Vector autoregressive models : a Gini approach

Stéphane Mussard, Oumar N'Diaye

risques chroniques émergents

CHROME EA 7352

Document de recherche #2017-01

Vector Autoregressive Models: A Gini Approach

Stéphane Mussard
Oumar Ndiaye

UNÎMES UNIVERSITÉ

Vector Autoregressive Models: A Gini Approach

Stéphane Mussard*
Chrome
Université de Nîmes

Oumar Ndiaye†
Chrome
Université de Nîmes

Abstract

In this paper, it is proven that the usual VAR approach may be performed in the Gini sense, that is, on a ℓ_1 metric space. The Gini regression is robust to outliers. As a consequence, when the data are contaminated by extreme values, we show that semi-parametric VAR-Gini regressions may be used to obtain robust estimators. The inference on the estimators is made with the ℓ_1 norm. Also, impulse response functions and Gini decompositions for prevision errors are introduced. Finally, Granger's causality tests are properly derived based on U-statistics.

Keywords: Causality, Gini Regression, Response function, U-statistics, VAR.

*Corresponding author: stephane.mussard@unimes.fr, Research fellow Lameta University of Montpellier, Grédi University of Sherbrooke, and Liser Luxembourg.

†oumar.ndiaye@unimes.fr.

1 Introduction

Vector autoregressive models – VAR here after – have been extensively used thanks to the seminal work of Sims (1980). Instead of considering a single random variable explained with many one-dimensional real independent variables, a vector of random variables is explained by multivariate regressors. This opens the door to inductive approaches for the study of correlation and causality between random variables in a system of equations.

As far as we know, VAR models are based on usual estimators such as generalized least squares, the general method of moments, the maximum likelihood, and many other ones. The main drawback of such techniques is their possible close interrelation to the basic Euclidean distance, *i.e.*, the ℓ_2 metric.

In their innovative works, Yitzhaki and Schechtman (2013) point out the difficulty inherent to the estimators lying in the ℓ_2 metric. Particularly, outliers in samples, even for a small percentage of contamination, imply serious problems on the estimators, for instance explosive variance and/or sign inversions of the coefficient estimates.

Yitzhaki and Schechtman (2013) explain that the use of the coGini operator, instead of the usual covariance, provides robust estimates. They argue that the traditional statistical methods may be performed with the Gini index instead of the variance. We investigate this possibility in order to propose semi-parametric Gini estimators for VAR models, which have not been studied before.

In this paper, we show how it is possible to derive robust estimators from semi-parametric VAR-Gini regressions. We first review the difference between the variance and the Gini index (Section 2). In Section 3, VAR-Gini estimators are proposed: standard semi-parametric Gini regression, generalized Gini regression, and non-linear Gini regressions. The inference on the estimators is made with the ℓ_2 metric, but it is shown that relaxing the strong assumption on the existence of second moments, the inference can be made with the ℓ_1 metric *via* U-statistics (Section 4). Also, impulse response functions and Gini decomposition for prevision errors are investigated (Section 5). Finally, Granger's (non) causality tests are properly derived with U-statistics (Section 6). We close the paper in Section 7.

2 Gini regressions: the standard approaches

We provide in this section a review of the Gini regressions, as it is exposed in Ka and Mussard (2016), see also Yitzhaki and Schechtman (2013).

2.1 The semi-parametric regression

Consider a simple model $y = a + bx$ with x, y some $n \times 1$ vectors. The semi-parametric Gini (simple) regression introduced by Olkin and Yitzhaki (1992), consists in averaging tangents b_{ij} (between observations i and j) with weights v_{ij}. Let the values of x be ranked by ascending order ($x_1 \leqslant \cdots \leqslant x_n$), then the semi-parametric Gini estimator of the slope coefficient is given by:

$$\hat{b}^G = \sum_{i<j} v_{ij} b_{ij}, \text{ with } v_{ij} = \frac{(x_i - x_j)}{\sum_{i<j}(x_i - x_j)} \tag{1}$$

$$\text{and } b_{ij} = \frac{(y_i - y_j)}{(x_i - x_j)} \; \forall i < j \; ; \; i = 1, \ldots, n. \tag{2}$$

Olkin and Yitzhaki (1992) also demonstrate that if the weights v_{ij} are replaced by quadratic ones such as

$$w_{ij} = \frac{(x_i - x_j)^2}{\sum_{i<j}(x_i - x_j)^2}, \tag{3}$$

then the standard Ordinary Least squares (OLS) estimator of the slope coefficient is obtained:

$$\hat{b}^{OLS} = \sum_{i<j} w_{ij} b_{ij}. \tag{4}$$

Since it depends on quadratic weights, the OLS slope coefficient is shown to be heavily sensitive to outliers.

2.2 The parametric regression

The parametric Gini regression (Olkin and Yitzhaki, 1992) solves the minimization of Gini index of the residuals ($e_i = y_i - \hat{y}_i$) and provides the following estimator (only numerically in the multiple regression case):

$$\hat{b}^{PG} = \arg\min_b G(\mathbf{e}) = \arg\min_b \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} |e_i - e_j| \, .$$

Based on all pairwise "city-block" distances, the parametric and non-parametric Gini regressions are equivalent ($\hat{b}^{PG} = \hat{b}^{G}$) if, and only if, the linearity of the model $y = ax + b$ is assessed.

2.3 OLS vs. Gini

The semi-parametric Gini regression may be defined according to the cogini operator, *i.e.* $\text{cog}(y, x) := \text{cov}(y, \mathbf{r}(x))$ and $\text{cog}(x, x) := \text{cov}(x, \mathbf{r}(x))$ where $\mathbf{r}(x)$ is the rank vector of x:[1]

$$\hat{b}^{G} = \frac{\text{cog}(y, x)}{\text{cog}(x, x)} \text{ , whereas } \hat{b}^{OLS} = \frac{\text{cov}(y, x)}{\text{cov}(x, x)}.$$

The semi-parametric Gini multiple regression depends on the rank matrix of the regressors. Let X be the $n \times K$ matrix of the regressors and R_x its rank matrix, which contains in columns the rank vectors $\mathbf{r}(x_k)$ of the regressors x_k for all $k = 1, \ldots, K$. 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, the highest one by n). The semi-parametric Gini multiple regression yields the following estimator (a $K \times 1$ vector):

$$\hat{\mathbf{b}}^{G} = (R_x' X)^{-1} R_x' y \ . \tag{5}$$

The semi-parametric Gini estimator looks like that of instrumental variables in which the instruments are the rank vectors. Durbin (1954) suggested this estimator without being aware that it corresponds to a Gini framework, initiated by Yitzhaki and Schechtman (2004). It is worth mentioning that the cogini index is closed to the Gini coefficient, the so-called Gini Mean Difference:

$$GMD = \mathbb{E}\,|x_i - x_j| = 4\text{cov}(x, F(x)),$$

where $F(x)$ stands for the c.d.f. of the random variable x. Two main approaches have been developed in the literature for analyzing the variability of one random variables. The first one is the variance based on the covariance operator:

$$\sigma^2 = \text{cov}(x, x) = \frac{1}{2}\mathbb{E}\,(x_i - x_j)^2 \ .$$

[1]The rank vector of x (of size $n \times 1$) is obtained by replacing the elements of x by their rank, the smallest value of x is 1 and the highest one is n. For ties in the regressors, the values of the rank vector must be estimated as mid-points, see Yitzhaki and Schechtman (2013, p. 212-213).

The second one is based on the covariance between the c.d.f. of x expressed as $\text{cov}(F(x), F(x))$. This is Spearman's method defined to be the rank method. The cogini operator can be seen as a mixture of the variance and Spearman's pure rank approach. The difference between the the variance and the GMD is the metrics: ℓ_2 and ℓ_1 norm, respectively. Accordingly, the estimator $\hat{\mathbf{b}}^G$ is less sensitive to extreme values thanks to the cogini operator based on the ℓ_1 norm.

2.4 Gini-Grenander conditions

Some existence conditions on the matrix R'_x have to be imposed. Grenander conditions used in OLS regressions are modified in order to get a well-defined Gini regression, see Ka and Mussard (2016).

(i) The first condition postulates that no variable degenerates in a sequence of zero, that is, in the Gini sense:

$$\lim_{n \to +\infty} \mathbf{r}'(x_k) x_k \neq 0, \; k \in \{1, \dots, K\}. \tag{6}$$

(ii) The matrix X must be a full rank matrix, otherwise $R'_x X$ is non invertible. An additional assumption is necessary for Gini regressions, indeed the vectors x_k cannot be comonotonic. Two vectors x and y are comonotonic if, and only if, $\mathbf{r}(x) = \mathbf{r}(y)$. If at least two regressors x_k among $k = 1, \dots, K$ are comonotonic, then $R'_x X$ is non invertible. Let $\mathcal{M}^c$ be the set of all comonotonic matrices with at least two comonotonic vectors x_k. Note that the full rank hypothesis is a necessary condition but it is not sufficient. It is necessary to require in addition the non comonotonicity of the regressors. Then, a condition of identification is:

$$X \notin \mathcal{M}^c \text{ and } X \text{ is a full rank matrix.} \tag{7}$$

(iii) Another condition has to be added with regard to the second moments. As explained in the previous subsection, the Gini estimator does not rely on the second moments of X as this is the case for OLS. Let $\mathbb{E}(x_k^2)$ denotes the second moment of regressor x_k:

$$\text{if } \mathbb{E}(x_k^2) \to \infty \text{ or } \mathbb{E}(x_k^2) = \emptyset \implies \hat{\mathbf{b}}^G \text{ exists.} \tag{8}$$

Finally, the Gini semi-parametric approach relies on a few assumptions:
- no linearity hypothesis is needed ;
- we can have outliers such that $\mathbb{E}(x_k^2) \to \infty$.

