Estimating the Income Distribution of Some Muslim Countries Based on Entropy Measures

Omar Abdul Rahman Kittaneh
Assistant Professor Natural Sciences, Mathematics, and Technology Unit, Effat University, Jeddah, Saudi Arabia

ABSTRACT. Having biased and unjustified income distribution in a country can have many unpleasant consequences such as, poor health, frustration and dissatisfaction, increase in social problems and criminal activities, unreliable education systems, etc. Therefore, determining the income distribution of a country is of great importance and benefit for social safety and economic growth. Using real data for the year 2016 for three Muslim countries, Saudi Arabia (KSA), United Arab Emirates (UAE) and Egypt, this paper utilizes the so-called entropy measures in the problem of estimating the income probability distribution, using the least available information or incomplete information such as the average and range of income. The paper also provides rough estimates of the Coefficient of Variation (CV) and Gini Index for each country.

Keywords: Income distribution, Income inequality metrics, Entropy, Maximum entropy principle, Coefficient of variation, Gini coefficient.

JEL Classification: O15, D63, C61

KAUJIE Classification: M0
1. Introduction

The income distribution of a country plays an important role in describing social stratification. The main reasons that affect income distribution are: access to high quality education (Dahill-Brown, Witte, & Wolfe, 2016), medical care (Klavus & Hakkinen, 1996), basic needs (Heerink & Folmer, 1994), equality of opportunity for advancement (Corak, 2013), social mobility (Kearney & Levine, 2014), government policies and regulations and their enforcement (Wu, Perloff, & Golan, 2002), corrupt practices (You & Khagram, 2005), and the tax system and its enforcement (Adam, Kammas, & Lapatinas, 2015). The income distribution of a certain country is the way the national income is distributed among the population or segments of the society of that country. If everyone has the same share of income, then the income distribution is perfectly equal. If the income is concentrated on one person or one segment of the society, then the income distribution is perfectly unequal. In practice, a country’s income distribution is between complete equality and complete inequality (O’Sullivan & Sheffrin, 2003). Correct estimation of income distribution and the corresponding inequality indices are of great importance as they are sources of information about the structure of the population income (Cowell & Flachaire 2007). Such data could be the base of any considerations related to taxation policies, governmental aid programs, the amount of savings in societies, and zakāh distributions in Muslim countries in particular (Abdul Latif, 1998).

Income distributions are usually estimated by assuming the model of the distribution (to be Normal, Pareto, Lognormal, Beta, etc.), and applying surveys and sampling techniques to estimate the parameters of the suggested model. The most commonly used models for income distribution are the right-skewed models which include, Pareto, Lognormal, Weibull, and Gamma distributions. For example, Chotikapanich, Valenzuela, and Rao (1997) studied Gini inequality coefficients for several countries. Their work was based on the assumption that the income distribution in each country follows a lognormal distribution. Griffiths, Chotikapanich, and Rao (2005), estimated the Australian income distribution using the Bayesian approach based on seven different models. In more recent work, Chotikapanich, Griffiths, Rao, and Valencia (2012) used beta distribution as a model of the income distribution. Bricker, Henriques, Krimmel, and Sabelhaus (2016) estimated the income density and the Wealth Shares of US using administrative and survey data. However, these methods are costly and time-consuming, and sometimes the required information to conduct such methods may not be available. Therefore, there is a need to estimate such distributions based on few given information such as the average income and without any further assumptions. Estimating the income distribution using entropy measures has not been thoroughly featured in literature and most of the works used simulated and not real data (as in Wu, 2003).

In the current paper, the Shannon entropy measure (Shannon, 1948, p. 2) is utilized in estimating the income distribution per household in three Muslim countries, KSA, UAE and Egypt. This is done based on real data of the average monthly income for the year 2016 and the range of wages in the three countries in that year. In addition to this, estimations of Gini index and coefficient of variation are also provided for the three countries based on the estimated income distributions. The reason for choosing these three countries is that they have different economic features and population structures. Egypt has very low income and negligible expat population. KSA has very high income and moderate percentage of expats around 33%. UAE has very high income with excessive expat population of 88%. The diversity in the three economies would enable us to properly test the proposed method and enrich the discussions and conclusions.

