Martha FLORES-SOSA, PhD (correspondin author)
martha.flores@uadeo.mx
University Autonomous of Occidente, Culiacán, Mexico

Berenice MENDOZA-CARBAJAL, PhD
berenice.mendoza@umich.mx
Michoacan University of San Nicolás de Hidalgo, Morelia, Mexico

Jose M. MERIGÓ, PhD
Jose.Merigo@uts.edu.au
University of Technology Sydney, Sydney, Australia

The Gender Multiple Regression with OWA Operators

Abstract. *The wage gap between men and women is an issue that has become relevant in recent years. Factors such as experience and years of study can affect their salaries differently. This paper presents a multiple regression where wages by gender are analysed using OWA aggregation operators. With the use of aggregation operators, numerous scenarios can be analysed when the obtained OWA parameters are overestimated or underestimated. In our case, the regression has given us information on how the gap behaves, becoming larger according to the educational level, this information taking into account the ENIGH database of INEGI in Mexico.*

Keywords: *OWA Operators, salary, gender, weighting methods, financial decision making.*

JEL Classification: C44, D63, D81, G11.

1. Introduction

In the world, since all times, men and women with the same job do not have the same salaries. In particular, the wage gap that exists in Mexico is greater than 10%, depending on the educational level of men and women, and can even reach close to 40%. There are different tools with which the salary gap is estimated; one of them is simply taking the average salary, another slightly more complex methodology is to consider the Mincer Equation and observe the behaviour of the salary. Mendoza (Mendoza et al., 2022) took these wages and classified them by running regressions.

It is important to point out that despite the fact that the wage gap has been calculated, there are still several pending issues to be addressed, such is the case of making more precise weightings. The OWA operator allows us to perform a reordering, consider maximums and minimums, in addition to giving importance to certain variables of interest, as is the case of taking IOWA (Merigó & Gil-Lafuente, 2009a).

Since the OWA operator began to be studied, it has been used in various applications. The ordered weighted average operator is a very common aggregation

DOI: 10.24818/18423264/59.1.25.16

method that includes maximum, minimum, and average as special cases. It also provides a parameterised family of aggregation operators (Yager, 1988).

An operator in which the reordering step is induced by another device such that the order and position of the arguments depend on the values of their associated order-inducing variables (constructed by Yager and Filev as an extension of the OWA operator) is called induced ordered weighted average (IOWA) operator (Yager et al., 1999), so the difference with respect to OWA is that the reordering step is not carried out with the values of the arguments. With this operator, we can create other types of operators such as the generalised OWA operator (GOWA) which, as its name indicates, generalises the OWA operator using generalised means (Yager, 2004). GOWA generalises a variety of average operators, such as the Induced Generalised Ordered Weighted Averaging (IGOWA) this operator uses the moving average and the induced operator (Merigó & Gil-Lafuente, 2009b).

Another extension of the OWA operator is called the ordered weighted averaging–weighted average (OWAWA) and the induced ordered weighted averaging–weighted average (IOWAWA) operator, which have been created by Merigó (Merigó, 2011; Merigó et al., 2015).

2. Preliminaries

The section presents a review of the OWA operators and its family used, and additionally a brief of linear regression is presented.

2.1 OWA operators

An aggregation operator that provides a parameterised family of operators to consider the arithmetic mean, the maximum, and minimum is the OWA operator (Yager, 1988). This uses a vector of weights that accompany a vector of arguments reordered according to specific criteria. It is defined as follows:

Definition 1. It is an OWA operator with dimensions n if exist a model $OWA: R^n \rightarrow R$ such that it has a weights vector W thus the components are $w_i \in [0,1]$ and $\sum_{i=1}^{n} w_i = 1$, and then:

$$\mathrm{OWA}(a_1, a_2, \dots, a_n) = \sum_{j=1}^{n} w_j b_j, \tag{1}$$

where b_j is the jth largest argument a_i. The OWA operator considers several properties as monotonicity, idempotence, and symmetry (Yager, 1988).

An popular extension of the family OWA is the induced OWA (IOWA) operator (Yager & Filev, 1999). The IOWA operator introduces an induced vector that reorders the arguments. The new order can model the arguments in different situations (Yager, 2003). It can be defined as follows:

Definition 2. An IOWA operator of dimension n is a mapping $IOWA: R^n \rightarrow R$ with a weights vector $W = [w_1, w_2, .., w_n]^T$ where $0 \leq w_i \leq 1$ and $\sum_{i=1}^{n} w_i = 1$, and an induced vector $U = [u_1, u_2, .., u_n]^T$. Then:

$$\mathrm{IOWA}(\langle u_1, a_1 \rangle \langle u_2, a_2 \rangle, \dots, \langle u_n, a_n \rangle) = \sum_{j=1}^{n} w_j b_j\ , \tag{2}$$

where b_j is the argument a_i that have the jth largest u_i. The IOWA operator also includes the proprieties monotonicity, idempotence, and symmetry.