3 VAR Gini Regressions: Estimation

Starting with a multivariate process $y_t \in (y_{1t}, \ldots, y_{kt}) \in \mathbb{R}^k$, it is possible to define a vector autoregressive of order 1, *i.e.* $VAR(1)$ as follows:

$$y_t = \alpha + By_{t-1} + \varepsilon_t, \ \forall t = 1, \ldots, T \ ; \tag{9}$$

with $\alpha \in \mathbb{R}^k$ a column vector, B a $k \times k$ matrix of real coefficients, and $\varepsilon_t \sim \mathcal{N}$ a column vector representing the error of the model. A $VAR(p)$ for $p \in \mathbb{N}^*$ ($\mathbb{N}^*$ being the set of positive integer) may be expressed as:

$$y_t = \alpha + \Phi_1 y_{t-1} + \cdots \Phi_p y_{t-p} + \varepsilon_t, \tag{10}$$

with Φ_i being $k \times k$ matrices such that $i = 1, \ldots, p$. For instance, for $k = 2$ and $p = 2$ we get that:

$$\begin{cases} y_{1t} = \alpha_1 + \Phi_{11} y_{1t-1} + \Phi_{12} y_{1t-2} + \Phi_{13} y_{2t-1} + \Phi_{14} y_{2t-2} + \varepsilon_{1t} \\ y_{2t} = \alpha_2 + \Phi_{21} y_{1t-1} + \Phi_{22} y_{1t-2} + \Phi_{23} y_{2t-1} + \Phi_{24} y_{2t-2} + \varepsilon_{2t}, \end{cases} \tag{11}$$

with $\mathbf{\Sigma}_\varepsilon$ the variance-covariance matrix of $(\varepsilon_{1t}, \varepsilon_{2t})$. Or equivalently in a matrix form:

$$\begin{cases} y_1 = \boldsymbol{\beta}_1 X_1 + \varepsilon_{1t} \\ \quad \vdots \\ y_k = \boldsymbol{\beta}_k X_k + \varepsilon_{kt}, \end{cases} \tag{12}$$

with X_i the matrices with the lagged variables $y_{kt-1}, \ldots, y_{kt-p}$.

3.1 The basic semi-parametric case

The Gini estimation of a $VAR(p)$ may be performed by concatenating the equations in the system (12) in order to estimate the following,

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta}_g + \varepsilon_t, \tag{13}$$

with

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \end{pmatrix} \ ; \mathbf{X} = \begin{pmatrix} X_1 & 0 & 0 & \cdots & 0 \\ 0 & X_2 & 0 & \cdots & 0 \\ 0 & \cdots & \cdots & \cdots & \vdots \\ 0 & \cdots & \cdots & \cdots & X_k \end{pmatrix} \ ; \ \boldsymbol{\beta}_g = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix} \tag{14}$$

with $\mathbf{y}$ a $Tk \times 1$ vector, $\mathbf{X}$ a $k \times (kp+1)$ matrix, β a $(kp+1) \times 1$ vector, and ε_t a $Tk \times 1$ white noise. The semi-parametric Gini regression yields an estimator of $\boldsymbol{\beta}_g$:

$$\hat{\boldsymbol{\beta}}_g = (\mathbf{R}_{\mathbf{x}}^{'}\mathbf{X})^{-1}\mathbf{R}_{\mathbf{x}}^{'}\mathbf{y}, \tag{15}$$

where $\mathbf{R}_{\mathbf{x}}$ is the rank matrix of $\mathbf{X}$. As in the standard semi-parametric Gini regression, Gini estimators are of particular relevance when outliers arise in the data.

3.2 Generalized Gini regressions

Let us now propose a new way of estimating VAR models in the Gini sense. Let $\boldsymbol{\Sigma}_\varepsilon$ be the variance-covariance matrix of ε_t. Since it is definite positive, then it exists a matrix $\mathbf{P}$ such that $\boldsymbol{\Sigma}_\varepsilon = \mathbf{P}\mathbf{P}'$. In this respect, thanks to the Cholesky decomposition, the model may be rewritten as:

$$\mathbf{P}^{-1}\mathbf{y} = \mathbf{P}^{-1}\mathbf{X}\boldsymbol{\beta}_g + \mathbf{P}^{-1}\varepsilon_t \iff \mathbf{y}^* = \mathbf{X}^*\boldsymbol{\beta}_g + \boldsymbol{\varepsilon}_t^*. \tag{16}$$

We can check that:

$$\mathbb{E}(\boldsymbol{\varepsilon}_t^*\boldsymbol{\varepsilon}_t^{*'}) = \mathbb{E}(\mathbf{P}^{-1}\varepsilon_t\varepsilon_t'\mathbf{P}^{-1'}) = \mathbf{P}^{-1}\boldsymbol{\Sigma}_\varepsilon\mathbf{P}^{-1'} = \mathbf{P}^{-1}\mathbf{P}\mathbf{P}'\mathbf{P}^{-1'} = \mathbb{I}_{Tk}. \tag{17}$$

Since the previous model has been purged from autocorrelation and heteroskedasticity, then applying the semi-parametric Gini regression we get the generalized Gini regression:

$$\hat{\boldsymbol{\beta}}_{gg} = (\mathbf{R}_{\mathbf{x}^*}'\mathbf{X}^*)^{-1}\mathbf{R}_{\mathbf{x}^*}'\mathbf{y}^*, \tag{18}$$

with $\mathbf{R}_{\mathbf{x}^*}'$ being the rank matrix of $\mathbf{X}^*$. Several special cases of the generalized Gini regression may be derived from the moel exposed above. For instance, if we now assume a contemporaneous correlation between the ε_{kt} and ε_{ht}, that is, $\sigma_{kh} := \text{cov}(\varepsilon_{kt}, \varepsilon_{ht})$:

$$\boldsymbol{\Sigma}_\varepsilon = \mathbb{E}(\varepsilon_t\varepsilon_t') = \begin{pmatrix} \sigma_{11}\mathbb{I}_T & \sigma_{12}\mathbb{I}_T & \cdots & \sigma_{1k}\mathbb{I}_T \\ \sigma_{21}\mathbb{I}_T & \sigma_{22}\mathbb{I}_T & \cdots & \sigma_{2k}\mathbb{I}_T \\ \cdots & \cdots & \cdots & \cdots \\ \sigma_{k1}\mathbb{I}_T & \sigma_{k2}\mathbb{I}_T & \cdots & \sigma_{kk}\mathbb{I}_T \end{pmatrix} \tag{19}$$

with $\mathbb{I}_T$ the $T \times T$ identity matrix, then this model is the well-known SURE regression, that is in our case, a SURE-Gini regression. In the special case where $\sigma_{kh} = 0$, for all $k \neq h$, then

$$\mathbf{\Sigma}_\varepsilon = \mathbb{E}(\varepsilon_t \varepsilon_t') = \begin{pmatrix} \sigma_{11}\mathbb{I}_T & 0 & \cdots & 0 \\ 0 & \sigma_{22}\mathbb{I}_T & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & \sigma_{kk}\mathbb{I}_T \end{pmatrix}. \tag{20}$$

This yields:

$$\mathbf{P}^{-1} := \begin{pmatrix} \sigma_{11}^{-\frac{1}{2}}\mathbb{I}_T & 0 & \cdots & 0 \\ 0 & \sigma_{22}^{-\frac{1}{2}}\mathbb{I}_T & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & \sigma_{kk}^{-\frac{1}{2}}\mathbb{I}_T \end{pmatrix}. \tag{21}$$

Assuming without loss of generality that the data are arranged such that $\sigma_{11} < \cdots < \sigma_{kk}$, then it follows that:

$$\mathbf{R}'_{\mathbf{x}^*} \equiv \mathbf{R}'(\mathbf{X}^*) = \mathbf{R}'(\mathbf{P}^{-1}\mathbf{X}) = \mathbf{R}'(\mathbf{X}) \equiv \mathbf{R}'_{\mathbf{x}}. \tag{22}$$

Therefore, we retrieve the classical semi-parametric Gini estimator:

$$\begin{aligned} \hat{\boldsymbol{\beta}}_{gg} &= (\mathbf{R}'_{\mathbf{x}^*}\mathbf{X}^*)^{-1}\mathbf{R}'_{\mathbf{x}^*}\mathbf{y}^* \\ &= (\mathbf{R}'_{\mathbf{x}}\mathbf{P}^{-1}\mathbf{X})^{-1}\mathbf{R}'_{\mathbf{x}}\mathbf{P}^{-1}\mathbf{y} \\ &= (\mathbf{R}'_{\mathbf{x}}\mathbf{X})^{-1}\mathbf{P}\mathbf{P}^{-1}\mathbf{R}'_{\mathbf{x}}\mathbf{y} \\ &= \hat{\boldsymbol{\beta}}_g. \end{aligned} \tag{23}$$

The same result holds true when the same regressors are used for each equation, *i.e.*,

$$\mathbf{X} = \begin{pmatrix} X & 0 & \cdots & 0 \\ 0 & X & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & X \end{pmatrix}. \tag{24}$$

This yields (21) and thereby, as shown above, $\hat{\boldsymbol{\beta}}_{gg} = \hat{\boldsymbol{\beta}}_g$.