2. The Maximum Entropy Distribution

Shannon (1948, p. 2) was the first to introduce entropy, known as Shannon information measure, into information theory. For a discrete random variable \( X \) with support \( \mathcal{X} = \{ x_1, x_2, \ldots, x_n \} \), and probability mass function \( p_i = \Pr\{ X = x_i \} \), \( x_i \in \mathcal{X} \), Shannon entropy is defined as

\[
S(X) = S_e(p_1, p_2, \ldots, p_n) = -\sum_{i=1}^{n} p_i \log p_i = -E(\log p(X))
\]

(1)

The properties and virtues of \( S(X) \) have been thoroughly investigated by Shannon (1948) and Wiener (1961). The measure is non-negative, attains its max-
imum when \( p_i \)'s are equal (uniform distribution), vanishes for the degenerate distributions, and finally, is additive. Consequently, Shannon entropy became important in quantifying randomness present in several fields such as financial analysis (Sharpe, 1985), data compression (Salomon, 2008), statistics (Kullback, 1959), and information theory and data transmission (Cover & Thomas, 1991). Recently, modified versions of this measure have been utilized in lifetime studies (Kittaneh & Akbar, 2016) and image processing (Kittaneh, Khan, Akbar, & Bayoud, 2016).

Jaynes (1957) was the first to introduce the principle of maximum entropy and the induced maximum entropy distribution. The concept was extended by Kullback (1959), Levine (1979) and (Levine, 1980). The maximum entropy distribution is the unique and most probable probability distribution that maximizes Shannon entropy subject to some constraints such as the population average as in this work. In fact, it is a powerful technique that is used to estimate the probabilities in a probability distribution given that all what we know is the expected value of some function(s). In other words, the technique of maximum entropy determines the most uniform probability distribution that matches the observations. The formula of the maximum entropy distribution can be obtained by maximizing Shannon entropy (equation 1) (Cover & Thomas, 1991, pp. 409-411) subject to the average constraint \( E(X) = \mu \). This can be done by the method of Lagrange multipliers, that is, by partially differentiating the following operator with respect to \( p_i \) :

\[
L = -\sum_{i=1}^{n} p_i \log p_i + A\left(\sum_{i=1}^{n} p_i - 1\right) + B\left(\sum_{i=1}^{n} x_i p_i - \mu\right)
\]

After equating the partial derivative to zero and simplifying the result we get the desired distribution

\[
p_i = \frac{e^{\lambda x_i}}{\sum_{i=1}^{n} e^{\lambda x_i}},
\]

where the parameter \( \lambda \) is chosen so that \( \sum_{i=1}^{n} x_i p_i = \mu \) with \( x_i \)'s are the realizations of the probability distribution under consideration and \( n \) is the number of those realizations.

3. Estimating the Income Distribution in UAE, KSA and Egypt

The average income of a country is defined to be the total income divided by the total population. The average monthly income per household can be obtained using the estimated data available in some formal websites such as the World Bank\(^{(1)}\) or other websites and surveys such as WorldData.info\(^{(2)}\). All such sources give close and comparable estimations to this important quantity. The estimated values of the average monthly income per household in KSA, UAE, and Egypt are \( \mu = $1669, $3135, \) and \( $293, \) respectively. In order to get appropriate estimation of the income distributions of the three countries, we need to have reasonable estimations of the maximum and minimum family monthly income. Unfortunately, these figures are not officially released; instead we will use some representatives of the maximum and minimum wage classes recorded in some websites such as Trading Economics\(^{(3)}\) or surveys\(^{(4)}\). Again, reasonable estimations of the representatives of the maximum wage classes in KSA, UAE and Egypt are approximately \( Max = $39643, $35085, \) and \( $5025, \) respectively, whereas, the minimum class representatives are \( Min = $666, $300, \) and \( $130, \) respectively.

