

**ONE BAD FORMULA CAN SPOIL EVERYTHING:  
A SIMPLE ADJUSTMENT THAT WOULD IMPROVE THE  
UN'S GENDER INEQUALITY INDEX**

MAX MCDONALD AND NEAL KOBLITZ

A simple change in one part of the definition of the United Nations Gender Inequality Index (GII) could go a long way toward making it less problematic and more useful. In the definition of the GII we propose to replace the quantity

$$a_F = \sqrt{\frac{1}{mt}},$$

where  $m$  is the maternal mortality rate and  $t$  is the teen pregnancy rate, by

$$A_F = \sqrt{\frac{C_1}{m + C_1} \cdot \frac{C_2}{t + C_2}},$$

where  $C_1$  and  $C_2$  are constants. This alternative formula is a natural one, both because of how it combines with the other factors that go into the GII and because it is similar to the way that the problem of near-zero denominators is handled in other areas of applied mathematics. The precise meaning of  $t$  and  $m$  and suggested values for  $C_1$  and  $C_2$  will be given below. We will explain the computation of the GII, discuss the other three indicators that it is based on, and show that the simple change from  $a_F$  to  $A_F$  would have useful consequences without altering the conceptual underpinning of the GII.

There can be dangers as well as benefits when mathematical methods are used to answer questions that have a large human component. In the middle of the last century Darrell Huff [5] wrote of the myriad ways that statistics are used in marketing and politics to mislead the public. In our time Felix Salmon [14] has described how David Li's mathematical model for collateralized debt obligations was seized upon by Wall Street as a justification for irresponsible investment practices that precipitated the financial crisis of 2008; John Ewing [3] has decried the misuse of mathematics to evaluate teachers in his exposé of value-added models in the *AMS Notices*; and Cathy O'Neil [12] has explained many of the ways that data science has been used to undermine democratic values — for example, by reinstituting a type of racial discrimination in lending (called “redlining”) that had supposedly been outlawed in the U.S. fifty years ago.

Even people with the best of intentions who use mathematical methods to help achieve worthy goals can easily be led astray. A single unfortunate

---

*Date:* 9 February 2019.

choice of a formula can lead to results that undermine these objectives. Such is the case with the United Nation’s Gender Inequality Index (GII).

## 1. MEASURING GENDER INEQUALITY

For the last two or three decades, international private and governmental aid organizations have been increasingly interested in finding quantitative tools to guide policy-making about allocation of resources and to help evaluate the effectiveness of ongoing projects. Among the various metrics and indexes that have been developed, the ones that carry the imprimatur of the United Nations are especially influential.

In 1990 the United Nations Development Program (UNDP) launched its annual *Human Development Report*, containing the country-by-country Human Development Index (HDI). Created by the prominent Pakistani economist Mahbub ul Haq, Nobel prize-winner Amartya Kumar Sen, and other economists, its object was to influence policy-makers to shift their focus from measures of wealth to human-centered measures of well-being.

Five years later, along with the HDI the UNDP started including two gender-related indexes in the *Report*. Although many welcomed this as an overdue recognition of the central importance of women’s rights in discussions of national progress, the two indexes were widely criticized for being too closely tied to the overall economic level of a country, and for giving the impression that economic development is the key to women’s equality, without any need for gender-specific efforts.

In response to these criticisms, in 2010 the UNDP introduced a new Gender Inequality Index (GII) that was not tied so directly to economic metrics. A central objective of the GII was to convey the message that achieving gender equity requires dedicated programs and resources; the problem of discrimination and mistreatment of women is not automatically solved by improving a country’s economic indicators.

## 2. CONSTRUCTION OF THE GII

The GII is a composite index that uses some fairly complicated formulas — involving the geometric, arithmetic, and harmonic means — to combine five indicators into a single number. Two indicators relate to *reproductive health* — the maternal mortality rate (per 10,000 births) and the teen pregnancy rate (per 1,000 females age 15 to 19). The next two indicators are related to *empowerment* — the proportions of adult women and men who have completed some secondary schooling, and the proportions of women and men in the national parliament. The fifth indicator is the proportions of adult women and men in the labor force.