Finally, it can be noticed that the analytic form of $\hat{\boldsymbol{\beta}}_{gg}$ is close to that of instrumental variables (IV) – see Yitzhaki and Schechtman (2004) for the link between Gini regressions and IV. Indeed setting the matrix of instruments $\mathbf{Z}' := \mathbf{R}'_{\mathbf{x}^*}\mathbf{P}^{-1}$, we get an IV estimator:

$$\hat{\boldsymbol{\beta}}_{gg} = (\mathbf{R}'_{\mathbf{x}^*}\mathbf{X}^*)^{-1}\mathbf{R}'_{\mathbf{x}^*}\mathbf{y}^* = (\mathbf{R}'_{\mathbf{x}^*}\mathbf{P}^{-1}\mathbf{X})^{-1}\mathbf{R}'_{\mathbf{x}^*}\mathbf{P}^{-1}\mathbf{y} = (\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{y} \equiv \hat{\boldsymbol{\beta}}_{IV}. \tag{25}$$

Off-course, the estimator $\hat{\boldsymbol{\beta}}_{IV}$ does not rely on the same assumptions compared with the traditional OLS-IV estimator. Thanks to the Gini approach, there is no need to postulate neither exogeneity of the error term, nor the existence of the second moments of the regressors, nor linearity.

3.3 Non linearity

Let us introduce a semi-parametric Gini regression by relaxing the linearity hypothesis. Let us assume the existence of the following regression curve,

$$\mathbf{y} = g(\mathbf{X}, \boldsymbol{\beta}) + \varepsilon_t,$$

that is, $\mathbf{y}$ is a non-linear function of the regressors $\mathbf{X}$ in which extreme values arise. Paralleling the well-known non-linear least squares, we make a Taylor expansion of $g(\mathbf{X}, \boldsymbol{\beta})$ on an exogenous vector $\boldsymbol{\beta}^0$:

$$g(\mathbf{X}, \boldsymbol{\beta}) \approx g(\mathbf{X}, \boldsymbol{\beta}^0) + \sum_k \frac{\partial g(\mathbf{X}, \boldsymbol{\beta}^0)}{\partial \beta_k^0}(\beta_k - \beta_k^0).$$

Thus:

$$\begin{aligned} g(\mathbf{X}, \boldsymbol{\beta}) &\approx \left[g(\mathbf{X}, \boldsymbol{\beta}^0) - \sum_k \beta_k^0 \frac{\partial g(\mathbf{X}, \boldsymbol{\beta}^0)}{\partial \beta_k^0} \right] + \sum_k \beta_k \underbrace{\frac{\partial g(\mathbf{X}, \boldsymbol{\beta}^0)}{\partial \beta_k^0}}_{\mathbf{x}_k^0} \\ &= \left[g^0 - \sum_k \beta_k^0 \mathbf{x}_k^0 \right] + \sum_k \beta_k \mathbf{x}_{k^0} \\ &= g^0 - \mathbf{X}^0\boldsymbol{\beta}^0 + X^0\boldsymbol{\beta}. \end{aligned}$$

The model becomes:

$$\mathbf{y} = g(\mathbf{X}, \boldsymbol{\beta}) + \varepsilon_t = g^0 - \mathbf{X}^0\boldsymbol{\beta}^0 + \mathbf{X}^0\boldsymbol{\beta} + \varepsilon_t.$$

Thus:

$$\underbrace{\mathbf{y} - g^0 + \mathbf{X}^0\boldsymbol{\beta}^0}_{\mathbf{y}^0} = \mathbf{X}^0\boldsymbol{\beta} + \varepsilon_t.$$

Then, we loop the semi-parametric Gini regression (18), where the estimated vector $\hat{\boldsymbol{\beta}}_g^t$ in step t will be the vector used in the next step $t+1$ with $\hat{\boldsymbol{\beta}}_g^0$ the initial vector at step 0 such that:

$$\hat{\boldsymbol{\beta}}_g^{t+1} = [\mathbf{R}'^t_{\mathbf{x}^0}\mathbf{X}^{0t}]^{-1}\mathbf{R}'^t_{\mathbf{x}^0}\mathbf{y}^t, \tag{26}$$

where $\mathbf{R}^t_{\mathbf{x}^0}$ is the rank matrix of $\mathbf{X}^{0t}$ at iteration t. The algorithm stops when $\hat{\boldsymbol{\beta}}_g^{t+1} - \hat{\boldsymbol{\beta}}_g^t \approx 0$.[2]

3.4 Selection of the VAR order

For all semi-parametric Gini regressions studied above, the number of lags p has to be determined. The classical methods may be employed without any problem such as Akaike (AIC), Schwarz (BIC), Hannan et Quinn (HQ), and other ones. For instance, the Akaike information criterion may be computed as,

$$AIC = \log\det(\hat{\boldsymbol{\Sigma}}_\varepsilon) + \frac{2k(kp+1)}{T}, \tag{27}$$

or in the small sample case, see Hurvich and Tsay (1989), as follows

$$AIC = \log\det(\hat{\boldsymbol{\Sigma}}_\varepsilon) + \frac{2k(kp+1)}{T-(kp+1)}. \tag{28}$$

The usual discussion on the choice of the desirable criterion is analyzed for instance in Colletaz (2017).

4 Inference

In this Section it is proven that the inference on the coefficient estimates of Gini-VAR regressions may be done with or without the existence of second moments. Before, let us analyze the existence of the bias.

[2]The same demonstration holds true for generalized Gini regression based on $\mathbf{X}^*$.

4.1 Bias

The estimators studied above are proven to be unbiased. Let us start from the following result. Under the linear approximation $\varepsilon_t = \mathbf{y} - \mathbf{X}\boldsymbol{\beta}_g$ with $\mathbb{E}(\varepsilon_t) = 0$, the estimator $\hat{\boldsymbol{\beta}}_g$ is an unbiased estimator of $\boldsymbol{\beta}_g$,

$$\begin{aligned}\hat{\boldsymbol{\beta}}_g &= (\mathbf{R}_{\mathbf{x}}'\mathbf{X})^{-1}\mathbf{R'}_{\mathbf{x}}\mathbf{y} = (\mathbf{R'}_{\mathbf{x}}\mathbf{X})^{-1}\mathbf{R}_{\mathbf{x}}'(\mathbf{X}\boldsymbol{\beta}_g + \varepsilon_t) \\ &= \boldsymbol{\beta}_g + (\mathbf{R}_{\mathbf{x}}'\mathbf{X})^{-1}\mathbf{R}_{\mathbf{x}}'\varepsilon_t.\end{aligned}$$

Thus,

$$\mathbb{E}(\hat{\boldsymbol{\beta}}_g) = \boldsymbol{\beta}_g + \mathbb{E}((\mathbf{R}_{\mathbf{x}}'\mathbf{X})^{-1}\mathbf{R}_{\mathbf{x}}'\varepsilon_t) = \boldsymbol{\beta}_g. \tag{29}$$

4.2 Existence of second moments

A first possibility of inference is to examine the standard approach, *i.e.*, to consider that second moments of any given distribution exist. Let us assume that the following matrices have constant terms,

$$\begin{aligned}&\text{plim } \frac{1}{T}\mathbf{R}_{\mathbf{x}}'\mathbf{R}_{\mathbf{x}} =: \mathbf{Q}_{\mathbf{RR}}, \ \text{plim } \frac{1}{T}\mathbf{X}'\mathbf{R}_{\mathbf{x}} =: \mathbf{Q}_{\mathbf{xR}} \\ &\text{plim } \frac{1}{T}\mathbf{R}_{\mathbf{x}}'\mathbf{X} =: \mathbf{Q}_{\mathbf{Rx}}, \ \text{plim } \frac{1}{T}\mathbf{R}_{\mathbf{x}}'\varepsilon_t = \mathbf{0},\end{aligned} \tag{30}$$

with $\mathbf{Q}_{\mathbf{Rx}}$ positive definite (ans so invertible).[3] It follows from (29) that:

$$\text{plim}\left[\sqrt{T}\left(\hat{\boldsymbol{\beta}}_g - \boldsymbol{\beta}\right)\right] = \text{plim}\left[\left(\frac{\mathbf{R}_{\mathbf{x}}'\mathbf{X}}{T}\right)\right]^{-1}\left(\frac{1}{\sqrt{T}}\right)\mathbf{R}_{\mathbf{x}}'\varepsilon_t = \mathbf{Q}_{\mathbf{RX}}^{-1}\left(\frac{1}{\sqrt{T}}\right)\mathbf{R}_{\mathbf{x}}'\varepsilon_t.$$