In this study, the income distributions in all the three countries are divided into five classes or segments: very low income, low income, average income, high income and very high income. The following values are respectively chosen as representatives of each class: \( Min, \frac{Min+\mu}{2}, \mu, \frac{\mu+Max}{2}, Max, \) where \( Min, \mu \) and \( Max \) are the minimum, average and maximum incomes, respectively.

Basing on formula (3) and the above available information, we estimate the income distributions in

---

(3) Trading Economics, 2018: https://tradingeconomics.com/countries
the three countries using MATLAB 7 software. The results are given in Table 1.

It is easy to observe that the three income distributions are right-skewed, and this means that in all cases, the median is between the mode and the mean. For comparison purposes, the three income distributions are displayed in Figure 1. It is clear from the figure that the three distributions start at almost the same point (the very low-income class) then strictly decrease with a big jump when moving from the average-income class into the high-income class. This means that the majority of the population in each of the three countries are on or below the average. Although the income distributions of KSA and Egypt have similar behaviors, they have different variability and inequality indices as will be seen in sections five and six of this paper. Also, it can be observed that the curve of UAE is above the other two curves at the high and very high classes and it is below them otherwise.

<table>
<thead>
<tr>
<th>Class($)</th>
<th>Very Low</th>
<th>Low</th>
<th>Average</th>
<th>High</th>
<th>Very High</th>
</tr>
</thead>
<tbody>
<tr>
<td>KSA</td>
<td>666</td>
<td>1167.5</td>
<td>1669</td>
<td>20656</td>
<td>39643</td>
</tr>
<tr>
<td>percentage</td>
<td>34.86%</td>
<td>32.59%</td>
<td>30.03%</td>
<td>2.34%</td>
<td>0.18%</td>
</tr>
<tr>
<td>UAE</td>
<td>300</td>
<td>1717.5</td>
<td>3135</td>
<td>19110</td>
<td>35085</td>
</tr>
<tr>
<td>percentage</td>
<td>34.94%</td>
<td>30.65%</td>
<td>26.88%</td>
<td>6.13%</td>
<td>1.40%</td>
</tr>
<tr>
<td>Egypt</td>
<td>130</td>
<td>206.5</td>
<td>293</td>
<td>2646.5</td>
<td>5000</td>
</tr>
<tr>
<td>percentage</td>
<td>34.87%</td>
<td>32.18%</td>
<td>29.71%</td>
<td>2.95%</td>
<td>0.29%</td>
</tr>
</tbody>
</table>

Figure (1) The Income Distributions of KSA, UAE and Egypt

Source: Calculated by author.

4. Income Inequality Metrics

Income inequality metrics are measures of statistical dispersion intended to describe the income or wealth distributions of a nation’s residents. Income inequality metrics should possess the following four fundamental mathematical characteristics that agree with our intuitive understanding of the concept of inequality: symmetric, invariant under scalar multiplication, free of the population size, transferable. Whereas satis-
fying one or more of the following properties is not mandatory but preferable: being non-negative, equal to zero for degenerate distributions, bounded above and/or below. For more details, the reader is referred to (Atkinson, 1970).

The most commonly used inequality metrics are the Gini coefficient (Gini, 1921), the Atkinson index (Atkinson, 1975), the coefficient of variation (Campano & Salvatore, 2006, p. 80), the generalized entropy index (Cowell, 1995), the decile ratios (Gold, Kawachi, Kennedy, Lynch, & Connell, 2001), etc.

The conventional approach in nearly all empirical work is to adopt some summary statistic of inequality metrics with no explicit reason being given for preferring one measure rather than another. In this paper, we select the Gini coefficient and the coefficient of variation to describe the inequality in the aforementioned countries.

5. Coefficient of Variation (CV)

This measure of income variability is the ratio of the standard deviation of the income distribution to its mean. More equal income distributions will have smaller standard deviations; consequently, the CV will be smaller in more equal societies. For comparison purposes, based on the three income distributions that are displayed in Table 1, the standard deviation (σ) and coefficient of variation (CV) are computed and recorded in Table 2. The reason for using the coefficient of variation rather than the standard deviation is that the averages are widely different.