Let  $a_F$ ,  $b_F$ ,  $c_F$  for women and  $a_M$ ,  $b_M$ ,  $c_M$  for men denote the three components *reproductive health*, *empowerment*, and *labor force participation*, respectively. The UNDP computes them as follows. Proceeding from right to left,  $c_F$  and  $c_M$  are the respective proportions (percent written as a

decimal) of adult women and men participating in the workforce. Next, each of the numbers  $b_F$  and  $b_M$  is the *geometric* mean of the proportion of the corresponding gender in parliament and the proportion (age 25 or older) that has had some secondary schooling. The number  $a_M$  is set equal to 1. We'll discuss  $a_F$  after describing the UNDP's computation of the GII.

Let  $a_{av}$  be the *arithmetic* mean  $(a_F + a_M)/2$ , and similarly let  $b_{av} = (b_F + b_M)/2$  and  $c_{av} = (c_F + c_M)/2$ . Let  $G_F = \sqrt[3]{a_F \cdot b_F \cdot c_F}$  and define  $G_M$  similarly. Let  $H$  be the *harmonic* mean of  $G_F$  and  $G_M$ , defined as  $2/(1/G_F + 1/G_M)$ . Finally, the GII is defined to be

$$(1) \quad 1 - \left( \frac{H}{\sqrt[3]{a_{av} \cdot b_{av} \cdot c_{av}}} \right).$$

### 3. A PROBLEMATIC FORMULA

It's the formula for  $a_F$  that causes huge problems. It is composed of two indicators: maternal mortalities  $m$  (per 10,000 births) and teen pregnancies  $t$  (per 1,000 teenage girls). Unlike the other numbers that go into the GII, these are not proportions between 0 and 1. After taking reciprocals (and replacing  $m$  by 1 if  $m < 1$ ) these indicators become positive numbers less than or equal to 1. The geometric mean of those reciprocals is the reproductive-health component  $a_F$ ; thus,

$$(2) \quad a_F = \frac{1}{\sqrt{mt}}.$$

The Technical Notes in [15] explain that it is in order to avoid giving countries with  $m < 1$  an undue advantage that the value of  $m$  is truncated at 1, that is, 1 is the least value that is substituted for  $m$  in (2). No truncation is specified for  $t$ , but since no country has  $t < 1$ , in practice no truncation of  $t$  is necessary. The explanation given in the Technical Notes for the truncation of  $m$  at 1 maternal death per 10,000 births is that at the low end the difference between 1 and an even lower rate does not reflect a significant difference in medical care for pregnant women. That is, one wants to avoid giving Portugal (which has  $m = 1.0$ ) the penalty coming from a multiplicative factor of 2.5 compared to Kuwait (which has  $m = 0.4$ ). (We shall achieve this without truncation by modifying the formula (2).)

It is important, however, to note that even with the truncation  $a_F$  is very sensitive to differences in  $m$  at the low range. For example, in comparing Portugal with a country such as Costa Rica that has 2.5 maternal deaths per 10,000 births, the difference *does* result in a penalty factor of 2.5 in the value of  $m$  for Costa Rica, despite the fact that 2.5 deaths per 10,000 births is quite low for a low-income country (less than 1/8 the worldwide average maternal mortality rate). A factor of 2.5 in  $m$  produces a lowering of  $a_F$  by a factor of  $0.63 \approx 1/\sqrt{2.5}$ . In order to produce a similar penalty in either of the other two components  $b_F$  or  $c_F$  there would have to be a large

and meaningful gender disparity. Thus, a problem with the mathematical formula for  $a_F$  is that it is far too sensitive to increases in small values of  $m$  and  $t$ .

Another deficiency in the formula (2) is that the values of  $a_F$  are clustered near zero, especially for the low-income countries. For example, if a country has the worldwide averages  $m = 21.6$  and  $t = 49$  [18, 20], then its value for  $a_F$  is 0.031. Values near zero lead to an exaggerated impact on the GII. In contrast, the impact of the labor force participation component is much less, because the ratio of  $c_F$  to  $c_M$  is never close to zero. Neither is the ratio of  $b_F$  to  $b_M$ , except in the unusual case when there are 0 women in parliament (in which case the UNDP truncates the political empowerment indicator at 0.001). Thus, in general, it is because of the formula (2) that the reproductive-health component has a disproportionate effect on the GII.