With the hypothesis of i.i.d. variables, with no dominated observations (Grenander conditions), applying the central limit theorem we get,

$$\left(\frac{1}{\sqrt{T}}\right)\mathbf{R}_{\mathbf{x}}'\varepsilon_t \overset{a}{\sim} \mathcal{N}\left(\mathbf{0}, \sigma_g^2\mathbf{Q}_{\mathbf{RR}}\right),$$

with $\sigma_g^2\mathbb{I}_T = \mathbb{E}[\varepsilon_t'\varepsilon_t]$. It then follows that:

$$\left(\frac{\mathbf{R}_{\mathbf{x}}'\mathbf{X}}{T}\right)^{-1}\left(\frac{1}{\sqrt{T}}\right)\mathbf{R}_{\mathbf{x}}'\boldsymbol{\varepsilon}_g \overset{a}{\sim} \mathcal{N}\left[\mathbf{0}, \sigma_g^2\ \mathbf{Q}_{\mathbf{RX}}^{-1}\ \mathbf{Q}_{\mathbf{RR}}\ \mathbf{Q}_{\mathbf{XR}}^{-1}\right],$$

[3] See Schechtman, Yitzhaki and Pudalov (2011) for the demonstration of $\mathbf{R}_{\mathbf{x}}'\varepsilon_t = \mathbf{0}$ in the standard Gini semi-parametric case.

in other terms,

$$\sqrt{T}\left(\hat{\boldsymbol{\beta}}_g - \boldsymbol{\beta}\right) \overset{a}{\sim} \mathcal{N}\left[\mathbf{0}, \sigma_g^2\ \mathbf{Q}_{\mathbf{RX}}^{-1}\ \mathbf{Q}_{\mathbf{RR}}\ \mathbf{Q}_{\mathbf{XR}}^{-1}\right]. \tag{31}$$

Let us prove this result. Under linearity and the hypothesis that $\mathbb{E}(\varepsilon_t'\varepsilon_t) = \sigma_g^2\mathbb{I}_T$, the variance covariance matrix of the semi-parametric Gini estimator is defined as follows:

$$V(\hat{\boldsymbol{\beta}}_g|\mathbf{X}) = \sigma_g^2(\mathbf{R}_{\mathbf{x}}'\mathbf{X})^{-1}(\mathbf{R}_{\mathbf{x}}'\mathbf{R}_{\mathbf{x}})[(\mathbf{R}_{\mathbf{x}}'\mathbf{X})^{-1}]'. \tag{32}$$

Indeed, we get from (29):

$$V(\hat{\boldsymbol{\beta}}_g|\mathbf{X}) = \mathbb{E}[(\hat{\boldsymbol{\beta}}_g - \boldsymbol{\beta}_g)(\hat{\boldsymbol{\beta}}_g - \boldsymbol{\beta}_g)'] = \mathbb{E}[(\mathbf{R}_{\mathbf{x}}'\mathbf{X})^{-1}\mathbf{R}_{\mathbf{x}}'\varepsilon_t\varepsilon_t'\mathbf{R}_{\mathbf{x}}[(\mathbf{R}_{\mathbf{x}}'\mathbf{X})^{-1}]'].$$

Then,

$$V(\hat{\boldsymbol{\beta}}_g|\mathbf{X}) = \sigma_g^2(\mathbf{R}_{\mathbf{x}}'\mathbf{X})^{-1}(\mathbf{R}_{\mathbf{x}}'\mathbf{R}_{\mathbf{x}})[(\mathbf{R}_{\mathbf{x}}'\mathbf{X})^{-1}]'.$$

We proceed exactly in the same way for non linear models depending on the matrix $\mathbf{X}^0$. In this case,

$$\sqrt{T}\left(\hat{\boldsymbol{\beta}}_g - \boldsymbol{\beta}\right) \overset{a}{\sim} \mathcal{N}\left[\mathbf{0}, \sigma_g^2\ \mathbf{Q}_{\mathbf{RX}^0}^{-1}\ \mathbf{Q}_{\mathbf{RR}}\ \mathbf{Q}_{\mathbf{X}^0\mathbf{R}}^{-1}\right]. \tag{33}$$

It is worth mentioning that results (31) and (33) necessitates strong assumptions. If $\mathbf{R}_i$, $\mathbf{X}_i$ and ε_i denotes line i of respectively the matrices $\mathbf{R}$ and $\mathbf{X}$ and of the vector ε_t, then $[\mathbf{R}_i, \mathbf{X}_i, \varepsilon_i]$ must be a sequence of *i.i.d.* random variables. Moreover, the second moments of $\mathbf{X}$ and $\mathbf{R}$ must exist.

On the other hand, for the generalized Gini estimator:

$$\hat{\boldsymbol{\beta}}_{gg} = (\mathbf{R}_{\mathbf{x}^*}'\mathbf{X}^*)^{-1}\mathbf{R}_{\mathbf{x}^*}'\mathbf{y}^* = (\mathbf{R}_{\mathbf{x}^*}'\mathbf{P}^{-1}\mathbf{X})^{-1}\mathbf{R}_{\mathbf{x}^*}'\mathbf{P}^{-1}\mathbf{y},$$

the necessary hypotheses are similar:

$$\begin{aligned}
&\text{plim}\ \frac{1}{T}\mathbf{R}_{\mathbf{x}^*}'\mathbf{R}_{\mathbf{x}^*} =: \mathbf{Q}_{\mathbf{RR}}^*,\ \text{plim}\ \frac{1}{T}\mathbf{X}^{*'}\mathbf{R}_{\mathbf{x}^*} =: \mathbf{Q}_{\mathbf{xR}}^* \\
&\text{plim}\ \frac{1}{T}\mathbf{R}_{\mathbf{x}^*}'\mathbf{X}^* =: \mathbf{Q}_{\mathbf{Rx}}^*,\ \text{plim}\ \frac{1}{T}\mathbf{R}_{\mathbf{x}^*}'\varepsilon_t^* = \mathbf{0}.
\end{aligned} \tag{34}$$

Let,

$$\mathbf{Q}_{\mathbf{xxR}}^* := \text{plim}\ \left[\left(\frac{1}{T}\mathbf{X}^{*'}\mathbf{R}_{\mathbf{x}^*}\right)\left(\frac{1}{T}\mathbf{R}_{\mathbf{x}^*}'\mathbf{R}_{\mathbf{x}^*}\right)^{-1}\left(\frac{1}{T}\mathbf{R}_{\mathbf{x}^*}'\mathbf{X}\right)\right]^{-1}\left(\frac{1}{T}\mathbf{X}^{*'}\mathbf{R}_{\mathbf{x}^*}\right)\left(\frac{1}{T}\mathbf{R}_{\mathbf{x}^*}'\mathbf{R}_{\mathbf{x}^*}\right)^{-1}.$$

The result follows the one issued from generalized least squares with instrumental variables,

$$\hat{\boldsymbol{\beta}}_{gg} \overset{a}{\sim} \mathcal{N}\left[\mathbf{0}, \frac{\sigma_g^2}{T}\mathbf{Q}^*_{\mathbf{xxR}} \text{ plim } \left(\frac{1}{T}\mathbf{R}'_{\mathbf{x}^*}\boldsymbol{\Omega}\mathbf{R}_{\mathbf{x}^*}\right)\mathbf{Q}^{*\prime}_{\mathbf{xxR}}\right], \tag{35}$$

with $\boldsymbol{\Omega}^{-1/2} = \mathbf{P}^{-1}$.[4] Again, this is a crude result since strong assumptions are necessary. Indeed, in addition to the existence of second moments of $\mathbf{X}$ and $\mathbf{R}$, in the case of auto-correlation of the error terms, it is necessary to postulate ergodicity and stationarity in order to get a central limit theorem, see Greene (2003, Chapter 12).