Another interesting OWA extension is when the arguments are drawn from the unit interval. The Generalised Ordered Weighted Aggregation operator (GOWA) (Yager, 2004) provides an additional way to analyses more scenarios on OWA means. So:

Definition 3. It is a GOWA operator with mapping $GOWA: R^n \to R$ if it has an associated weights vector W thus $0 \leq w_i \leq 1$ and $w_i + \cdots + w_n = 1$, and a lambda parameter $\lambda \in [-\infty, \infty]$ is considered. So:

$$\mathrm{GOWA}(a_1, a_2, \dots, a_n) = \left(\sum_{j=1}^{n} w_j {b_j}^{\lambda}\right)^{1/\lambda}, \qquad (3)$$

where b_j is the jth largest a_i. According to the value of the λ parameter, the GOWA operator analyse different special cases. If $\lambda = 1$, the GOWA operator is the OWA operator (Yager, 1988), When $\lambda = 0$, the GOWA operator is the OWG operator (Chiclana, Herrera, & Herrera-Viedma, 2000). When $\lambda = -1$ the aggregation is the OWHA operator (Chen et al., 2004), and when $\lambda = 2$, we form the OWQA operator (Dyckhoff & Pedrycz, 1984).

An OWA extension that considers additional weights unifying the OWA operator and weighted average (WA) in the same formulation is the ordered weighted averaging-weighted average (OWAWA) operator (Merigó, 2011) It. The definition is as follows:

Definition 4. the OWAWA operator of dimension *n* is a mapping $OWAWA: R^n \to R$ that has two vectors, the W with a set $w_i =\in [0,1]$ and $\sum_{i=1}^{n} w_i = 1$, and the V such that $v_i =\in [0,1]$ and $\sum_{i=1}^{n} v_i = 1$, the equation is as follows:

$$OWAWA(a_1, a_2, \dots, a_n) = \sum_{j=1}^{n} \widehat{v_j}\, b_j, \qquad (4)$$

where b_j is the *j*th largest a_i. v_j is a vector of weights composed of vectors W and V, considering the degree of importance of each one. So, $\widehat{v_j} = \beta w_j + (1-\beta) v_j$ with $\beta \in [0, 1]$. The OWAWA operator has the same properties as the OWA operator.

The OWA family commonly unites two or more operators into one. Therefore, the IOWAWA operator can be combined with the IOWA operator. The IOWAWA operator uses a weighted average weighted vector and an induced vector with the attributes. This is:

Definition 5. An IOWAWA operator of dimension *n* is a model $IOWAWA: R^n \to R$ with a weight vector $W = [w_1, w_2, .., w_n]^T$ such that $w_i =\in [0,1]$ and $w_i + \cdots + w_n = 1$, and a second weight vector V with $\sum_{i=1}^{n} v_i = 1$ and $v_i \in [0,1]$, then an induced IOWA pair (v_i, a_i) is considered for the reorder of the arguments, so:

$$IOWAWA(\langle u_1, a_1\rangle\langle u_2, a_2\rangle, \dots, \langle u_n, a_n\rangle) = \sum_{j=1}^{n} \widehat{v_j}\, b_j, \quad (5)$$

where b_j is the attribute with the jth largest value of the induce vector. $\widehat{v_j} = \beta w_j + (1-\beta)v_j$ with $\beta \in [0, 1]$. The IOWAWA operator is monotonicity, idempotence, and symmetry.

The IOWAWA can also analyse more scenarios using generalised means. The IGOWAWA operator (Merigó, 2009) considers a second weight vector, with induced variables and a lambda parameter. The definition is as follows:

Definition 6. An IGOWAWA operator of dimension n is a model $IGOWAWA$: $R^n \rightarrow R$ with a weight vector W such that $0 \leq w_i \leq 1$ and $w_i + \cdots + w_n = 1$, and an additional weight vector V with $\sum_{i=1}^{n} v_i = 1$ and $v_i \in [0,1]$, then an induced IOWA pair (v_i, a_i) is considered for the reorder of the arguments, so:

$$IGOWAWA(a_1, a_2, \dots, a_n) = \left(\sum_{j=1}^{n} \hat{v}_j {b_j}^{\lambda}\right)^{1/\lambda} \quad (6)$$

where b_j is the argument a_i that have the jth largest u_i. $\widehat{v_j} = \beta w_j + (1-\beta)v_j$ with $\beta \in [0, 1]$. $\lambda \in [-\infty, \infty]$.

2.2 Linear regression

In the area of modelling and estimation, linear regression is one of the fundamental methodologies. The LR is an approach to understanding the mutual effects of an independent variable on the response variables (Kneip et al., 2016; Schmidt & Finan, 2018). A model with two or more response variables is a multiple regression. The linear fitting in multiple linear regression is attempted by keeping constant all but one of the predictor variables (Moreira, 2016; Leung et al., 2017). It can be defined as follows:

Definition 7. A multiple regression is a set of variables (x_k, y_k, z_k) where $k = 1, \dots, K$: $x_k \in U^n$, $y_k \in U, z_k \in U^n$, then, exist a model $f_\theta: R^n \rightarrow R$, parameterised by a parameter vector $\theta = \alpha, \beta_2, \beta_3$. Multiple linear regression develops in a following equation:

$$\mathrm{y_j} = \alpha + \beta_1 \mathrm{x_j} + \beta_2 \mathrm{z_j}. \quad (7)$$