#### 4. REPAIRING THE FORMULA

The following change in the definition of  $a_F$  fixes these problems. Let  $m_W$  be the worldwide maternal mortality (per 10,000), and let  $t_W$  be the worldwide teen pregnancies (per 1,000), where we take these values at the baseline years 2015 and 2014, respectively. That is, using the figures in [18, 20], we take  $m_W = 21.6$  and  $t_W = 49$ . We now define

$$(3) \quad A_F = \sqrt{\frac{m_W}{m + m_W} \cdot \frac{t_W}{t + t_W}} = \sqrt{\frac{21.6}{m + 21.6} \cdot \frac{49}{t + 49}}.$$

What makes this modification seem natural to us is a loose analogy with the technique of *Tikhonov regularization* [7] in the theory of inverse problems. That method is used, for example, when one has to invert a matrix that may have very small eigenvalues, and one wants to avoid outliers caused by near-zero denominators. Our setting is a particularly simple, one-dimensional one, in which our analogous approach is to replace  $m$  by  $m + \text{const}_1$  and  $t$  by  $t + \text{const}_2$  in (2) before taking the inverses and the squareroot:

$$\sqrt{\frac{1}{m + \text{const}_1} \cdot \frac{1}{t + \text{const}_2}}.$$

However, to make the values range between 0 and 1, we need to replace 1 by  $\text{const}_1$  and  $\text{const}_2$  in the numerators. In addition, it's natural to choose the constants to be the worldwide average values (at a fixed point in time, such as 2016, so as to preserve uniformity from year to year), so that for a country with the average values  $A_F$  will equal 0.5. This leads to the formula (3).

Replacing  $a_F$  by  $A_F$  has several advantages:

- It is simple and does not require any arbitrary choice of truncation value.

- It does not depend on the choice of units for the mortality or pregnancy rates. In other words, if, for example, we changed  $m$  and  $m_W$  to number per million, the value of  $A_F$  would remain the same, whereas the value of  $a_F$  would decrease by a factor of 10.

- It reduces the disproportionate effect that  $a_F$  had on the composite index.

- Each factor under the radical in (3) is near 0 when  $m$  or  $t$  is large compared to the worldwide average and is near 1 when  $m$  or  $t$  is small. The values range steadily from 0 to 1 as  $m$  or  $t$  improves. In particular, when  $m$  and  $t$  are both small compared to the worldwide averages,  $A_F$  is near 1, which is the value always given to men's  $a_M$ . This contrasts with  $a_F$ , which is not close to 1 unless both  $m$  and  $t$  are very close to 1, and there is no country in the world with teen pregnancies per 1000 very close to 1.

We remark that arguments could be made to justify other choices for the constants  $C_1$  and  $C_2$ . For example, instead of world averages one could take aspirational values, such as the averages for the 25 countries with the lowest  $m$  or  $t$  values. If we did this, we would find (using data from 2015) that a country with average  $m$  and  $t$  would now have  $A_F = 0.165$  rather than 0.5. Most low-income countries would have  $A_F$  quite small, and this would have a major impact on the GII for those countries. In general, the effect on country rankings of using much lower values  $C_1 \ll m_W$  and  $C_2 \ll t_W$  would be to reduce the changes in rankings that occur when the original GII using  $a_F$  is replaced by the modified GII using  $A_F$ .

We next consider the effect of replacing  $a_F$  by  $A_F$  (with  $C_1 = m_W$  and  $C_2 = t_w$ ) but making no other change in the GII. The consequences of this simple replacement of an inadequate formula by a more reasonable one can be dramatic.

## 5. DOES CUBA *Really* HAVE GREATER GENDER INEQUALITY THAN THE US?

Let's compare the US and Cuba before and after our proposed modification. The input to the computations consists of  $m$ ,  $t$ , PRF and PRM (the percent of women and men in parliament), SEF and SEM (the percent of women and men with some secondary schooling), and LFF and LFM (the percent of women and men in the labor force). The values from [16] are given in Table 1.

In Table 1 we see that Cuba's PRF is 2 1/2 times as great as the PRF of the US. However, because of the formula  $a_F = 1/\sqrt{mt}$ , this advantage is wiped out by Cuba's 5.6 times greater value of  $m$  times  $t$  – despite the fact that Cuba's maternal mortality and teen pregnancy numbers are not high for a low-income country, whereas according to [9] the US has the highest maternal mortality rate in the developed world. Thus, the formula for  $a_F$  causes the major advantage Cuba has in women's equality in parliament to

TABLE 1. GII indicators for the US and Cuba.

	m	t	PRF	PRM	SEF	SEM	LFF	LFM
U.S.	1.4	22.6	19.5%	80.5%	95.4%	95.1%	56.0%	68.4%
Cuba	3.9	45.6	48.9%	51.1%	83.9%	86.7%	42.6%	68.6%

TABLE 2. Computation of the GII for the US and Cuba.