In the semi-parametric case, either for generalized Gini regressions or standard ones, it is also necessary to estimate the asymptotic variance-covariance matrix. Then, we require an estimator of σ_g^2. Let $\mathbf{y} - \mathbf{X}\boldsymbol{\beta}_g = \mathbf{y} - \mathbf{X}(\mathbf{R}'_{\mathbf{x}}\mathbf{X})^{-1}\mathbf{R}'_{\mathbf{x}}\mathbf{y}$. Since $\mathbf{y} = \mathbf{X}\boldsymbol{\beta}_g + \varepsilon_t$, we have $\hat{\varepsilon}_t = [\mathbb{I} - \mathbf{X}(\mathbf{R}'_{\mathbf{x}}\mathbf{X})^{-1}\mathbf{R}_{\mathbf{x}}]\,\varepsilon_t$. Thus:

$$\begin{aligned}
asym.\hat{\sigma}_g^2 &= \frac{\hat{\varepsilon}'_t\hat{\varepsilon}_t}{T} \\
&= \frac{\varepsilon'_t\varepsilon_t}{T} + \left(\frac{\varepsilon'_t\mathbf{R}_{\mathbf{x}}}{T}\right)\left(\frac{\mathbf{X}'\mathbf{R}_{\mathbf{x}}}{T}\right)^{-1}\left(\frac{\mathbf{X}'\mathbf{X}}{T}\right)\left(\frac{\mathbf{R}'_{\mathbf{x}}\mathbf{X}}{T}\right)^{-1}\left(\frac{\mathbf{R}'_{\mathbf{x}}\varepsilon_t}{T}\right) \\
&\quad - 2\left(\frac{\varepsilon'_t\mathbf{R}_{\mathbf{x}}}{T}\right)\left(\frac{\mathbf{X}'\mathbf{R}_{\mathbf{x}}}{T}\right)^{-1}\left(\frac{\mathbf{R}'_{\mathbf{x}}\varepsilon_t}{T}\right).
\end{aligned}$$

Since the semi-parametric regression implies $\mathbf{R}'_{\mathbf{x}}\varepsilon_t = \mathbf{0}$, then:

$$\text{plim } asym.\hat{\sigma}_g^2 = \frac{\varepsilon'_t\varepsilon_t}{T} = \sigma_g^2.$$

For small samples, it can be shown, in the same manner than in the OLS case, that:

$$\hat{\sigma}_g^2 = \frac{\hat{\varepsilon}'_t\hat{\varepsilon}_t}{T-K-1}.$$

4.3 Existence of first moments only

As shown in the previous subsection, some strong conditions must be imposed such as the existence of second moments in order to get the limiting distributions of $\hat{\boldsymbol{\beta}}_{gg}$. Yitzhaki and Schechtman (2013) show that all the estimators

[4] The matrix $\boldsymbol{\Omega}$ is the matrix issued from the singular value decomposition of $\boldsymbol{\Omega}$ with $\boldsymbol{\Sigma}_\varepsilon = \sigma_g^2\boldsymbol{\Omega}$. In this case, $\boldsymbol{\Omega} = \mathbf{C}\boldsymbol{\Lambda}\mathbf{C}'$ with $\boldsymbol{\Omega}^{-1} = \mathbf{P}'\mathbf{P}$, thus $\mathbf{P}' = \mathbf{C}^{-1'}\boldsymbol{\Lambda}^{-\frac{1}{2}} = \boldsymbol{\Omega}^{-\frac{1}{2}}$.

used in semi-parametric Gini regressions are U-statistics, which possess desirable and weaker asymptotic properties. Let us show that $\hat{\boldsymbol{\beta}}_{gg}$ is a function of U-statistics.

Let $\mathbf{x}_1$, $\mathbf{x}_2, \ldots, \mathbf{x}_m$ be m i.i.d. random variables, and $\phi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m)$ a symmetric function (the kernel) such that:

$$\phi^*(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m) = (n!)^{-1} \sum_{i_1, i_2, \ldots, i_m} \ldots \sum \phi(\mathbf{x}_{i_1}, \mathbf{x}_{i_2}, \ldots, \mathbf{x}_{i_m}),$$

where n is the smallest number of observations needed to estimate ϕ^*. The U-statistic for the parameter ϕ^*, which is an unbiased estimate of ϕ^*, is expressed as:

$$U(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m) = \binom{m}{n}^{-1} \sum_{i_1, i_2, \ldots, i_m} \ldots \sum \phi(\mathbf{x}_{i_1}, \mathbf{x}_{i_2}, \ldots, \mathbf{x}_{i_m}).$$

The variance of an U-statistics, $Var(U)$, for the parameter ϕ^* of degree m (degree of the kernel) is,

$$Var(U) = \binom{n}{m}^{-1} \sum_{i=1}^{n} \binom{n}{i} \binom{n-m}{m-i} \xi_i,$$

where,

$$\xi_i = Var[\phi_i^*(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m)] = \mathbb{E}(\phi_i^{*2}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m)) - \mathbb{E}(\phi_i^*(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m))^2.$$

An easier way to estimate the variance of U is the jackknife:

$$Var(U) = \frac{n-1}{n} \sum_{i=1}^{n} \left[U_{-i} - \frac{1}{n} \sum_{i=1}^{n} U_{-i} \right]^2, \tag{36}$$

where U_{-i} is the estimator issued from the sample of size $n-1$ *i.e.* without the ith observation.

In order to prove that $\hat{\boldsymbol{\beta}}_{gg}$ is a semi-parametric estimator, which is a function of U-statistics, we start from the estimation of the system (12):

$$\begin{cases} y_1 = \hat{\boldsymbol{\beta}}_{gg1} X_1 + \hat{\varepsilon}_{1t} \\ \quad \vdots \\ y_k = \hat{\boldsymbol{\beta}}_{ggk} X_k + \hat{\varepsilon}_{kt}, \end{cases} \tag{37}$$

For each equation of the system, we can prove that $\hat{\boldsymbol{\beta}}_{ggk}$ is an U-statistics. Let us start with the k-th equation, which as been estimated *via* generalized Gini regression, then for p lags and r regressors, we have:

$$y_{kt} = \hat{\beta}_{ggk,1} y_{1t-1} + \cdots + \hat{\beta}_{ggk,\ell} y_{rt-p} + \hat{\varepsilon}_{kt}. \tag{38}$$

We follow a technique initiated by Yitzhaki and Schechtman (2013, Chapter 8) for the standard semi-parametric Gini regression. We set the following identity issued from (38):

$$\begin{aligned}
\text{cov}(y_{kt}, \mathbf{r}_1) &= \hat{\beta}_{ggk,1}\text{cov}(y_{1t-1}, \mathbf{r}_1) + \cdots + \hat{\beta}_{ggk,\ell}\text{cov}(y_{rt-p}, \mathbf{r}_1) + \text{cov}(\hat{\varepsilon}_{kt}, \mathbf{r}_1)\\
\text{cov}(y_{kt}, \mathbf{r}_2) &= \hat{\beta}_{ggk,1}\text{cov}(y_{1t-1}, \mathbf{r}_2) + \cdots + \hat{\beta}_{ggk,\ell}\text{cov}(y_{rt-p}, \mathbf{r}_2) + \text{cov}(\hat{\varepsilon}_{kt}, \mathbf{r}_2)\\
&\vdots\\
\text{cov}(y_{kt}, \mathbf{r}_\ell) &= \hat{\beta}_{ggk,1}\text{cov}(y_{1t-1}, \mathbf{r}_\ell) + \cdots + \hat{\beta}_{ggk,\ell}\text{cov}(y_{rt-p}, \mathbf{r}_\ell) + \text{cov}(\hat{\varepsilon}_{kt}, \mathbf{r}_\ell),
\end{aligned}$$

with $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_\ell$ the rank vectors of, respectively, $y_{1t-1}, y_{1t-2}, \ldots, y_{rt-p}$. Dividing the previous equations by, respectively, $\text{cov}(y_{1t-1}, \mathbf{r}_1), \ldots, \text{cov}(y_{rt-p}, \mathbf{r}_\ell)$ yields:

$$\begin{aligned}
\hat{\beta}_{01} &= \hat{\beta}_{ggk,1} + \cdots + \hat{\beta}_{ggk,\ell}\hat{\beta}_{\ell 1} + \hat{\beta}_{\varepsilon 1}\\
\hat{\beta}_{02} &= \hat{\beta}_{ggk,2} + \cdots + \hat{\beta}_{ggk,\ell}\hat{\beta}_{\ell 2} + \hat{\beta}_{\varepsilon 2}\\
&\vdots\\
\hat{\beta}_{0\ell} &= \hat{\beta}_{ggk,1}\hat{\beta}_{1\ell} + \cdots + \hat{\beta}_{ggk,\ell} + \hat{\beta}_{\varepsilon\ell},
\end{aligned} \tag{39}$$

with, for $j = 1, \ldots, \ell$,

$$\hat{\beta}_{0j} := \frac{\text{cov}(y_{kt}, \mathbf{r}_j)}{\text{cov}(y_{jt-j}, \mathbf{r}_j)} \; ; \; \hat{\beta}_{\ell j} := \frac{\text{cov}(y_{rt-p}, \mathbf{r}_j)}{\text{cov}(y_{jt-j}, \mathbf{r}_j)} \; ; \; \hat{\beta}_{\varepsilon j} := \frac{\text{cov}(\hat{\varepsilon}_{kt}, \mathbf{r}_j)}{\text{cov}(y_{jt-j}, \mathbf{r}_j)}. \tag{40}$$

Now, let us set the vectors $\hat{\mathbf{b}}_0 := (\hat{\beta}_{01}, \ldots, \hat{\beta}_{0\ell})$ and $\hat{\mathbf{b}}_\varepsilon := (\hat{\beta}_{\varepsilon 1}, \ldots, \hat{\beta}_{\varepsilon\ell})$, then we get,

$$\begin{pmatrix} \hat{\beta}_{ggk,1} \\ \vdots \\ \hat{\beta}_{ggk,\ell} \end{pmatrix} = \begin{pmatrix} 1 & \hat{\beta}_{21} & \ldots & \hat{\beta}_{\ell 1} \\ \vdots & \vdots & \ldots & \vdots \\ \hat{\beta}_{1\ell} & \hat{\beta}_{2\ell} & \ldots & 1 \end{pmatrix}^{-1} \begin{pmatrix} \hat{\beta}_{01} - \hat{\beta}_{\varepsilon 1} \\ \vdots \\ \hat{\beta}_{0\ell} - \hat{\beta}_{\varepsilon\ell} \end{pmatrix} \tag{41}$$

$$\iff \hat{\boldsymbol{\beta}}_{gg} = \hat{\mathbf{B}}^{-1}\left[\hat{\mathbf{b}}_0 - \hat{\mathbf{b}}_\varepsilon\right].$$

