

$$U_a = 4\binom{n}{m}^{-1} \sum_{i=1}^{n} (2i - 1 - n) y_{x_{(i)}},$$

where $y_{x_{(i)}}$ is the value of y concomitant to the ith order statistic of $x_1, \ldots, x_n$. On the other hand, the U-statistic of $4\text{cov}(X, F(X))$ is (Proposition 9.1 of Yitzhaki and Schechtman 2013):

$$U_b = 4\binom{n}{m}^{-1} \sum_{i=1}^{n} (2i - 1 - n) x_{(i)}.$$

Estimators U_a and U_b are unbiased and consistent estimators of $4\text{cov}(X, G(Y))$ and $4\text{cov}(X, F(X))$, respectively.[6] The estimators $\hat{\beta}^*_{\varepsilon j}$, $\hat{\beta}^*_{0j}$ and $\hat{\beta}^*_{kj}$ are ratios of two dependent U-statistics, such as $U_l := U_a/U_b$. By Slutzky's theorem, because U_a and U_b are consistent estimators, then U_l is also a consistent estimator. By Theorem 10.4 in Yitzhaki and Schechtman (2013), if there exists a real-valued function $g(\theta_1, \ldots, \theta_l)$ of parameters θ_i of the population, and if there exist U-statistics $\mathbf{U} = (U_1, \ldots, U_l)$ corresponding to $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_l)$ such that g and its derivatives are continuous in the neighbourhood of $(\theta_1, \ldots, \theta_l)$, then $\sqrt{n}(g(\mathbf{U}) - g(\boldsymbol{\theta})) \overset{a}{\sim} \mathcal{N}$. Because the estimation of the variance of $g(U_1, \ldots, U_l)$ can be made by jackknife, there is no need to postulate the existence of the second moments $\mathbb{E}\{\phi^2(X_1, \ldots, X_m)\}$. From Equation (A1), since each element $\hat{\beta}_{agk}$ of $\hat{\boldsymbol{\beta}}_{ag}$ is a function of ratios of U-statistics being consistent, then applying Theorem 10.4 in Yitzhaki and Schechtman (2013), it comes that $\hat{\beta}_{agk}$ is a consistent estimator of β_{agk} such that $\hat{\beta}_{agk} \overset{a}{\sim} \mathcal{N}$, for all $k = 1, \ldots, K$.

References

Aitken, Alexander Craig. 1935. On least squares and combinations of observations. *Proceedings of the Royal Society of Edinburgh* 55: 42–48. [CrossRef]

Carcea, Marcel, and Robert Serfling. 2015. A Gini autocovariance function for time series modeling. *Journal of Time Series Analysis* 36: 817–38. [CrossRef]

Ka, Ndéné, and Stéphane Mussard. 2016. ℓ_1 Regressions: Gini estimators for fixed effects panel data. *Journal of Applied Statistics* 43: 1436–46. [CrossRef]

Kakwani, Nanak. 1967. The unbiasedness of Zellner's seemingly unrelated regression equation estimators. *Journal of the American Statistical Association* 82: 141–42. [CrossRef]

Mussard, Stéphane, and Oumar Hamady Ndiaye. 2018. Vector autoregressive models: A Gini approach. *Physica A* 492: 1967–79. [CrossRef]

Olkin, Ingram, and Shlomo Yitzhaki. 1992. Gini regression analysis. *International Statistical Review* 60: 185–96. [CrossRef]

Schechtman, Edna, and Shlomo Yitzhaki. 1987. A Measure of association based on Gin's mean difference. *Communications in Statistics—Theory and Methods* 16: 207–31. [CrossRef]

Schmidt, Peter. 1976. *Econometrics*. New York: Marcel Dekker.

Serfling, Robert J. 1980. *Approximation Theorems of Mathematical Statistics*. New York: John Wiley & Sons.

Shelef, Amit, and Edna Schechtman. 2011. A Gini-based methodology for identifying and analyzing time series with non-normal innovations. *SSNR Electronic Journal* 1–26. [CrossRef]

Yitzhaki, Shlomo, and Edna Schechtman. 2004. The Gini Instrumental Variable or the "double instrumental variable" estimator. *Metron* LXII: 287–313.

[6] Note that, for small samples, both U_a and U_b are biased downward.

Yitzhaki, Shlomo, and Edna Schechtman. 2013. *The Gini Methodology. A Primer on a Statistical Methodology*. New York: Springer.

Zellner, Arnold. 1962. An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American Statistical Association* 57: 348–68. [CrossRef]