	$(a_F)$ $A_F$	$b_F$	$c_F$	$a_M$	$b_M$	$c_M$	$(a_{av})$ $A_{av}$
U.S.	(0.178) 0.802	0.431	0.560	1	0.875	0.684	(0.589) 0.901
Cuba	(0.075) 0.662	0.641	0.426	1	0.666	0.686	(0.538) 0.831

	$b_{av}$	$c_{av}$	$(G_F)$ $G_F$	$G_M$	(H) H	(GII) GII
U.S.	0.653	0.622	(0.350) 0.579	0.843	(0.495) 0.686	(0.203) 0.041
Cuba	0.653	0.556	(0.274) 0.565	0.770	(0.404) 0.652	(0.304) 0.028

be negated by reproductive health statistics that, if one takes into account the tremendous difference in wealth of the two countries, do not reflect poorly on Cuba or well on the US.

In Table 2 we show the computation of the GII for the US and Cuba both before and after our proposed adjustment in the formula for the reproductive health component. In the table the numbers in parentheses are the old numbers, based on  $a_F$ , and the numbers not in parentheses result from replacing  $a_F$  by  $A_F$ .

With the UNDP formula (2) for  $a_F$ , Cuba's GII was 50 percent worse than that of the US; now with our modified formula (3) it's 30 percent better!

## 6. REMARKS

First, our critique of the GII is based on technical considerations, and, as mentioned in the introduction, we are not concerning ourselves with broader issues. Among such issues would be the conspicuous absence from the reproductive health component of statistics on availability of abortion and other birth control services. Nor is the gender wage gap included in the workforce component. The five indicators that go into the GII are not adequate proxies for these measures of gender inequality, which, if included, would undoubtedly favor Cuba over the United States.

Second, we should not be surprised that for both countries the modified GII is a small number. That does not indicate that the US and Cuba have no significant gender inequality – quite the contrary. We have to remember that the way the GII measures inequality is to compute the deviation of the harmonic mean from the arithmetic mean. To illustrate this in a simplified situation with only one indicator, suppose that a certain high-tech company

pays its female employees who have a Ph.D. an average salary  $s_F$ =\$160,000 and its male employees with a Ph.D. an average salary  $s_M$ =\$240,000. The measure of inequality would then be

$$1 - \frac{((1/s_F + 1/s_M)/2)^{-1}}{(s_F + s_M)/2} = 1 - \frac{4s_F s_M}{(s_F + s_M)^2} = 0.040.$$

However, in this illustrative example the company is paying men 50 percent more than women! Because a small value for the modified GII could easily be misinterpreted as indicating almost no gender inequality, it might be wise to simply scale these measures by a factor of 100. Thus, our hypothetical company would have a measure of inequality 4.0 rather than 0.040.

By replacing  $a_F$  by  $A_F$ , we have removed the disproportionate impact of  $m$  and  $t$  on the GII, thereby bringing the actual numbers down to a level that's similar to the inequality measure 0.040 in the above example. Those numbers should not be misinterpreted as an indication that the gender inequality is small. What is interesting is the rankings, not the actual numbers. (By "ranking" we mean a country's order with respect to GII, from lowest to highest inequality index.)

Third, it should be noted that the inordinate impact of  $m$  and  $t$  on the GII affects not only the poor countries, but even comparisons among the wealthy countries. For example, in [16] South Korea has a high GII ranking (10th), above Canada (18th) and France (19th). The reason is that South Korea's reproductive health indicator  $a_F \approx 0.75$  is the largest (closest to 1) of any country in the world. However, abortion there is illegal in most circumstances, and the rate of illegal abortion is estimated to be one of the highest in the world [2].

Let us also compare rank-1 Switzerland with a GII of 0.040 and rank-5 Iceland with a GII of 0.051 (before our modification of the GII). Looking at the five indicators that go into the GII, we see that Iceland's parliament has roughly a 50 percent greater proportion of women than Switzerland's, Iceland has a slightly higher proportion of women than men with secondary schooling (whereas it's the reverse in Switzerland), and it has 10 percent more men than women in the labor force (as opposed to 20 percent more men than women in the labor force in Switzerland). The maternal mortality per 10,000 is 0.5 in Switzerland and 0.3 in Iceland, but both are defaulted to 1 because of the truncation. Why, then, does Switzerland come out substantially ahead of Iceland in the GII? It's entirely because of the  $t$ -values, which are 6.1 for Iceland and 2.9 for Switzerland. But 6.1 teen pregnancies per 1000 women age 15–19 is a low figure – the US figure is 3.7 times as great – and it strains credulity to interpret that teen pregnancy rate as a significant indication of gender inequality in Iceland.