The estimator $\hat{\boldsymbol{\beta}}_{gg}$ is a function of slope coefficients of semi-parametric simple Gini regressions, see (1). Thereby, it is a semi-parametric Gini estimator. The estimators $\hat{\beta}_{0j}$, $\hat{\beta}_{\varepsilon j}$ and $\hat{\beta}_{\ell j}$ are function of U-statistics, see Yitzhaki and Schechtman (2013, Chapter 9). If $\hat{\mathbf{B}}$ is a full rank matrix, so invertible, then $\hat{\boldsymbol{\beta}}_{gg}$ is a function of U-statistics. By Slutzky's theorem, $\hat{\boldsymbol{\beta}}_{gg}$ is a consistent estimator and it is asymptotically normal:

$$\hat{\boldsymbol{\beta}}_{gg} \overset{a}{\sim} \mathcal{N}(\boldsymbol{\beta}_{gg}, \hat{\sigma}_J(\hat{\boldsymbol{\beta}}_{gg})), \tag{42}$$

with $\hat{\sigma}_J(\hat{\boldsymbol{\beta}}_{gg})$ the Jackknife standard deviation estimated with (36).

4.4 Testing for linearity

As mentioned by Yitzhaki and Schechtman (2013), in the case of cross sectional data, the parametric Gini regression may be used to test for the linearity of the model. It is worth mentioning that the parametric Gini regression relies on linearity. Then, it is a tool to test whether the coefficients of the semi-parametric Gini regressions are coincident with those derived under linearity. Let $\boldsymbol{\beta}_{pg}$ and $\boldsymbol{\beta}_{gg}$ be respectively the parametric and semi-parametric Gini coefficients of any given $VAR(p)$ model. Then, for any given regressor k the test is the following,

$$\left\| \begin{array}{l} H_0 : \beta_{pgk} - \beta_{ggk} = 0 \ \forall k \\ H_1 : \beta_{pgk} - \beta_{ggk} \neq 0 \ \forall k \end{array} \right. \iff \left\| \begin{array}{l} H_0 : \text{linearity of the model} \\ H_1 : \text{non linearity of the model.} \end{array} \right.$$

From (42) we know that $\boldsymbol{\beta}_{gg} \overset{a}{\sim} \mathcal{N}$. On the other hand, if $\boldsymbol{\beta}_{pg} \overset{a}{\sim} \mathcal{N}$, then

$$\hat{\beta}_{ggk} - \hat{\beta}_{pgk} \overset{a}{\sim} \mathcal{N}(\beta_{ggk} - \beta_{pgk}, \hat{\sigma}_J(\hat{\beta}_{ggk} - \beta_{pgk})), \tag{43}$$

with $\hat{\sigma}_J(\hat{\beta}_{ggk} - \beta_{pgk})$ the Jackknife standard deviation estimated with (36). If the null is rejected, then the semi-parametric approach provides a regression curve. Also, in this case, it is possible to define a specific non-linear form and use the semi-parametric non-linear Gini regression studied above (26).

5 Impulse Response and Gini Decomposition

In this section, we study the impulse response function of the Gini VAR model. First, we briefly review the condition of stationarity necessary to derive the impulse response function. Then, we develop the Gini impulse response function and the Gini decomposition of the errors.

5.1 Stationarity

It is well-known that any $VAR(p)$ may be reduced to a $VAR(1)$ as follows,

$$Y_t = \mathbf{c} + \mathbf{C}Y_{t-1} + \varepsilon_t, \tag{44}$$

such that,

$$Y_t = \begin{pmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-p+1} \end{pmatrix} \; ; \; \mathbf{c} = \begin{pmatrix} c \\ 0 \\ \vdots \\ 0 \end{pmatrix} \text{ and } \mathbf{C} = \begin{pmatrix} \Phi_1 & \Phi_2 & \cdots & \Phi_p \\ \mathbb{I}_n & 0 & \cdots & 0 \\ 0 & \mathbb{I}_n & 0 & \cdots \\ \vdots & \cdots & \cdots & \vdots \\ 0 & \cdots & \mathbb{I}_n & 0 \end{pmatrix}. \tag{45}$$

As a consequence, if a $VAR(p)$ is stationary or stable then its reduced form (44) is also stable. It is noteworthy that a $VAR(p)$ is stationary if, and only if, the eigen values of $\mathbf{C}$ are no greater than one in absolute values. Or, equivalently, defining L the lag operator, then stationarity implies that the polynomial,

$$|\mathbb{I} - \Phi_1 L - \Phi_2 L^2 - \cdots \Phi_p L^p|, \tag{46}$$

has roots outside the complex circle of radius one. On the other hand, it is well-known that any $VAR(p)$ may be rewritten as a Vector Mooving Average process VMA. An infinite vector mooving average process, $VMA(\infty)$, is expressed as follows,

$$y_t = \mu + \varepsilon_t + \Psi_1 \varepsilon_{t-1} + \Psi_2 \varepsilon_{t-2} + \cdots = \mu + \Psi(L)\varepsilon_t, \tag{47}$$

with $\mu := \mathbb{E}(y_t)$ and $\varepsilon_{t-i} \sim \mathcal{N}$. In this case, the coefficients of VAR and VMA are linked in the following way,

$$(\mathbb{I} - \Phi_1 L - \Phi_2 L^2 - \cdots \Phi_p L^p)(\Psi_0 + \Psi_1 L + \Psi_2 L^2 + \cdots) = \mathbb{I}. \tag{48}$$

This provides a well-known result in the literature of VAR models:

$$\Psi_0 = \mathbb{I} \tag{49}$$

$$\Psi_1 = \Phi_1 \tag{50}$$

$$\Psi_2 = \Phi_1 \Psi_1 + \Phi_2 \tag{51}$$

$$\vdots$$

$$\Psi_i = \Phi_1 \Psi_{i-1} + \Phi_2 \Psi_{i-2} + \cdots + \Phi_p \Psi_{i-p}. \tag{52}$$

5.2 The simple impulse response function

Thanks to the $VMA(\infty)$ model, it is possible to capture the impact of any given innovation ε_{kt} on the whole system y_t. Since the $VMA(\infty)$ is given by,

$$y_t = \mu + \varepsilon_t + \Psi_1 \varepsilon_{t-1} + \Psi_2 \varepsilon_{t-2} + \cdots + \Psi_i \varepsilon_{t-i} + \cdots \tag{53}$$

then, the simple response function is given by,

$$\frac{\partial y_t}{\partial \varepsilon_{t-i}} = \Psi_i. \tag{54}$$

Since Ψ_i is a $k \times k$ matrix, the simple impulse response function provides the impact at time t of the innovations $\varepsilon_1, \ldots, \varepsilon_k$ on the dependent variable y_{t+i}, precisely on each $y_{1t+i}, \ldots, y_{kt+i}$.

5.3 The orthogonal impulse response function

As seen in the previous sections, the generalized Gini regression provides an estimator $\hat{\boldsymbol{\beta}}_{gg}$ based on the variance-covariance matrix $\mathbf{\Sigma}_\varepsilon$:

$$\mathbf{\Sigma}_\varepsilon := \mathbb{E}(\varepsilon_t \varepsilon_t') = \begin{pmatrix} \sigma_{11}\mathbb{I}_T & \sigma_{12}\mathbb{I}_T & \cdots & \sigma_{1k}\mathbb{I}_T \\ \sigma_{21}\mathbb{I}_T & \sigma_{22}\mathbb{I}_T & \cdots & \sigma_{2k}\mathbb{I}_T \\ \cdots & \cdots & \cdots & \cdots \\ \sigma_{k1}\mathbb{I}_T & \sigma_{k2}\mathbb{I}_T & \cdots & \sigma_{kk}\mathbb{I}_T \end{pmatrix}. \tag{55}$$

As a consequence, it is of interest to deal with orthogonal innovations in order to get proper estimates of impulse responses.