A common framework for estimating the parameter vector is the ordinary least-squares OLS (Gujarati & Porter, 2009; Bun & Harrison, 2019). This minimising the sum of the squared error between the predicted and actual observations as $\sum \hat{u}_i^2$ as $Min \sum \hat{u}_i^2 = \sum(y_i - \alpha - \beta_1 x_i - \beta_2 z_i)^2$. The formulation for each parameter is as follows:

$$\alpha = \overline{\mathrm{Y}} - \beta_1 \overline{\mathrm{X}} - \beta_2 \overline{\mathrm{Z}} \quad (8)$$

$$\beta_1 = \frac{(\sum y_i x_i)(\sum z_i^2) - (\sum y_i z_i)(\sum x_i z_i)}{(\sum x_i^2)(\sum z_i^2) - (\sum x_i z_i)^2} = \frac{[\mathrm{Cov}(y,x)][\mathrm{var}(z)] - [\mathrm{Cov}(y,z)][\mathrm{Cov}(x,z)]}{[\mathrm{var}(x)][\mathrm{var}(z)] - [\mathrm{Cov}(x,z)]^2} \quad (9)$$

$$\beta_2 = \frac{(\sum y_i z_i)(\sum x_i^2) - (\sum y_i x_i)(\sum x_i z_i)}{(\sum x_i^2)(\sum z_i^2) - (\sum x_i z_i)^2} = \frac{[\mathrm{Cov}(y,z)][\mathrm{var}(x)] - [\mathrm{Cov}(y,x)][\mathrm{Cov}(x,z)]}{[\mathrm{var}(x)][\mathrm{var}(z)] - [\mathrm{Cov}(x,z)]^2} \quad (10)$$

where $\bar{x}$, $\bar{y}$, and $\bar{z}$ are the averages in the sets x_k, y_k, z_k severally.

3. OWA operator in multiple linear regression

The main idea of the multiple linear regression with OWA operators is to estimate the means in the OLS process by ordering and weighting the arguments. This is, OWA variances and covariances are used (Merigó et al., 2015; Blanco-Mesa et al., 2019).

Therefore, if weighted weights are used on the means of the OLS estimate, we obtain the MLR-OWAWA. It can be used in cases where the uncertainty of the data demands greater complexity in the estimation process. The definition can be developed as follows:

Definition 8. An OWAWA multiple linear regression with two response variables of dimension *n* is a model $OWAWA: R^n \rightarrow R$ given the parameters $x_k \in U^n$, $y_k \in U$ and $z_k \in U^n$ such have two weights vector W and V with $w_i = \in [0,1]$; $\sum_{i=1}^{n} w_i = 1$, and $0 \leq v_i \leq 1$;, where the model is as follows:

$$y_{OWAWA} = \alpha_{OWAWA} + \beta_{1_{OWAWA}} x_j + \beta_{2_{OWAWA}} z_j \quad (11)$$

where α_{OWAWA}, $\beta_{1_{OWAWA}}$ and $\beta_{2_{OWAWA}}$ are estimated through OLS method with OWAWA variances and covariances as follows:

$$\beta_{1OWAWA} = \frac{[Cov_{OWAWA}(y,x)][var_{OWAWA}(z)] - [Cov_{OWAWA}(y,z)][Cov_{OWAWA}(x,z)]}{[var_{OWAWA}(x)][var_{OWAWA}(z)] - [Cov_{OWAWA}(x,z)]^2}$$

$$\beta_{1OWAWA} = \frac{\left[\sum_{k=1}^{k} w_j (y_j - v)(x_j - \mu)\right]\left[\sum_{k=1}^{k} w_j (z_j - v)^2\right] - \left[\sum_{k=1}^{k} w_j (y_j - v)(z_j - v)\right]\left[\sum_{k=1}^{k} w_j (x_j - \mu)(z_j - v)\right]}{\left[\sum_{k=1}^{k} w_j (x_j - \mu)^2\right]\left[\sum_{k=1}^{k} w_j (z_j - v)^2\right] - \left[\sum_{k=1}^{k} w_j (x_j - \mu)(z_j - v)\right]^2} \quad (12)$$

where x_j, z_j and y_j is the jth largest data in the variables *x*, *z* and *y* severally, and μ, ν and υ are OWAWA means.

For β_2 estimation the formula is developed as follows:

$$\beta_{2OWAWA} = \frac{[Cov_{OWAWA}(y,z)][var_{OWAWA}(x)] - [Cov_{OWAWA}(y,x)][Cov_{OWAWA}(x,z)]}{[var_{OWAWA}(x)][var_{OWAWA}(z)] - [Cov_{OWAWA}(x,z)]^2}$$

$$\beta_{2OWAWA} = \frac{\left[\sum_{k=1}^{k} w_j(y_j-\upsilon)(z_j-\nu)\right]\left[\sum_{k=1}^{k} w_j(x_j-\mu)^2\right] - \left[\sum_{k=1}^{k} w_j(y_j-\upsilon)(x_j-\mu)\right]\left[\sum_{k=1}^{k} w_j(x_j-\mu)(z_j-\nu)\right]}{\left[\sum_{k=1}^{k} w_j(x_j-\mu)^2\right]\left[\sum_{k=1}^{k} w_j(z_j-\nu)^2\right] - \left[\sum_{k=1}^{k} w_j(x_j-\mu)(z_j-\nu)\right]^2}, \quad (13)$$

where x_j, z_j and y_j is the jth largest arguments in the variables *x*, *z* and *y* severally, and μ, ν and υ are OWAWA means. So, let us into the estimation of α using β_1 and β_2 as follows:

$$\alpha_{OWAWA} = \upsilon - \beta_{1OWAWA}\mu - \beta_{2OWAWA}\nu, \quad (14)$$

where μ, ν and υ are OWAWA means.