Meanwhile, if we use our adjusted formula with  $a_F$  replaced by  $A_F$ , the comparison between these two countries changes radically: the adjusted GII for Switzerland is now 0.016 and for Iceland is 0.003. While there is no denying that Switzerland has come a long way since 1991, when it was

the last country in Europe to allow women to vote in all elections, it is hard to imagine that any reasonable methodology would put it first in the world in women's equality. It is interesting to note that the World Economic Forum [19] ranks Iceland first in the world in closing the gender gap, whereas Switzerland's rank is 21.

What we have seen in the comparisons between the US and Cuba and between Switzerland and Iceland holds more generally. In Table 3 we compare the correlations of the GII with its three components, both before and after the replacement of  $a_F$  by  $A_F$ . The reproductive health component of the GII given by the formula (2) has an undue influence, whereas the three components are much more evenly balanced in the revised GII using the formula (3).

## 7. EARLIER CRITIQUES OF THE GII

While the GII was broadly recognized as a major improvement over the UNDP's earlier gender-related indexes, several authors [1, 10, 13] criticized the construction of the GII on various grounds — most notably, for having a strong bias in favor of the richer countries. Of the five indicators, only the fourth one (representation in parliament) is free of this bias. Maternal mortality would be expected to correlate more strongly with resource shortages in the health care system than with actual gender inequality. Pregnancy before age 20 would naturally be higher in countries with large rural populations, where early marriages are more common than in the wealthy countries. Educational inequality in the wealthy countries would not be so much at the secondary level as at the university level and beyond, and inequality at that level is not reflected in the GII. And because labor force statistics usually do not include the disproportionately female informal sectors of the economy — which are vitally important in many developing countries — the labor force indicator is also skewed against the poorer countries.

Some of these authors [13, 10] have also objected to the mixing of comparative and absolute statistics. Permanyer [13] writes that “the inclusion of indicators that compare the relative performance of women vis-à-vis men together with absolute women-specific indicators obscures even more the interpretation of an already complicated index that turns out to penalize the performance of low-income countries.” He claims that the reproductive-health component of the GII is largely responsible for the high negative correlation between a country's GII and its Gross Domestic Product (GDP) per capita, presumably because a good health care system is necessary in order to get a low inequality index. He computes the correlation between the GII and the logarithm of the GDP per capita using data from [15] to be  $-0.87$ , and then observes that if the GII were computed based only on the other three indicators, the correlation would be just  $-0.34$ .

We disagree with this general criticism of the reproductive health component. We do not see anything wrong with mixing comparative indicators



TABLE 3. Pearson correlation coefficients of the GII with its three components.

	with $a_F$ or $A_F$	with $b_F/b_M$	with $c_F/c_M$
Original GII	(with $a_F$ ) -0.902	-0.511	-0.1682
Revised GII	(with $A_F$ ) -0.545	-0.768	-0.4577

with women-specific ones (with the male side of the latter indicators set equal to 1.0, corresponding to ideal circumstances). It can be argued that, since men are free of this threat to well-being, any shortfall in a country's reproductive health (causing the female side of the indicator to fall below 1.0) should indeed count toward gender inequality, just as in the case of the relative-performance indicators.

Moreover, Permanyer is incorrect when he blames the inclusion of the women-specific indicators for the GII's bias in favor of the wealthy countries. Rather, we pinpointed the main source of the distortions not in the inclusion of the reproductive health indicators as such, but rather in the particular way in which they were incorporated into the mathematical formulas defining the GII.

Permanyer [13] writes:

As a matter of fact, if some male-specific variables are fixed at the “perfect achievement level” in a gender inequality measure that is supposed to capture the relative performance of women vis-à-vis men, the values of the aggregate index will naturally tend to be dominated precisely by those variables (which will tend to have larger gaps than the others).

However, our revised GII still includes the reproductive-health component with male achievement level set at  $a_M = 1$ . Thus, contrary to what Permanyer suggests, it is not the fact that  $a_M = 1$  that caused the domination of the GII by that component. Table 3 shows that the problem can be solved by changing the formula (2), with no need to remove the reproductive health component or to change the value  $a_M = 1$ .