For that purpose, the Cholesky decomposition is used in the same manner than in the generalized Gini estimation (16). The matrix $\mathbf{P}^{-1}$ enables the the errors $\mathbf{P}^{-1}\varepsilon_t$ to be neutral in the sense that,

$$\mathbb{E}(\boldsymbol{\varepsilon}_t^* \boldsymbol{\varepsilon}_t^{*'}) = \mathbb{E}(\mathbf{P}^{-1}\varepsilon_t \varepsilon_t' \mathbf{P}^{-1'}) = \mathbf{P}^{-1}\mathbf{\Sigma}_\varepsilon \mathbf{P}^{-1'} = \mathbf{P}^{-1}\mathbf{P}\mathbf{P}'\mathbf{P}^{-1'} = \mathbb{I}_{Tk}. \tag{56}$$

Consequently, the corrected innovations $\boldsymbol{\varepsilon}_t^* = \mathbf{P}^{-1}\varepsilon_t$ allows for a better estimation of the responses since they are uncorrelated and orthogonal to each others. Since $\mathbf{P}\boldsymbol{\varepsilon}_t^* = \varepsilon_t$ the $VMA(\infty)$ is rewritten as:

$$y_t = \mu + \varepsilon_t + \Psi_1 \mathbf{P}\boldsymbol{\varepsilon}_{t-1}^* + \Psi_2 \mathbf{P}\boldsymbol{\varepsilon}_{t-2}^* + \cdots + \Psi_i \mathbf{P}\boldsymbol{\varepsilon}_{t-i}^* + \cdots \tag{57}$$

Setting $\Theta_i := \Psi_i \mathbf{P}$, then

$$y_t = \mu + \varepsilon_t + \Theta_1 \boldsymbol{\varepsilon}_{t-1}^* + \Theta_2 \boldsymbol{\varepsilon}_{t-2}^* + \cdots + \Theta_i \boldsymbol{\varepsilon}_{t-i}^* + \cdots \tag{58}$$

The orthogonal response function is given by:

$$\frac{\partial y_t}{\partial \boldsymbol{\varepsilon}^*_{t-i}} = \Theta_i. \tag{59}$$

The previous approaches are quite standard. However in the case of outliers in y_t, the standard VAR models would overestimate the innovations and the same thing for the impulse response functions. In our approach, since the estimator $\hat{\boldsymbol{\beta}}_{gg}$ is robust to outliers then robustness is also ensured for the impulse response functions, which depend on $\hat{\boldsymbol{\beta}}_{gg}$.

It is also possible to robustify the variance-covariance matrix by the use of the coGini matrix:

$$\boldsymbol{G}_\varepsilon := \mathbb{E}(\varepsilon_t F'(\varepsilon_t)) = \begin{pmatrix} \text{cog}_{11}\mathbb{I}_T & \text{cog}_{12}\mathbb{I}_T & \cdots & \text{cog}_{1k}\mathbb{I}_T \\ \text{cog}_{21}\mathbb{I}_T & \sigma_{22}\mathbb{I}_T & \cdots & \text{cog}_{2k}\mathbb{I}_T \\ \cdots & \cdots & \cdots & \cdots \\ \text{cog}_{k1}\mathbb{I}_T & \text{cog}_{k2}\mathbb{I}_T & \cdots & \text{cog}_{kk}\mathbb{I}_T \end{pmatrix} \tag{60}$$

where F is the c.d.f. of ε_t, and where,

$$\text{cog}_{ij} = \text{cov}(\varepsilon_{it}, F(\varepsilon_{jt})). \tag{61}$$

In this case, when outliers occur in y_y, even if the semi-parametric Gini regression attenuates the presence of outliers in the errors terms, the use of $\boldsymbol{G}_\varepsilon$ enables to get a higher degree of robustness compared with $\boldsymbol{\Sigma}_\varepsilon$ thanks to the ℓ_1 metric. In this case, the Cholesky decomposition applied to $\boldsymbol{G}_\varepsilon$ yields an impulse response function in the Gini sense,

$$\frac{\partial y_t}{\partial \boldsymbol{\varepsilon}^*_{t-i}} = \tilde{\Theta}_i, \tag{62}$$

with $\tilde{\Theta}_i = \Psi_i \tilde{\mathbf{P}}$ and with $\tilde{\mathbf{P}}$ issued from the Cholesky decomposition of $\boldsymbol{G}_\varepsilon$.

5.4 The Gini decomposition of the errors

From the $VMA(\infty)$ it is possible to derive the prevision error at period $t+\ell$. For variable y_k, this error is, for lags $l = 1, \ldots, \ell$:

$$\begin{aligned} er_{kt+\ell} &= y_{kt+\ell} - \mathbb{E}(y_{kt+\ell}) \\ &= \sum_{l=1}^{\ell} \Theta_{\ell-l,k1}\varepsilon^*_{1,t+1} + \sum_{l=1}^{\ell} \Theta_{\ell-l,k2}\varepsilon^*_{2,t+1} + \cdots + \sum_{l=1}^{\ell} \Theta_{\ell-l,kk}\varepsilon^*_{k,t+1}, \end{aligned} \tag{63}$$

where $\Theta_{\ell-l,k1}$ is the coefficient estimate of variable k associated with innovation ε_1^*. The variance of the previous equation is:

$$\mathrm{Var}(er_{kt+\ell}) = \sum_{l=1}^{\ell} \Theta_{\ell-l,k1}^2 + \cdots + \sum_{l=1}^{\ell} \Theta_{\ell-l,kk}^2. \tag{64}$$

The contribution of innovation ε_k^* to the variance of the prevision error is:

$$\mathrm{Cont}_k = \frac{\sum_{l=1}^{\ell} \Theta_{\ell-l,kk}^2}{\sum_{l=1}^{\ell} \Theta_{\ell-l,k1}^2 + \cdots + \sum_{l=1}^{\ell} \Theta_{\ell-l,kk}^2}. \tag{65}$$

For the Gini view, we have to compute the coGini between the prevision errors and its rank vector $\mathbf{r}_{er}$:

$$\mathrm{cov}(er_{kt+\ell}, \mathbf{r}_{er}) = \sum_{l=1}^{\ell} \Theta_{\ell-l,k1} \mathrm{cov}(\varepsilon_{1,t+1}^*, \mathbf{r}_{er}) + \cdots + \sum_{l=1}^{\ell} \Theta_{\ell-l,kk} \mathrm{cov}(\varepsilon_{k,t+1}^*, \mathbf{r}_{er}). \tag{66}$$

Multiplying both sides of the previous expression by $4/n$ yields the Gini index of the prevision error $G(er_{kt+\ell})$:

$$G(er_{kt+\ell}) = \frac{4}{n} \sum_{l=1}^{\ell} \Theta_{\ell-l,k1} \mathrm{cov}(\varepsilon_{1,t+1}^*, \mathbf{r}_{er}) + \cdots + \frac{4}{n} \sum_{l=1}^{\ell} \Theta_{\ell-l,kk} \mathrm{cov}(\varepsilon_{k,t+1}^*, \mathbf{r}_{er}). \tag{67}$$

The contribution of innovation ε_k^* to the Gini prevision error is given by:

$$\text{Cont-}G_k = \frac{\sum_{l=1}^{\ell} \Theta_{\ell-l,kk} \mathrm{cov}(\varepsilon_{k,t+1}^*, \mathbf{r}_{er})}{\sum_{l=1}^{\ell} \Theta_{\ell-l,k1} \mathrm{cov}(\varepsilon_{1,t+1}^*, \mathbf{r}_{er}) + \cdots + \sum_{l=1}^{\ell} \Theta_{\ell-l,kk} \mathrm{cov}(\varepsilon_{k,t+1}^*, \mathbf{r}_{er})}. \tag{68}$$

6 Granger Causality Test

In order to test whether a variable (does not) cause(s) y_t, in Granger's sense, let us start with a $VAR(1)$ model:

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{pmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \end{pmatrix} + \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}. \tag{69}$$

The variable y_{1t} does not cause the variable y_{2t}, at the order 1, then in this case $y_{1t} \overset{(1)}{\nrightarrow} y_{2t}$. If the variable y_{1t} causes the variable y_{2t}, at the order 1, then in this case $y_{1t} \overset{(1)}{\rightarrow} y_{2t}$. In the latter case, this means that y_{2t} is a predictor of y_{1t}, that is, y_{2t} enables to predict y_{1t} with a time horizon of one period. It is also noteworthy that the non causality test is linear, however the regression curve issued from the semi-parametric Gini regression does not necessarily imply a linear model.

The (non) causality tests, between the variables y_{1t} and y_{2t}, stemming from the $VAR(1)$ model are as follows:

$$\left\| \begin{array}{l} H_0 : \Phi_{12} = 0 \\ H_1 : \Phi_{12} \neq 0 \end{array} \right. \iff \left\| \begin{array}{l} H_0 : \ y_{2t} \overset{(1)}{\nrightarrow} y_{1t} \\ H_1 : \ y_{2t} \overset{(1)}{\rightarrow} y_{1t}, \end{array} \right. \tag{70}$$

$$\left\| \begin{array}{l} H_0 : \Phi_{21} = 0 \\ H_1 : \Phi_{21} \neq 0 \end{array} \right. \iff \left\| \begin{array}{l} H_0 : \ y_{1t} \overset{(1)}{\nrightarrow} y_{2t} \\ H_1 : \ y_{1t} \overset{(1)}{\rightarrow} y_{2t}. \end{array} \right. \tag{71}$$

Since $\hat{\Phi}_{ij} \equiv \hat{\beta}_{gg,ij}$ and since, as seen before,

$$\hat{\beta}_{gg,ij} \overset{a}{\sim} \mathcal{N}(\beta_{gg,ij}, \hat{\sigma}_J(\beta_{gg,ij})),$$

then the (non) causality test may be implemented directly by testing the nullity of $\beta_{gg,ij}$.