The income distribution in Egypt has the smallest coefficient of variation, also in the coming section we will see that it also has the smallest Gini index value, whereas, although, UAE has a moderate coefficient variation it has the highest Gini index. From Table 1, one would think that the three countries have similar income distributions, but actually, they have many intrinsic differences as we can see from Table 2.

Table 2: μ, σ and CV of the Income Distributions in UAE, KSA and Egypt

<table>
<thead>
<tr>
<th></th>
<th>μ</th>
<th>σ</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>KSA</td>
<td>1669</td>
<td>3385.9</td>
<td>2.03</td>
</tr>
<tr>
<td>UAE</td>
<td>3135</td>
<td>5775.7</td>
<td>1.84</td>
</tr>
<tr>
<td>Egypt</td>
<td>293</td>
<td>489.2</td>
<td>1.67</td>
</tr>
</tbody>
</table>

Source: Calculated by author.

6. Gini Coefficient

Despite the fact that CV is a simple measure of variability which can to some extent interpret inequality, it has not been usually utilized in studies on income inequality. This is due to the limitations of the CV measure as it is not bounded. In contrast to the Gini coefficient, making comparison between different income distributions using CV is more intractable; and also due to the sensitivity of CV as its two components (the mean and the standard deviation) are highly affected by the sudden fluctuation of the income scores and weights. This means that the CV would not be an appropriate measure of inequality unless certain conditions are satisfied such as normality.

Gini coefficient is the most commonly used indicator of income inequality or equality. Gini coefficient of a certain income distribution is the distance (in the area sense) between the equality line and the Lorenz curve. Lorenz curve matches the cumulative percentage of population with the cumulative percentage of income. Gini coefficient describes how far the Lorenz curve is from the equality line of 45°, where this line means complete equality or perfect uniformity. This can be done by dividing the area between the equality line and Lorenz curve by the total area under the equality line. This calculation produces a number between zero and 1, the higher the number, the greater the degree of income inequality. A perfectly equal income distribution will have a Gini coefficient of 0, while a perfectly unequal distribution will have a Gini coefficient of 1.

In our study we divide the population into five segments in terms of how much money they make as shown in Table 1. However, to estimate the Gini coefficient, we have to divide it equally according to the size of the segments and observe the variations of the income. The five segments become: the very low 20%, the low 20%, the average 20%, the high 20%,
and the very high 20%. We can then create a table detailing how much income each segment earns (on average) out of the total amount of income. This can be done by redistributing the data in Table 1 according to the new equal weights as shown in Table 3.

It can be checked that the average values of income in Table 3 are equal to those in Table 1, and the two quantities are equal to the actual averages that we started with for the three countries (the variations between the three answers are only due to rounding).

In order to look at the income distribution, we need to see what percentage of total income each segment has, rather than the actual amount of money each makes. Since the sizes of the segments are equal, we don’t need to worry about weighting the average incomes anymore and can do a straightforward calculation of each segment’s earnings.

### Table (3) The Income Distributions in UAE, KSA and Egypt using equal weights

<table>
<thead>
<tr>
<th></th>
<th>Very Low (20%)</th>
<th>Low (20%)</th>
<th>Average (20%)</th>
<th>High (20%)</th>
<th>Very High (20%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>KSA</strong></td>
<td>666</td>
<td>794.89</td>
<td>1167.5</td>
<td>1482.19</td>
<td>4232.25</td>
</tr>
<tr>
<td><strong>UAE</strong></td>
<td>300</td>
<td>658.63</td>
<td>1717.5</td>
<td>2738.81</td>
<td>10267.84</td>
</tr>
<tr>
<td><strong>Egypt</strong></td>
<td>130</td>
<td>149.62</td>
<td>206.5</td>
<td>262.51</td>
<td>708.39</td>
</tr>
</tbody>
</table>

Source: Calculated by author.