## 8. PROBLEMS WITH THE EMPOWERMENT COMPONENT

The replacement of  $a_F$  by  $A_F$  by no means completely solves the problem of the overall bias toward the wealthier countries. The GII is tabulated and is meant to be used for all countries, not just low-income countries, and so it should not ignore major sources of gender inequality in the wealthy countries. As part of the empowerment component the GII looks at the proportion of women and men who have had some secondary schooling. In most of the wealthy countries almost all youngsters of both sexes reach at least that level, so no significant inequality is detected. We would get a very different picture of the extent of inequality in the high-income countries if

TABLE 4. Correlations between values and between rankings.

Comparisons	Correlation of values	Correlation of rankings
Revised vs Original GII	0.754	0.841
Original GII vs $\ln(\text{GNI per capita})$	-0.802	-0.828
Revised GII vs $\ln(\text{GNI per capita})$	-0.462	-0.601

we looked, for example, at the proportion of women among full professors, among researchers with Ph.D. in the STEM fields, or among lab and institute directors.

The only area where the GII measures empowerment at a high level is the political arena. And, in fact, it is in the PRF/PRM where the most inequality is detected for the US (19.5 percent female, 80.5 percent male). If the GII had taken a much lower level of political empowerment – for example, voter turnout on election day – little or no gender disparity would have been detected in political power in the US. By the same token, the GII fails to detect the significant gender inequality in educational empowerment that exists in the developed world because the secondary-school metric is all that it uses.

## 9. NEW RANKINGS

In our revised GII the only change we made was to replace  $a_F$  by  $A_F$  (see formulas (2) and (3)). The most dramatic differences between the original GII rankings and our revised ones tend to be seen among African and Latin American countries (in a positive direction) and among Mideastern countries (in a negative direction). Most of the gains are in low-income countries, and large losses (that is, increased GII) tend to be in wealthy countries.

In Table 4 we compare the before and after GII with each other and with the natural logarithm of GNI (gross national income) per capita using Pearson correlation coefficients for the values and also for the rankings. (The use of the logarithm is standard in such situations, because the difference, for example, between a country with GNI/capita of \$500 and one with \$1,000 is similar to that between \$5,000 and \$10,000, rather than to that between \$5,000 and \$5,500.) Our conclusion from Table 4 is that, despite having a strong correlation with the old GII, our revised GII substantially reduces the magnitude of the GII-GNI correlation.

Figure 2 shows scatterplots of the original GII (at the top) and the revised GII (at the bottom) against the (logarithm of) GNI per capita. The difference between the top and bottom scatterplots is striking. As expected from the lower correlation coefficient, the data points from the revised GII do not cluster around a line.

**Cautionary remark.** Inevitably there is over-simplification when one tries to represent a complicated phenomenon such as gender inequality by

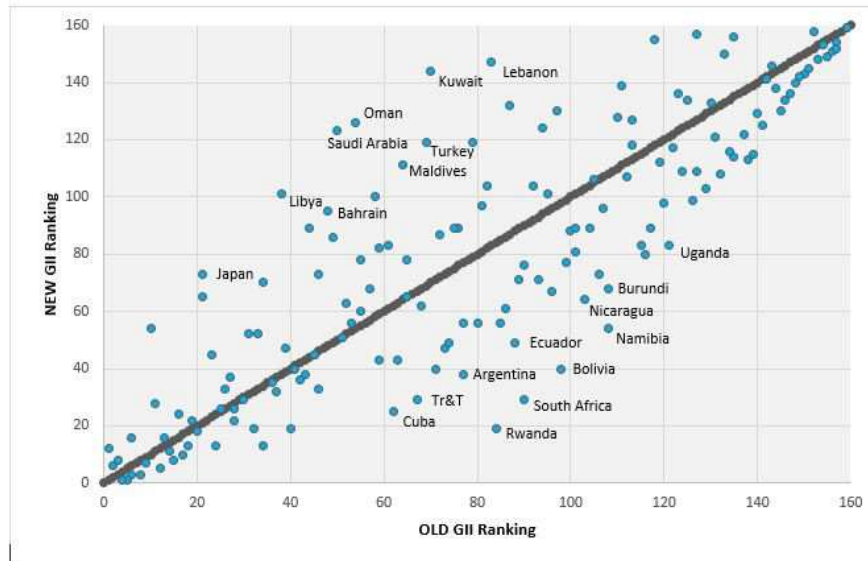


FIGURE 1. Our revised GII rankings plotted against original GII rankings. The labels identify countries most affected by the change; the countries listed below the diagonal are the ones that benefited most from the adjustment in the computation of the reproductive health component. Here Tr&T denotes Trinidad and Tobago.