The MLR-OWAWA shares some special cases with OWA operators. Depending on the ordering weights in variance and covariance, we can obtain maximus. Then:

The MAX- Var_{OWAWA} is as follows:

$$MAX - Var_{OWA} = \sum_{j=1}^{n} w_j D_j, \quad (15)$$

where D_j is the largest of the $(a_i - \mu)^2$, using μ as OWAWA minimum, and $w_j = 1$. In a similar way MAX- Cov_{OWAWA} is developed as follows:

$$MAX - Cov_{OWA} = \sum_{j=1}^{n} w_j K_j, \quad (16)$$

where K_j is the jth largest of the $(x_i - \mu)(y_i - \nu)$, and μ and ν are OWAWA minimum of X and Y respectively and $w_j = 1$.

Then, an MLR-OWAWA with maximums can be calculated as follows:

$$MLR - OWAWA_{MAX} = \alpha_{MAX} + \beta_{2MAX}x_i + \beta_{2MAX}z_i. \quad (17)$$

$$\beta_{1MAX} = \frac{[MAXCov_{OWAWA}(y,x)MAXvar_{OWAWA}(z)] - [MAXCov_{OWAWA}(y,z)MAX\ Cov_{OWAWA}(x,z)]}{[MAXvar_{OWAWA}(x)var_{OWAWA}(z)] - [MAXCov_{OWAWA}(x,z)]^2}. \quad (18)$$

$$\beta_{2MAX} = \frac{[MAXCov_{OWAWA}(y,z)MAXvar_{OWAWA}(x)] - [MAXCov_{OWAWA}(y,x)MAX\ Cov_{OWAWA}(x,z)]}{[MAXvar_{OWAWA}(x)MAXvar_{OWAWA}(z)] - [MAXCov_{OWAWA}(x,z)]^2}. \quad (19)$$

$$\alpha_{MAX} = \upsilon - \beta_{1MAX}\mu - \beta_{2MAX}\nu. \tag{20}$$

Note that the use of variances and covariances maximum can be used in different parts of the estimate. Therefore, a large number of combinations can be made at the convenience of the decision maker.

Because the OWAWA variances and covariances tend to zero when the minimum is investigated, the MLR-OWAWA calculation with minimums is indeterminate.

The multiple linear regression also can be joined with induced operators. The MLR-IOWAWA is a multiple regression estimated with variances and covariances using IOWAWA means, where the order of these depends on an induced vector. This is:

Definition 9. It is an MLR-IOWAWA with two response variables if exist a model $IOWAWA: R^n \to R$ given the sets X, Y, Z, such have two weights vector W and V with components that have values from zero to one and the sum of them is equal to one. The model is defined as follows:

$$y_{IOWAWA} = \alpha_{IOWAWA} + \beta_{1_{IOWAWA}} x_j + \beta_{2_{IOWAWA}} z_j \tag{21}$$

Where the parameters are estimated as follows:

$$\beta_{1IOWAWA} = \frac{[Cov_{IOWAWA}(y,x)][var_{IOWAWA}(z)] - [Cov_{IOWAWA}(y,z)][Cov_{IOWAWA}(x,z)]}{[var_{IOWAWA}(x)][var_{IOWAWA}(z)] - [Cov_{IOWAWA}(x,z)]^2}$$

$$\beta_{1IOWAWA} = \frac{\left[\sum_{k=1}^{k} w_j(y_j-\upsilon)(x_j-\mu)\right]\left[\sum_{k=1}^{k} w_j(z_j-\nu)^2\right]-\left[\sum_{k=1}^{k} w_j(y_j-\upsilon)(z_j-\nu)\right]\left[\sum_{k=1}^{k} w_j(x_j-\mu)(z_j-\nu)\right]}{\left[\sum_{k=1}^{k} w_j(x_j-\mu)^2\right]\left[\sum_{k=1}^{k} w_j(z_j-\nu)^2\right]-\left[\sum_{k=1}^{k} w_j(x_j-\mu)(z_j-\nu)\right]^2} \tag{22}$$

$$\beta_{2IOWAWA} = \frac{[Cov_{IOWAWA}(y,z)][var_{IOWAWA}(x)] - [Cov_{IOWAWA}(y,x)][Cov_{IOWAWA}(x,z)]}{[var_{IOWAWA}(x)][var_{IOWAWA}(z)] - [Cov_{IOWAWA}(x,z)]^2}$$

$$\beta_{2IOWAWA} = \frac{\left[\sum_{k=1}^{k} w_j(y_j-\upsilon)(z_j-\nu)\right]\left[\sum_{k=1}^{k} w_j(x_j-\mu)^2\right]-\left[\sum_{k=1}^{k} w_j(y_j-\upsilon)(x_j-\mu)\right]\left[\sum_{k=1}^{k} w_j(x_j-\mu)(z_j-\nu)\right]}{\left[\sum_{k=1}^{k} w_j(x_j-\mu)^2\right]\left[\sum_{k=1}^{k} w_j(z_j-\nu)^2\right]-\left[\sum_{k=1}^{k} w_j(x_j-\mu)(z_j-\nu)\right]^2}, \tag{23}$$

$$\alpha_{IOWAWA} = \upsilon - \beta_{1IOWAWA}\mu - \beta_{2IOWAWA}\nu, \tag{24}$$

where x_j, z_j and y_j is the jth largest data in the variables x, z and y, and μ, ν and υ are IOWAWA means.