Things are more complicated for a $VAR(p)$ model. Let us take, as before, only two variables. Then the $VAR(p)$ is given as,

$$\begin{aligned} \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} &= \begin{pmatrix} \Phi_{11,1} & \Phi_{12,1} \\ \Phi_{21,1} & \Phi_{22,1} \end{pmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \end{pmatrix} + \cdots + \\ &+ \begin{pmatrix} \Phi_{11,p} & \Phi_{12,p} \\ \Phi_{21,p} & \Phi_{22,p} \end{pmatrix} \begin{pmatrix} y_{1t-p} \\ y_{2t-p} \end{pmatrix} + \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}. \end{aligned} \tag{72}$$

A joint test may be performed to assess whether y_{1t} may (not) cause y_{2t} and inversely:

$$\left\| \begin{array}{l} H_0 : \Phi_{12,1} = \cdots = \Phi_{12,p} = 0 \\ H_1 : \exists \Phi_{12,i} \neq 0, \ i = 1, \ldots, p \end{array} \right. \iff \left\| \begin{array}{l} H_0 : \ y_{2t} \overset{(1)}{\nrightarrow} y_{1t} \\ H_1 : \ y_{2t} \overset{(1)}{\rightarrow} y_{1t}. \end{array} \right. \tag{73}$$

$$\left\| \begin{array}{l} H_0 : \Phi_{21,1} = \cdots = \Phi_{21,p} = 0 \\ H_1 : \exists \Phi_{21,i} \neq 0, \ i = 1, \ldots, p \end{array} \right. \iff \left\| \begin{array}{l} H_0 : \ y_{1t} \overset{(1)}{\nrightarrow} y_{2t} \\ H_1 : \ y_{1t} \overset{(1)}{\rightarrow} y_{2t}. \end{array} \right. \tag{74}$$

In order to perform the joint test, let us compare the model with the restriction imposed by the null H_0 and without it. For that purpose, we compute the goodness of fit of the model with and without restriction. The Gini R-squared is given by, for instance for the first equation,

$$GR^2 = 1 - \left(\frac{\text{cov}(\hat{\varepsilon}_{1t}, \mathbf{r}_{\hat{\varepsilon}_{1t}})}{\text{cov}(y_{1t}, \mathbf{r}_{y_{1t}})} \right)^2 , \tag{75}$$

where $\mathbf{r}_{\hat{\varepsilon}_{1t}}$ is the rank vector of $\hat{\varepsilon}_{1t}$ and $\mathbf{r}_{y_{1t}}$ the rank vector of y_{1t}. Setting $G\rho_0^2$ being the Gini R-squared of the model with constraint (under the null), and $G\rho^2$ the Gini R-squared without restriction, then (74) is equivalent to test for:

$$\left\| \begin{array}{l} H_0 : G\rho_0^2 - G\rho^2 = 0 \\ H_1 : G\rho_0^2 - G\rho^2 \neq 0 \end{array} \right. \iff \left\| \begin{array}{l} H_0 : \ y_{1t} \overset{(1)}{\nrightarrow} y_{2t} \\ H_1 : \ y_{1t} \overset{(1)}{\rightarrow} y_{2t}. \end{array} \right. \tag{76}$$

Let us denote $U_1 := \text{cov}(\hat{\varepsilon}_{1t}, \mathbf{r}_{\hat{\varepsilon}_{1t}})$ the coGini of the residuals and $U_2 := \text{cov}(y_{1t}, \mathbf{r}_{y_{1t}})$ the coGini of y_{1t}. Then, an estimator of $G\rho^2$ is given by:

$$U = \frac{U_2^2 - U_1^2}{U_2^2}. \tag{77}$$

Since U is a function of U-statistics it is an unbiased and convergent estimator of $G\rho^2$. In the same way, we get that U_0 is a consistent estimator of $G\rho_0^2$ issued from the model with constraints on the parameters. Consequently, the null is tested thanks to the following statistics:

$$U_0 - U \overset{a}{\sim} \mathcal{N}(G\rho_0^2 - G\rho^2, \hat{\sigma}_J(U_0 - U), \tag{78}$$

where $\hat{\sigma}_J(U_0 - U)$ is the Jackknife standard deviation of $U_0 - U$.

It is noteworthy that, in their original paper, Olkin and Yitzhaki (1992) made use of (75) which is obtained under the parametric approach, *i.e.*, with the assumption that:

$$\text{cov}(\hat{y}, \mathbf{r}_{\hat{\varepsilon}}) = 0. \tag{79}$$

The previous equation is true in the parametric case in which linearity is imposed. In order to assess the goodness of fit of the semi-parametric Gini regression, we propose another Gini analysis. Since,

$$y = \hat{y} + \hat{\varepsilon}, \tag{80}$$

then,

$$\text{cov}(y, \mathbf{r}_y) = \text{cov}(\hat{y}, \mathbf{r}_y) + \text{cov}(\hat{\varepsilon}, \mathbf{r}_y). \tag{81}$$

that is,

$$G_y = \frac{4}{n}\text{cov}(\hat{y}, \mathbf{r}_y) + \frac{4}{n}\text{cov}(\hat{\varepsilon}, \mathbf{r}_y), \tag{82}$$

with G_y the Gini index of y. As a consequence the Gini index of the dependent variable (*i.e.* its variability) is explained by the variability of the estimated dependent variable $\hat{y}$ and the variability of the residuals. We deduce that the Gini R-squared may be defined as follows:

$$\frac{G_y}{G_y} = \frac{\text{cov}(\hat{y}, \mathbf{r}_y)}{\text{cov}(y, \mathbf{r}_y)} + \frac{\text{cov}(\hat{\varepsilon}, \mathbf{r}_y)}{\text{cov}(y, \mathbf{r}_y)}, \tag{83}$$

that is,

$$\overline{GR}^2 = 1 - \frac{\text{cov}(\hat{\varepsilon}, \mathbf{r}_y)}{\text{cov}(y, \mathbf{r}_y)}. \tag{84}$$

Thereby, relaxing the hypothesis of linearity of the model used by Olkin and Yitzhaki (1992) implies that the (non) causality test has to be performed as in (78) thanks to the statistics (84).

7 Conclusion

In this paper, we have shown that taking recourse to inequality measurement, such as the Gini index, may have interesting features, see *e.g.* Palestini and Pignataro (2016) for the use of the Atkinson index.

The semi-parametric Gini regression has been shown to be an alternative to the usual estimators available in the literature of VAR models. It offers a wide range of estimators: parametric, non parametric, linear, non linear, etc. This flexibility enables one to test whether a model may be specified with linearity or not, as in the usual Gini regression initiated by Schechtman *et al.* (2011). Also it allows for dealing with non spherical disturbances such as heteroskedasticity and auto-correlation, the so-called generalized Gini regression.

Finally, it is shown that if outliers drastically affect the sample, then it is possible to make use of other tools: a Gini causality test in Granger's sense, a Gini (orthogonal) impulse response functions, and the Gini decomposition of the prevision errors.

References

[1] Colletaz, G. (2017), Les processus VAR : notes du cours de séries temporelles multivariées, *mimeo online*.

[2] Durbin, J. (1954), Errors in variables, *Review of the International Statistical Institute*, 22, 23-32.

[3] Greene, W. (2003), *Econometric Analysis*, Prentice-Hall, Pearson Education (5th edition).

[4] Hoeffding, W. (1948), A class of statistics with asymptotically normal distributions, *Annals of Statistics*, 19, 293-325.

[5] Huber, P. (1981), *Robust Statistics*, Chichester, John Wiley.

[6] Hurvich, C. M. and C.-L. Tsai, (1989), Regression and time series model selection in small samples, *Biometrika*, 76, 297-307.

[7] Ka, N. and S. Mussard (2016), ℓ_1 Regressions: Gini Estimators for Fixed Effects Panel Data, *Journal of Applied Statistics*, 43(8), 1436-1446.

[8] Koenker, R. and G. Bassett (1978), Regression Quantiles, *Econometrica*, 46(1), 33-50.

[9] Olkin, I. and S. Yitzhaki (1992), Gini Regression Analysis, *International Statistical Review*, 60(2), 185-196.

[10] Palestini, A. and G. Pignataro (2016), A graph-based approach to inequality assessment, *Physica A*, 455(1), 65-78.

[11] Schechtman E., Yitzhaki, S. and T. Pudalov (2011), Gini's multiple regressions: two approaches and their interaction, *Metron - International Journal of Statistics*, 0(1), 67-99.

Sims, C. (1980), Macroeconomics and Reality, *Econometrica*, 48(1), 1-48.

[12] Yitzhaki, S. (2003), Gini's Mean difference: a superior measure of variability for non-normal distributions, *Metron*, LXI(2), 285-316.

[13] Yitzhaki, S. and P. Lambert (2013), The Relationship Between the Absolute Deviation from a Quantile and Gini's Mean Difference, *Metron*, 71, 97-104.

[14] Yitzhaki, S. and E. Schechtman (2004), The Gini instrumental variable, or the "double instrumental variable" estimator, *Metron*, LXII(3), 287-313.

[15] Yitzhaki, S. and E. Schechtman (2013), *The Gini Methodology. A Primer on a Statistical Methodology*, Springer.