a single number – a number that of necessity is computed from the type of data that are readily available from almost all countries of the world. In the case of our modified GII, some of the Mideastern countries that do poorly on it actually have unusually *good* conditions for women in certain sectors of the economy. In particular, women professionals in high technology in the Gulf states enjoy greater opportunities, endure less discrimination, and suffer less sexual harassment than their counterparts in Silicon Valley in the US [11]. On the other hand, working-class immigrant women in the same countries are frequently mistreated and abused [8, 6]. Neither circumstance is captured by the GII, either before or after our modification.

## 10. CONCLUSION

We have seen that much of the distortion in the GII country rankings comes from the flawed formula (2) that is used to incorporate the reproductive health indicators. The modified formula (3) replaces the absolute numbers of maternal deaths per 10,000 births and teen pregnancies per 1,000 adolescents by a ratio that compares these numbers to worldwide averages. This simple adjustment substantially lowers the strong negative correlation

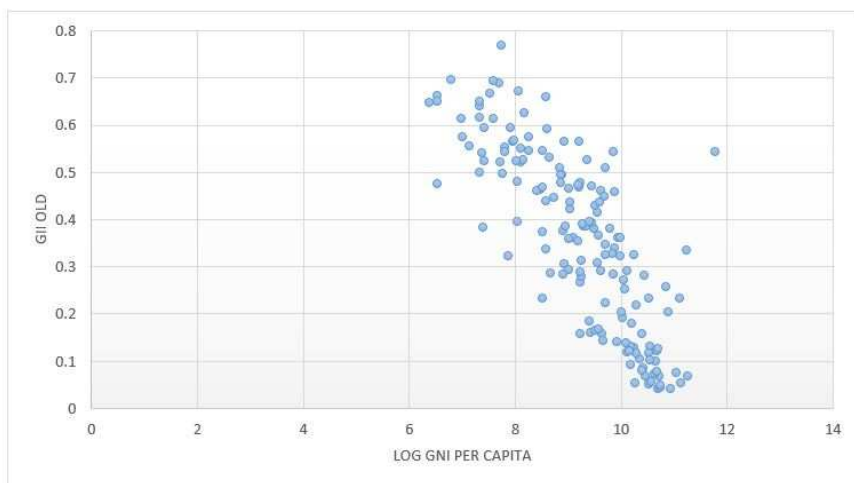
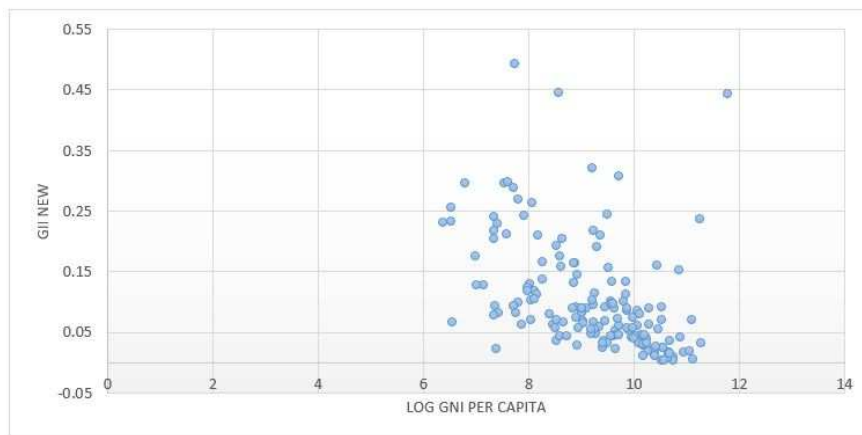


FIGURE 2. Original 2015 GII (top) and our revised 2015 GII (bottom) plotted against  $\ln$  of GNI per capita.



reported in [13] between GII and wealth, and thus reduces (but does not eliminate) the GII's bias in favor of the richer countries.

An important observation about women's empowerment that is emphasized by the UNDP is that improvement in GNI is not necessarily accompanied by reduced inequality, and this is why "governments need to pay dedicated attention to gender equality and not rely solely on growth to achieve it"[17]. However, the formula (2) artificially produces an apparent linear relation between GII and (the logarithm of) GNI per capita (see Figure 2), and this conveys the unintended message that economic growth by itself

improves the status of women. This unfortunate defect can be remedied by using (3) instead of (2).