The MLR-IOWAWA has the properties of OWA operators, so:

- The MLR-IOWAWA is monotonic if let a further ordered argument vector $w = [b_1, b_2, \dots, b_n]$ where $a_j \geq b_j$ then $F(y_{owa}(a_1, a_2, \dots, a_n)) \geq F(y_{OWA}(b_1, b_2, \dots, b_3))$.

- Symmetry if two ordered argument vectors $(A = a_1, a_2, \dots, a_n; A' = a_1' a_2', \dots, a_n')$ are A=A′, then $F\big(y_{IOWAWA}(a_1, a_2, \dots, a_n)\big) = F(y_{IOWAWA}(a_1', a_2', \dots, a_n')$.
- An MLR-IOWAWA is idempotent if $a_j = a$, for all $j = 1, \dots, n$, then $F(y_{IOWAWA}(a_1, a_2, \dots, a_n) = a$.

In order to analyse additional scenarios of the estimators in MLR-IOWAWA, the generalisation of these is proposed. The MLR-IGOWAWA is a tool that estimates parameters in a linear regression with two variables using means that consider two vectors of weights, the arguments are ordered according to induction, and a parameter lambda is considered. The definition is as follows:

Definition 10. It is an MLR-IGOWAWA with two response variables if there is a model $IGOWAWA: R^n \rightarrow R$ given three sets $x_k \in U^n$, $y_k \in U$ and $z_k \in U^n$ and a weight vector $W = [w_1, w_2, .., w_n]^T$ such that $w_i = \in [0,1]$; $\sum_{i=1}^{n} w_i = 1$, and a second weight vector $V = [v_1, v_2, .., v_n]^T$ $0 \leq v_i \leq 1$; $v_i + \cdots + v_n = 1$, additionally a parameter $\lambda \in [-\infty, \infty]$ is defined. So:

$$y_{IGOWAWA} = \alpha_{IGOWAWA} + \beta_{1_{IGOWAWA}} x_j + \beta_{2_{IGOWAWA}} z_j \tag{25}$$

Where the parameters are estimated as follows:

$$\beta_{1IGOWAWA} = \frac{[Cov_{IGOWAWA}(y,x)][var_{IGOWAWA}(z)] - [Cov_{IGOWAWA}(y,z)][Cov_{IGOWAWA}(x,z)]}{[var_{IGOWAWA}(x)][var_{IGOWAWA}(z)] - [Cov_{IGOWAWA}(x,z)]^2}$$

$$\beta_{1IGOWAWA} = \frac{\left[\sum_{k=1}^{k} w_j(y_j-\upsilon)(x_j-\mu)\right]\left[\sum_{k=1}^{k} w_j(z_j-\nu)^2\right]-\left[\sum_{k=1}^{k} w_j(y_j-\upsilon)(z_j-\nu)\right]\left[\sum_{k=1}^{k} w_j(x_j-\mu)(z_j-\nu)\right]}{\left[\sum_{k=1}^{k} w_j(x_j-\mu)^2\right]\left[\sum_{k=1}^{k} w_j(z_j-\nu)^2\right]-\left[\sum_{k=1}^{k} w_j(x_j-\mu)(z_j-\nu)\right]^2}, \tag{26}$$

$$\beta_{2IGOWAWA} = \frac{[Cov_{IGOWAWA}(y,z)][var_{IGOWAWA}(x)] - [Cov_{IGOWAWA}(y,x)][Cov_{IGOWAWA}(x,z)]}{[var_{IGOWAWA}(x)][var_{IGOWAWA}(z)] - [Cov_{IGOWAWA}(x,z)]^2}$$

$$\beta_{2IGOWAWA} = \frac{\left[\sum_{k=1}^{k} w_j(y_j-\upsilon)(z_j-\nu)\right]\left[\sum_{k=1}^{k} w_j(x_j-\mu)^2\right]-\left[\sum_{k=1}^{k} w_j(y_j-\upsilon)(x_j-\mu)\right]\left[\sum_{k=1}^{k} w_j(x_j-\mu)(z_j-\nu)\right]}{\left[\sum_{k=1}^{k} w_j(x_j-\mu)^2\right]\left[\sum_{k=1}^{k} w_j(z_j-\nu)^2\right]-\left[\sum_{k=1}^{k} w_j(x_j-\mu)(z_j-\nu)\right]^2}, \tag{27}$$

$$\alpha_{IGOWAWA} = \upsilon - \beta_{1IGOWAWA}\mu - \beta_{2IGOWAWA}\nu, \tag{28}$$

where x_j, z_j and y_j is the jth largest data in the variables *x*, *z* and *y* severally, and μ, ν and υ are IGOWAWA means.

An interesting issue in the application of the OWA operators is the weights and their analysis. Using different measures of weight vector, some information can be obtained.

The orness and the entropy of dispersion (Yager, 1988) can be defined as follows:

If occur that the first weight $w_1 = 1$, a pure "or" operator appears. Then, the closer all the total weights to being in w_1 is the degree of orness, the formulation is as following:

$$\alpha(W) = \sum_{j=1}^{n} w_j^* \left(\frac{n-j}{n-1}\right), \tag{29}$$

where w_j^* is the w_j weight with the jth largest a_i.