For total income we will use the sum of the five average incomes, and then compute the percentage of total income that each segment of the population earns, by dividing their income by the total income as shown in Table 4.

What Table 4 indicates is that the bottom fifth of the population gets less than 9% of the total income in KSA and Egypt, and less than 2% in UAE, while the top fifth of the population gets over 50%, 65% and 48%, in KSA, UAE and Egypt, respectively, of the total income, indicating a large degree of income inequality. In Table 5, we simply compute the cumulative percentages for income distribution in order to compute the Gini coefficient.

### Table (4) The Income Percentages of the Five Segments in UAE, KSA and Egypt

<table>
<thead>
<tr>
<th></th>
<th>Very Low (20%)</th>
<th>Low (20%)</th>
<th>Average (20%)</th>
<th>High (20%)</th>
<th>Very High (20%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>KSA</strong></td>
<td>7.98%</td>
<td>9.53%</td>
<td>13.99%</td>
<td>17.77%</td>
<td>50.73%</td>
</tr>
<tr>
<td><strong>UAE</strong></td>
<td>1.91%</td>
<td>4.20%</td>
<td>10.95%</td>
<td>17.46%</td>
<td>65.48%</td>
</tr>
<tr>
<td><strong>Egypt</strong></td>
<td>8.92%</td>
<td>10.27%</td>
<td>14.17%</td>
<td>18.02%</td>
<td>48.62%</td>
</tr>
</tbody>
</table>

Source: Calculated by author.

### Table (5) The Cumulative Income Percentages of the Five Segments in UAE, KSA and Egypt

<table>
<thead>
<tr>
<th></th>
<th>Very Low</th>
<th>Low</th>
<th>Average</th>
<th>High</th>
<th>Very High</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Uniform</strong></td>
<td>20%</td>
<td>40%</td>
<td>60%</td>
<td>80%</td>
<td>100%</td>
</tr>
<tr>
<td><strong>KSA</strong></td>
<td>7.98%</td>
<td>17.51%</td>
<td>31.50%</td>
<td>49.27%</td>
<td>100%</td>
</tr>
<tr>
<td><strong>UAE</strong></td>
<td>1.91%</td>
<td>6.11%</td>
<td>17.06%</td>
<td>34.52%</td>
<td>100%</td>
</tr>
<tr>
<td><strong>Egypt</strong></td>
<td>8.92%</td>
<td>19.19%</td>
<td>33.36%</td>
<td>51.38%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Source: Calculated by author.

Figures 2, 3, and 4 show for the three countries, respectively, how much each of the population segments earns (cumulatively), and compare the resulting curve to a perfectly equal income distribution (the equality line).

The Lorenz curves were prepared by fitting the data with appropriate five degree polynomials using MATLAB 7 software and the concerned areas were computed using the classical integration techniques. From the figures below, the Gini coefficient is equal to the area of A divided by the area of (A + B). Based on our estimates, the computed Gini coefficients are 0.3997, 0.5978 and 0.3715 for KSA, UAE and Egypt, respectively. As per the World Bank, the Gini...
coefficient of Egypt was 0.3180 in 2015\(^{(5)}\), which is quite close to our estimate. Unfortunately, there are no official records for the values of the Gini coefficient for many countries, in particular oil countries. However, some sources like Matar (2017, p. 22) reported that the Gini coefficient in KSA was 0.4590 in 2016 which is very close to our estimation, whereas it was 0.5300 in UAE. As per our estimates or the previous estimates, the Gini coefficients of KSA and UAE are higher than that of Egypt, especially in UAE. The UAE is one of the richest countries in the world, with a Gross National Income (GNI) per capita of USD 40,480 for 2016. However, the high inequality in its income distribution is due to the fact that the Emirati nationals are about 10% of the total population and the scale of their salaries and benefits are much higher than those for the 90% from all other nationalities.

