Slavery is an affront to the fundamental human right to liberty. In our June 2015 issue, we looked at estimates of the number of victims of slavery living in the UK. In this follow-up article, **Jacqueline Joudo Larsen**, **Monti Narayan Datta** and **Kevin Bales** discuss the crucial role extrapolation plays in assessing the number of global victims

Slavery, in its contemporary form, presents pressing challenges for the social scientist. The first is that it is ongoing, widespread, and destructive. The knowledge that any delay in illuminating and understanding this crime might have damaging or fatal results for those suffering enslavement requires a tenacious and timely response. As with public health statisticians confronting an unfolding epidemic, an unhurried research plan is not appropriate, however desirable.

The second pressing challenge is that slavery is a hidden crime. Illegal in all countries, slaveholders work hard to conceal their activities, offering a daunting task to anyone who would seek to measure the extent of slavery. This article presents both an extrapolation technique for estimating the prevalence of this crime at the national level, and a test of the accuracy of that technique.

Perhaps the last precise measurement of slavery occurred in 1860, when the United States Census enumerated those held in legal bondage. In that year, the total number of men, women, and children owned by slaveholders was 3 953 760, accounting for 13% of the US population.

From that time until the late twentieth century, however, there was no reliable measurement of the extent of slavery in any country, with the possible exception of the records of slave labourers kept by the Nazi regime during the Second World War. As the number of countries making slavery illegal increased, so did the difficulty of measuring this hidden criminal activity.

Advancing methods

With the end of the Cold War in 1989, criminal groups exploited newly opened international borders and human trafficking crime increased in the developed world. Responding to a growing concern about global trafficking and slavery in the 1990s, the first attempts were made to achieve a global estimate. In the short period since then, the methodology has advanced rapidly through three stages.

The first stage relied upon secondary sources, including governmental records, non-governmental organisations and service provider tallies, and reports in the media – in short, any source that might shed light on the extent of slavery. Even when sources were systematically assessed for reliability, these estimates[1,2] could only be seen as an approximation of the global situation.

One elaboration of this technique in the US was an attempt to triangulate secondary sources with surveys of service providers and government and law enforcement records. While the estimates derived in this first stage were not widely different from each other, there was no way to ensure their validity.

Perhaps the last precise measurement of slavery occurred in 1860, when the United States Census enumerated those held in legal bondage: 3.95m, or 13% of the US population

The second stage was set in motion by the pioneering work of Pennington et al.[3] in 2009. This team introduced a series of questions concerning human trafficking into a random sample health survey of five eastern European countries (Belarus, Ukraine, Moldova, Romania, and Bulgaria). Employing random sample surveys, they were able to build the first reliable estimate of the proportion of each country's population that had been enslaved, as many of those trafficked across borders often end up in hock to those who transported them and must work to pay off the costs of their journey.

The work of Pennington et al. was critical to the advancement of the methodology for two reasons. Firstly, it demonstrated that, at least in some countries and circumstances,

enslavement could be measured through random sample surveys of the full population. Secondly, by fixing valid data points for these five countries, it became possible to begin the process of estimating the *range* of prevalence of modern slavery across countries, as well as using these and other emerging survey results to extrapolate the prevalence of slavery in other countries.[4]

In addressing the question of range, by 2009 it had become clear that slavery cases were being reported in virtually all countries with a population over 100 000. For that reason, the low end of the global range of prevalence of modern slavery for countries in which measurement was possible was assumed to be greater than zero. In the same year, a report from the United States Agency for International Development and the Pan American Development Foundation[5] provided a random sample survey of child *restavèk* slavery in Haiti. The proportion of the Haitian population (2.3%) found to be in this pervasive form of slavery was assumed to be in the upper range of the global distribution of slavery.

In 2014, a more robust extrapolation became possible with this sense of the range of prevalence across countries, supported by an increase in the amount of data available from random sample surveys. In addition to data from the Pennington et al. and Haiti surveys, random sample surveys of slavery were also identified in three more countries (Niger, Namibia, and the Democratic Republic of Congo), and an additional seven (Pakistan, Indonesia, Brazil, Nigeria, Ethiopia, Nepal, and Russia) were commissioned by the Walk Free Foundation using the Gallup International World Poll (bit.ly/1FfBCN8), which also provided data for three other countries (Saudi Arabia, Qatar, and Malaysia), resulting in data points from a total of 19 random sample surveys.