The variability and the use of the inputs by the OWA weights are captured by the entropy of dispersion, so:

$$H(W) = -\sum_{j=1}^{n} w_j \ln(w_j). \tag{30}$$

Balance operator (Yager, 1996) measures the degree of relationship between favouring the higher valued elements or lower-valued elements in an OWA operator, we use:

$$BAL(W) = \sum_{j-1}^{n} \left(\frac{n+1-2j}{n-1}\right) w_j. \tag{31}$$

Finally, divergence (Yager, 2002) distinguishes between two OWA weights vectors. The formulation is as follows:

$$DIV(W) = \sum_{j=1}^{n} w_j \left(\frac{n-j}{n-1} - \alpha(W)\right)^2. \tag{32}$$

The measurement of the weight vector can measure the vectors in α, β_1 and β_2.

In order to clarify the operation of the MLR-IGOWMA, a numerical example is presented below.

Example. Considering the following set for the variables $(Y = 3, 5, 2)$, $(X = 2, 4, 3)$ and $(Z = 4, 5, 6)$, with a weighting vector $(W = 0.4, 0.3, 0.3.)$, the induced vector $(U = 1, 2, 4)$ and the second weight vector $(V = 0.5, 0.2, 0.3.)$ are considered. Then we have $\beta = 0.6$ and $\lambda = 2$. First the $\hat{v}_j$ is:

$$\hat{v}_j = 0.44, 0.26, 0.3$$

then the IOWA is defined:

$$IGOWAWA_y = [0.44(2)^2 + 0.26(5)^2 + 0.3(3)^2]^{1/2} = 3.31$$
$$IGOWAWA_x = [0.44(3)^2 + 0.26(4)^2 + 0.3(2)^2]^{1/2} = 3.05$$
$$IGOWAWA_z = [0.44(6)^2 + 0.26(5)^2 + 0.3(4)^2]^{1/2} = 4.98$$

The elements of variance and covariance are as follow:

$(y-\upsilon)$	$(x-\mu)$	$(z_j-\nu)$	$(x-\mu)^2$	$(z_j-\nu)^2$	$(x-\mu)(y-\upsilon)$	$(x-\mu)(z-\nu)$	$(y-\upsilon)(z-\nu)$
-1.31	-0.05	1.02	0.003	1.040	0.069	-0.054	-1.337
1.69	0.95	-0.98	0.897	0.960	1.600	-0.928	-1.656
-0.30	-1.05	-0.98	1.109	0.960	0.327	1.032	0.304

Variances and covariances IGOWAWA are calculated:

$$Var_{IGOWAWA}(x) = [0.44(0.03)^2 + 0.26(0.897)^2 + 0.3(1.109)^2]^{1/2} = 0.76$$

$$Var_{IGOWAWA}(z) = [0.44(1.040)^2 + 0.26(1.96)^2 + 0.3(1.96)^2]^{1/2} = 0.99$$

$$Cov_{IGOWAWA}(x,y) = [0.44(0.069)^2 + 0.26(1.600)^2 + 0.3(0.327)^2]^{1/2} = 0.83$$

$$Cov_{IGOWAWA}(y,z) = [0.44(-0.054)^2 + 0.26(-0.928)^2 + 0.3(1.032)^2]^{1/2} = 0.73$$

$$Cov_{IGOWAWA}(x,z) = [0.44(-1.337)^2 + 0.26(-1.656)^2 + 0.3(0.304)^2]^{1/2} = 1.23$$

The α, $\beta_1$and β_2 are estimated:

$$\beta_{1IGOWAWA} = \frac{(0.83 * 0.99) - (1.23 * 0.73)}{(0.76 * 0.99) - (0.73)^2} = -0.36$$

$$\beta_{2IGOWAWA} = \frac{(1.23 * 0.76) - (0.83 * 0.73)}{(0.76 * 0.99) - (0.73)^2} = 1.51$$

$$\alpha_{IGOWAWA} = 3.31 - (-0.36 * 3.05) - (1.51 * 4.98) = -3.09$$

We have the following model:

$$y_{IGOWAWA} = -3.09 - 0.36x + 1.51z$$

Note that this example can be applied in MLR-OWAWA and MLR-IOWAWA.

4. Estimating salary by gender in Mexico

We are interested in taking into consideration real data from Mexico, in which we can weigh the salaries between men and women with the previously described and achieve a result that allows us to make decisions. That is why this study is based on the National Survey of Household Income and Expenses (ENIGH) 2018 of the National Institute of Statistics and Geography (INEGI).

This article presents an estimate of salaries in Mexico by gender using a model that considers years of study and years of work experience. The models consider different OWA operators; therefore, six models are developed as follows:

For women:

1. $OWAWA_w$ $\quad S_{w_{OWAWA}} = \upsilon - \beta_{1OWAWA}YS_W - \beta_{2OWAWA}WE_W$
2. $IOWAWA_w$ $\quad S_{w_{IOWAWA}} = \upsilon - \beta_{1IOWAWA}YS_W - \beta_{2IOWAWA}WE_W$
3. $IGOWAWA_w$ $\quad S_{w_{IGOWAWA}} = \upsilon - \beta_{1IGOWAWA}YS_W - \beta_{2IGOWAWA}WE_W$

For men:

1. $OWAWA_m$ $S_{m_{OWAWA}} = \upsilon - \beta_{1OWAWA}YS_m - \beta_{2OWAWA}WE_m$
2. $IOWAWA_m$ $S_{m_{IOWAWA}} = \upsilon - \beta_{1IOWAWA}YS_m - \beta_{2IOWAWA}WE_m$
3. $IGOWAWA_m$ $S_{m_{IGOWAWA}} = \upsilon - \beta_{1IGOWAWA}YS_m - \beta_{2IGOWAWA}WE_m$

Where S is salary, YS is years of study and WE is work experience.