#### ACKNOWLEDGEMENTS

We wish to thank John Sylvester for explaining Tikhonov regularization, Edith Seier and Jorge Estrada Jr. for helpful comments, Ann Hibner Koblitz for editorial assistance, and an anonymous reviewer for useful suggestions.

#### REFERENCES

- [1] L. Beneria and I. Permanyer, The measurement of socio-economic gender inequality revisited, *Development and Change*, **41** (2010), pp. 375-399.
- [2] Choe Sang-Hun, South Korea confronts open secret of abortion, *The International Herald Tribune*, 6 January 2010.
- [3] J. Ewing, Mathematical intimidation: Driven by the data, *Notices of the Amer. Math. Soc.*, **58** (2011), pp. 667-673.
- [4] A. Gaye, J. Klugman, M. Kovacevic, S. Twigg, and E. Zambrano, Measuring key disparities in human development: The Gender Inequality Index, *Human Development Research Paper 2010/46*, 2010.
- [5] D. Huff, *How to Lie with Statistics*, W. W. Norton, 1954.
- [6] N. Ilesley, Gulf states fail to protect domestic workers from serious abuse, *Newsweek*, 16 October 2015.
- [7] J. Kaipio and E. Somersalo, Classical regularization methods, *Statistical and Computational Inverse Problems* (2005), pp. 7-48.
- [8] A. Kapiszewski, *Nationals and expatriates: Populations and labour dilemmas of the Gulf Cooperation Council states*, Berkshire, England: Garnet, 2001.
- [9] N. Kassebaum, R. Barber, Z. Bhutta, L. Dandona, P. Gething, S. Hay, Y. Kinfu, H. Larson, X. Liang, S. Lim, and A. López, Global, regional, and national levels of maternal mortality, 1990-2015: A systematic analysis for the Global Burden of Disease Study 2015, *The Lancet*, **388** (2016), pp. 1775-1812.
- [10] S. Klasen and D. Schüler, Reforming the Gender-Related Index and the Gender Empowerment Measure: Implementing some specific proposals, *Feminist Economics* **17** (2011), pp. 1-30.
- [11] A. H. Koblitz, Life in the fast lane: Arab women in science and technology, *Bulletin of Science, Technology & Society*. **36** (2016), pp. 107-117.
- [12] C. O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, Crown, 2016.
- [13] I. Permanyer, A critical assessment of UNDP's Gender Inequality Index, *Feminist Economics*, **19** (2013), pp. 1-32.
- [14] F. Salmon, Recipe for disaster: The formula that killed Wall Street, *Wired Magazine*, **17** (3) (2009).
- [15] United Nations Development Program (UNDP), *Human Development Report*, New York: Oxford University Press, 2010.
- [16] UNDP, Table 5: Gender Inequality Index, *Human Development Report* (2016), <http://hdr.undp.org/en/composite/GII> (accessed November 2017).
- [17] UNDP, Box 1.8: Five misconceptions about women's economic empowerment, *Human Development Report* (2016), [http://hdr.undp.org/sites/default/files/2016-human\\_development\\_report.pdf](http://hdr.undp.org/sites/default/files/2016-human_development_report.pdf) (accessed November 2017).
- [18] United Nations Children's Fund (UNICEF), *UNICEF Data: Monitoring the Situation of Children and Women* (2015), <https://data.unicef.org/topic/maternal-health/maternal-mortality/> (accessed December 2017).

- [19] World Economic Forum (WEF), *The Global Gender Gap Report 2017*, [http://www3.weforum.org/docs/WEF\\_GGGR\\_2017.pdf](http://www3.weforum.org/docs/WEF_GGGR_2017.pdf) (accessed December 2017).
- [20] World Health Organization (WHO), *Adolescent Pregnancy* (2014), <http://www.who.int/mediacentre/factsheets/fs364/en/> (accessed December 2017).

CENTER FOR STUDIES IN DEMOGRAPHY & ECOLOGY, EVANS SCHOOL OF PUBLIC  
POLICY & GOVERNANCE, UNIVERSITY OF WASHINGTON, SEATTLE, WA 98195 U.S.A.  
*E-mail address:* `mdjmcDonald@gmail.com`

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF WASHINGTON, SEATTLE, WA 98195  
U.S.A., 206-543-4386 (OFFICE) 206-522-6479 (HOME)  
*E-mail address:* `koblitz@uw.edu`