\(\text{(5) GINI index (World Bank estimate), 2018: https://data.worldbank.org/indicator/SI.POV.GINI}\)
Figure (2) Lorenz curve of KSA

Figure (3) Lorenz curve of UAE

Source: Prepared by author.
7. Conclusions and Recommendations

This paper suggests an entropy-based method for providing a rough estimate of the monthly family income distributions in KSA, UAE, and Egypt using only reasonable estimates of the average and range of incomes. We establish here on the merits of the Gini index that the UAE has the highest value while Egypt has the lowest ranking. The main distinguishing factor in the case of Egypt as compared to the KSA and the UAE, which justifies such a high level of income inequality, is that Egypt does not have a notable expat population. In both the KSA and the UAE, most of the expat work force is employed as laborers, cashiers, nurses, housemaids, etc. The wages in these professions are negligible as opposed to other higher professions. Based on our calculations, it is therefore recommended that the wages of the low-income class be raised in order to reduce the inequality, particularly in the case of the UAE.

In order to bring down the value of the Gini index, we advocate that the range and the huge gap between the very low and very high income classes be reduced duly. This can be achieved by increasing the salaries of the low and very low-income classes. In a subsequent study, we wish to explore the possibility of bridging such a gap further in the context of Muslim countries by proposing a more efficient method of zakāh distribution.

Having said that, we do not claim the results of this paper to be exhaustively accurate because the resulted probability distributions may often suffer from great disparities due to irregularities in, for example, the number of classes chosen, the reliability and accuracy of the given figures, including income averages and the maximum and minimum bounds of income. However, we believe that the procedure presented in this paper is highly significant in the sense that it offers a primary conception and visualization of the income distribution based on the minimum available information. In addition, the overall precision of the procedure presented in this paper can be conveniently substantiated by drawing a comparison between the estimated values of the Gini index based on the proposed method and those estimated by the World Bank and other references cited here.
References


**Omar Abdul Rahman Kittaneh** is an assistant professor at Effat University, Jeddah, Saudi Arabia. He is also the director of the Natural Sciences, Mathematics and Technology Unit at the same university. He earned his PhD from the University of Jordan in 2008. His field of specialty is mathematical statistics. His research interests include: statistical testing, information theory, reliability and image processing, among others. He has published many papers in international refereed journals such as, *the IMA Journal of Mathematical Control and Information*, *Oxford University Press, Communications in Statistics – Theory and Methods Journal, Taylor & Francis, The American Statistician, American Statistical Association*, among others.

E-mail: okitanneh@effatuniversity.edu.sa
قدير توزيع الدخل لبعض الدول الإسلامية بناءً على مقاييس عشوائية

عمرو عبد الرحمن كاتنة
أستاذ مساعد، وحدة العلوم الطبيعية والرياضيات والتكنولوجيا
جامعة عفت، جدة، المملكة العربية السعودية

 المستخلص. هناك نتائج وتبعات سيئة لتوزيع الدخل غير العادل وغير المبرر في أي دولة مثلا: تدني مستوى الصحة، والإحباط النفسي، وعدم الرضا، وتفاقم المشكلات الاجتماعية وزيادة معدلات الجريمة، علاوة على فقدان الثقة في نظم التعليم ومخرجاته، من هنا تكم أهمية عدالة توزيع الدخل وانعكاسها على صحة المجتمع ونموه الاقتصادي. يستخدم البحث بيانات حقيقية لتوزيع الدخل وقياساتها في عام 2016م في ثلاث دول إسلامية هي المملكة العربية السعودية والإمارات العربية المتحدة وجمهورية مصر العربية. ويعتبر البحث مقاييس عشوائية لتقدير توزيع الدخل، بناء على الحد الأدنى المتوافر من المعلومات، أو معلومات غير كاملة مثل متوسط ونطاق الدخل. كما يقدم البحث تقديرات تقريبية لمعامل التغير (CV) ومؤشر Gini لكل دولة من تلك الدول.

الكلمات الدلالة: توزيع الدخل، مقاييس عدم مساواة الدخل، الانترنوبيا، مبدأ الانترنوبيا الفصوى، Gini، معامل التباين، مؤشر

O15, D63, C61: JEL

 تصنيف KAUJIE: M0