Calculating the parameters of the models

The parameters in OWA models by gender are calculated as equations 11, 21 and 25 using OWAWA, IOWAWA and IGOWAWA. For this, we perform the following steps:

Step 1: Defining the number of elements considered in the variable models. Fifteen ranges have been developed in the information of salaries, years of education, and work experience. The average has been calculated from these ranges, which are the series with which this application works. Table 1 shows the data.

Table 1. Model and vectors weights

Sw	YSw	WEw	Sm	YSf	WEf	w	v
15858.86	15.28	2.72	35611.06	16.75	75.46	0.07	0.10
23138.39	16.35	6.65	34668.78	16.51	70.38	0.08	0.09
26336.59	16.75	11.25	32856.17	16.35	65.24	0.09	0.09
28451.30	16.51	16.49	32093.36	16.11	59.70	0.04	0.08
30304.01	16.11	21.89	30304.01	15.72	54.11	0.07	0.08
32093.36	15.72	27.28	28458.70	15.57	48.54	0.07	0.07
34668.78	15.57	32.43	28451.30	15.57	42.95	0.06	0.07
35611.06	15.57	37.43	26336.59	15.28	37.43	0.06	0.06
32856.17	15.05	42.95	24932.93	15.05	32.43	0.08	0.06
28458.70	14.46	48.54	23138.39	14.46	27.28	0.09	0.06
24932.93	13.89	54.11	22369.85	13.89	21.89	0.05	0.06
22369.85	13.30	59.70	22025.56	13.30	16.49	0.05	0.05
20493.54	12.76	65.24	20898.75	12.76	11.25	0.06	0.05
22025.56	12.62	70.38	20493.54	12.62	6.65	0.08	0.04
20898.75	12.54	75.46	15858.86	12.54	2.72	0.05	0.04

Source: Own elaboration based on data from the National Survey of Household Income and Expenditure (ENIGH) 2018, INEGI.

Step 2: the construction of weights vector *w* for IOWA and IOWAWA are developed. Weights vectors are shown in Table 1. The vector of weights *w* is defined randomly, while the vector *v* is overestimating the first elements of the ordering, which in this case will be the largest values.

Step 3: the OWA means are calculated for the different OWAs and models. Note that, $\boldsymbol{\omega}$ is the salary OWAs mean, ρ is the years education OWAs mean and φ is the OWAs mean in work experience. The table shows the results:

Table 2. Analysis of OWAs means

	ω	ρ	φ
OWAWAw	24931.48	15.07	41.42
IOWAWAw	24479.96	15.04	34.95
IGOWAWAw	23995.41	14.82	12.54
OWAWAm	27342.89	15.03	41.36
IOWAWAm	26621.03	15.01	34.98
IGOWAWAm	25308.13	14.87	14.03

Source: Own elaboration based on data from the National Survey of Household Income and Expenditure (ENIGH) 2018, INEGI.

Step 4: For the variance and covariance OWA, we consider the same weight vector used in OWA means. Two variances are considered for years of education and work experience, and the covariances are for the relationship between the three variables. The results are in Table 3.

Table 3. Variances and covariances OWA

	Var(YE)	Var(WE)	Cov(S, YE)	Cov(YE, WE)	Cov(S, WE)
OWAWAw	3.25	538.78	5374.67	41.29	70414.79
IOWAWAw	3.17	535.54	1649.77	-38.56	-327.66
IGOWAWAw	0.47	35.11	710.78	-60.25	-31625.60
OWAWAm	1.92	522.15	7377.45	30.86	125913.92
IOWAWAm	1.89	516.31	3162.58	-28.31	-3958.05
IGOWAWAm	0.26	39.74	4461.25	-1979.86	-54041.51

Source: Own elaboration based on data from the National Survey of Household Income and Expenditure (ENIGH) 2018, INEGI.

Step 4: the parameters are estimated: α, β_1, and β_2 are calculated with the variance and covariance OWA. Table 4 shows $the\ \beta_1$ parameters.

Table 4. Results

	α	$\boldsymbol{\beta_1}$	$\boldsymbol{\beta_2}$
OWAWAw	-264.71	150.97	22,666.65
OWAWAm	-672.56	280.88	25,831.72
IOWAWAw	4,132.51	296.94	-48,048.09
IOWAWAm	8,792.62	474.36	-121,933.65
IGOWAWAw	520.33	-7.77	16,380.21
IGOWAWAm	27.25	-2.24	24,934.36

Source: Own elaboration based on data from the National Survey of Household Income and Expenditure (ENIGH) 2018, INEGI.

The results are represented in the following model:

$$S_{w_{OWAWA}} = -264.72 + 150.98YS_W + 22{,}666.66WE_W$$
$$S_{w_{IOWAWA}} = 4132.51 + 296.95YS_W - 48{,}048.10WE_W$$
$$S_{w_{IGOWAWA}} = 520.34 - 7.78YS_W + 16{,}380.22WE_W$$
$$S_{m_{OWAWA}} = -672.56 + 280.89YS_m + 25{,}831.72WE_m$$
$$S_{m_{IOWAWA}} = 8{,}792.63 + 474.36YS_m - 121{,}933.65WE_m$$
$$S_{m_{IGOWAWA}} = 27.25 - 2.25YS_m + 24{,}934.36WE_m$$

The general results describe how the years of experience have a more significant impact than years of study on salaries for both women and men. When we use a second weighting vector, giving greater importance to the highest salaries on the IOWAWA models, the years of study and expertise multiply men's wages to a greater extent than women's. In this case an interesting situation is observed because without years of study and without experience, women would have lower salaries than men. Additionally, when greater importance is given to higher salaries, it is observed that experience has a more negative impact on men than on women. In the case of the IGOWAWA models, they also confirm that years of study and expertise give men a higher salary than women.