At the most basic level the process of extrapolation assumed that the 167 most populous countries were ranged along a continuum from those with the lowest

prevalence of slavery to those with the highest prevalence of slavery. Further, the data points derived from the 19 random sample surveys available were considered to be indicative of similar countries along the continuum.

In the 2014 Global Slavery Index, the process of extrapolating estimates of modern slavery followed five steps. In addition to the collection of data on prevalance, a large number of variables were collected for all 167 countries to measure vulnerability to enslavement across five dimensions: state stability, social discrimination, presence or absence of human rights protections, economic and social development, and presence of governmental slavery policy. The first step involved applying a K-means cluster analysis (see box) to group the 167 countries into seven distinct groups using these dimensions of vulnerability. The research team employed K-means as a clustering algorithm, deciding on the final cluster out of a dozen trials that had the highest pseudo-F score. Seven groups were chosen as there were seven prevalence surveys that used the same instruments and data collection methods, and ideally, each survey could be applied to a unique group.

The seven groupings were sufficiently distinct on overall mean values, although the minimum and maximum values did indicate some overlap among countries at the bottom of one list and top of the next.

The extrapolation process then followed these steps:

1. Countries within each group were ranked according to their mean vulnerability score, from low to high. This aggregated countries that are more alike as compared to those that are less alike.
2. Geography was also important to consider within and among clusters, given that some countries share many similar attributes that correlate highly with geography. Where available and relevant, the final calculation of estimated enslavement would incorporate survey data for a country within the same geographical region to reflect regional similarities.
3. Once this process was concluded, the research team examined countries on a case-by-case basis to determine if the extrapolation process corresponded to data collected from secondary source estimates. Country-level adjustments were also made in discussion with experts on specific countries.

K-means cluster analysis

K-means is a statistical method which groups similar items into clusters, ensuring that items not in the same cluster are as different as possible. This is achieved by allocating an item to the cluster with the nearest centroid, or the mean of the cluster. The cluster's mean is then recalculated and the process of allocating items to clusters begins again until no items change groups, or those changes do not make a substantial difference in the definition of a cluster. The Calinski–Harabasz pseudo-F score describes the ratio of between-cluster variance to within-cluster variance, whereby large values indicate distinct clustering.

4. A final, downward, adjustment was made for small island developing states in order to ensure that estimates for these nations erred on the conservative side in light of (in most cases) their remoteness, small but growing populations, limited resources, susceptibility to natural disasters, and limited potential for economic growth, among other development challenges.

This extrapolation process yielded a prevalence estimate for each country calculated as a proportion of the total population that was enslaved within that country. For all 167 countries this produced a total global estimate in 2014 of 35.8 million enslaved people.

A valid estimate?

Following the release of the Global Slavery Index in 2014, a further 19 surveys were commissioned by the Walk Free Foundation. In addition, multiple systems estimation (MSE) was used to compute a new prevalence estimate for the UK.[6] These new surveys allow a test of the extrapolation procedures, and can be added to previous surveys that were carried out after extrapolation estimates were made.

The survey data to which the extrapolation estimates can be compared is in three sets. Seven countries were surveyed in 2014, a further 19 in 2015, and surveys by other researchers were found for two other countries.[7,8] The addition of the UK's MSE based estimate brings the total number of countries to 29. Altogether these countries, and thus these surveys, represent 3.3 billion people, approximately 46% of the global population.

A comparison of the extrapolation estimates of slavery to the estimates derived from random sample representative surveys yields very promising results. As Table 1 shows, all but one of the extrapolation estimates of the prevalence of slavery within a national population fell within one percentage point of the estimates arrived through random sample surveys.

After the publication of extrapolation estimates in the 2013 and 2014 editions of the Global Slavery Index, there were concerns and reservations about the reliability of this technique. While we will continue to practise continuous and incremental testing and improvement of the methodology, the findings presented here provide a level of confidence in the use of extrapolation estimates in the absence of further random sample surveys.

The extrapolation technique presented here is not the only, or the best, means by which to estimate the prevalence of slavery. But at this time, and for many countries, it is the *only* method available, and will continue to be well into the future. This may be because ongoing conflict precludes on-the-ground survey work, or because the relevant government refuses to allow research to take place, or simply because of a lack of resources needed to carry out extensive random sample surveys. This test of the extrapolated estimates against the subsequent random sample survey results, however, supports the continued use of this technique until more precise methods are available.

Table 1. Difference in the percentage of the population estimated to be enslaved

	2013 extrapolation-based estimate, to 2014 survey-based estimate	2014 extrapolation-based estimate, to 2015 survey-based estimate
Pakistan	0.08	
Indonesia	0.20	
Brazil	0.03	
Nigeria	0.07	
Ethiopia	0.30	
Nepal	0.12	
Russia	0.37	
Namibia		0.21
Democratic Republic of Congo		0.43
India*		0.49
The Philippines		0.13
South Africa		0.26
Ghana		0.36
Thailand		0.69
Mauritania		1.59
Bangladesh		0.52
Vietnam		0.21
Cambodia		0.61
Myanmar		0.52
Sri Lanka		0.29
Tunisia		0.46
Guatemala		0.63
Chile		0.05
Dominican Republic		0.81
Poland		0.42
Hungary		0.03
Bolivia		0.21
Mexico		0.19

* In the case of India, this is a preliminary result based upon a nationally representative, random sample survey ($n = 3000$). The final estimate, however, will integrate data from a further 12 surveys, with samples representative at a state level. The combined sample size of 15 000 will provide a more robust analysis of India.

We have moved quickly to disseminate and publish these findings not simply because methodological testing and transparency are the absolute foundation of science, but also because of the immediate and pressing threat that slavery represents to human well-being. We are committed to ending the suffering and cost of slavery. We are certain that clear and objective measurement produced within the framework of globally accepted scientific process (open data, methodological transparency, and anonymous peer review) can provide the crucial metrics against which the progress of liberation – of human freedom – can be measured.

Acknowledgements

References

1. Bales, K. (1999) *Disposable People: New Slavery in the Global Economy*. Berkeley: University of California Press.

2. International Labour Office (2005) *A Global Alliance against Forced Labour: Global Report under the Follow-up to the ILO Declaration on Fundamental Principles and Rights at Work 2005*. Geneva: ILO. bit.ly/19WWQ8H

3. Pennington, J. R., Ball, W. A., Hampton, R. D. and Soulakova, J. N. (2009) The cross-national market in human beings. *Journal of Macromarketing*, 29(2), 119–134.

4. Datta, M. N. and Bales, K. (2013) Slavery in Europe: Part 1, Estimating the dark figure. *Human Rights Quarterly*, 35(4), 817–829.

5. Pierre, Y.-F., Smucker, G. R. and Tardieu, J.-F. (2009). *Lost Childhoods in Haiti: Quantifying Child Trafficking, Restavèks & Victims of Violence*. Washington, DC: Pan American Development Foundation. bit.ly/1N6TLVv

6. Bales, K., Hesketh, O. and Silverman, B. (2015) Modern slavery in the UK: How many victims? *Significance*, 12(3), 16–21.

7. Johnson, K., Scott, J., Rughita, B., Kisielewski, M., Asher, J. and Ong, R. (2010) Association of sexual violence and human rights violations with physical and mental health in territories of the eastern Democratic Republic of the Congo. *Journal of the American Medical Association*, 304(5) 553–562.

8. Directorate of Labour Market Services, Ministry of Labour and Social Welfare (2005) *Namibia Child Activities Survey 2005*. Windhoek, Namibia: Ministry of Labour and Social Welfare. bit.ly/1hulJNp

Jacqui Joudo Larsen is senior research manager, Walk Free Foundation. **Monti Narayan Datta** is associate professor of Political Science, University of Richmond. **Kevin Bales** is professor of Contemporary Slavery, Wilberforce Institute for the Study of Slavery and Emancipation, University of Hull