5. Conclusions

The differences in salaries between men and women have been studied in recent years. Knowing the scenarios and magnitude where this occurs brings us closer to identifying the problem and trying to overcome it.

This work proposes a multiple regression with OWA aggregation operators in order to obtain estimates where the parameters can be underestimated and overestimated depending on the scenario to be analysed. The proposal has been called MLR-IGOWAWA.

The application of this tool has focused on the effect that experience and years of study have on the salaries of men and women. The results show that men tend to earn higher salaries, and this wage increases as higher salaries are taken into account. In other words, there is a difference in salaries between men and women, and this difference grows larger as salaries increase.

The proposed tool has been applied to a particular country. That is why future research will seek to corroborate results in other countries.

References

[1] Chen, H., Liu, C., Sheng, Z. (2004), *Induced ordered weighted harmonic averaging (IOWHA) operator and its application to combination forecasting method. Chinese Journal of Management Science,* 12, 35-44.

[2] Chiclana, F., Herrera, F., Herrera-Viedma, E. (2000), *The ordered weighted geometric operator:Properties and application. In: Proc of 8th Int Conf on Information Processing and Management of Uncertainty in Knowledge-based Systems*, 985-991, Madrid, Spain.

[3] Blanco-Mesa, F., León-Castro, E., Merigo, J., Herrera-Viedma, E. (2019), *Variances with Bonferroni means and ordered weighted averages. Variances with Bonferroni means and ordered weighted averages,* 34(11), 3020-3045.

[4] Bun, M., Harrison, T. (2019), *OLS and IV estimation of regression models including endogenous interaction terms. Econometric Reviews,* 38(7), 814-827.

[5] Dyckhoff, H., Pedrycz, W. (1984), *Generalized means as model of compensative connectives. Fuzzy Sets and Systems,* 14, 143-54.

[6] Gujarati, D., Porter, D. (2009), *Basic econometrics* (Vol. 5ta ed). New York, NY: McGraw-Hill Irwin, SUA.

[7] Hamilton, J. (1995), *Time series analysis.* New Jersey: Princeton University, SUA.

[8] Kneip, A., Poss, D., Sarda, P. (2016), *Functional linear regression with points of impact. The Annals of Statistics,* 44(1), 1-30.

[9] Leung, H., Schiereck, D., Schroeder, F. (2017), *Volatility spillovers and determinants of contagion: Exchange rate and equity markets during crises. Economic Modelling,* 61, 169-180.

[10] Merigó, J. (2009), *Induced generalized aggregation operators in the weighted average. Intelligent Decision Making Systems*, 625-630.

[11] Merigó, J. (2011), *A unified model between the weighted average and the induced OWA operator. Expert Systems with Applications,* 38(9), 11560-11572.

[12] Merigó, J., Gil-Lafuente, A. (2009), *The induced generalized OWA operator. Information Sciences,* 179(6), 729-741.

[13] Merigó, J., Guillen, M., Sarabia, J. (2015), *The ordered weighted average in the variance and the covariance. International Journal of Intelligent Systems,* 30(9), 985-1005.

[14] Moreira, D. (2016), *The microeconomic impact on growth of SMEs when the access to finance widens: evidence from internet & high-tech industry. Procedia - Social and Behavioral Sciences,* 220, 278-287.

[15] Schmidt, A., Finan, C. (2018), *Linear regression and the normality assumption. Journal of Clinical Epidemiology, 98*, 146-151.

[16] Seber, G., Lee, A. (2003), *Linear regression analysis.* New Jesey: John Wiley and Sons, SUA.

[17] Yager, R. (1988), *On ardered weighted averaging aggregation operators in multi-criteria decision making. IEEE Transactions Systems. Man Cybernetics B,* 18, 183-190.

[18] Yager, R. (1988), *On ordered weighted averaging aggregation operators in multi-criteria decision making. IEEE Tranactions System Man Cybernetics B,* 18, 183-190.

[19] Yager, R. (1988), *On ordered weighted averaging aggregation operators in multi-criteria decision making. IEEE Transactions on Systems, Man and Cybernetics B, 18,* 183-190.

[20] Yager, R. (1992), *On generalized measures of realization in uncertain environments. Theory and Decision,* 33, 41-69.

[21] Yager, R. (1996), *Constrained OWA aggregation*. *Fuzzy Sets and Systems,* 81, 89-101.

[22] Yager, R. (2002), *Heavy OWA operators*. *Fuzzy Optimization and Decision Making,* 1, 379-297.

[23] Yager, R. (2003), *Induced aggregation operators*. *Fuzzy Sets and Systems,* 137, 59-69.

[24] Yager, R. (2004), *Generalized OWA aggregation operators*. *Fuzzy Optimization and Decision Making,* 3, 93-107.

[25] Yager, R., Filev, D. (1999), *Induced ordered weighted averaging operators*. *IEEE Trans Syst Man Cybern B,* 29, 141-150.

[26] Yager, R., Engemann, K., Filev, D. (1995), *On the concept of immediate probabilities*. *International Journal of Intelligent Systems,* 10, 373-397.