Prevalence estimation for the UK Modern Slavery Strategy and for US Cities

Sir Bernard Silverman

Professor of Modern Slavery Statistics

University of Nottingham, UK

www.bernardsilverman.co.uk

Measuring modern slavery is hard. Even if we can agree what a 'slave' is, finding them is difficult, since most of them are deliberately hidden from view. But without good data on where slaves are, how they become slaves, and what happens to them, anti-slavery policy will remain guesswork. Yet there has been significant innovation in this field recently, both in counting-based methodologies and in inferential or estimation methodologies.... The UK Home Office has developed a pioneering methodology based on Multiple Systems Evaluation to estimate the "dark figure" of hidden slaves, placing the number in the UK at 10,000 to 13,000

James Cockayne, United Nations University, 2015

Questions to consider

- How do you analyse data when many characteristics you'd like to know are not available?
- How should results be presented to policy makers and the wider public?
- How can we make data available in an easily available public repository to facilitate further research, without compromising confidentiality, and why does this matter?
- What are the features of data in this area which demand specific methodology?

Modern Slavery Act 2015

A CENTRAL PART OF THE MODERN SLAVERY STRATEGY NOV 2014

Professor Sir Bernard Silverman FRS
Chief Scientific Adviser to the Home Office 2010–2017

Provide Home Secretary, ministers and officials with **impartial and expert advice**

Leadership of **Home Office Science:** 500 staff in many different areas

Responsible for **science and research** in the Home Office and its agencies

Member of the cross-government **Chief Scientific Adviser network**

Foster links to industry and academia, and internationally

Facilitate **science advisory committees:** Drugs, Animals in Science, DNA Ethics

Bring own broad expertise to individual projects and topics

The National Referral Mechanism (NRM)

A framework for identifying victims of human trafficking and ensuring they receive the appropriate protection and support. It is run by the National Crime Agency (NCA), which is an operationally independent non-ministerial government department.

The NRM collates data from a range of sources to produce statistics on “potential victims of trafficking” and produces an annual strategic assessment.

In 2013, the NCA Strategic Assessment identified 2,744 potential victims

The sources of information can be grouped

LA	Local authorities
NG	Non-governmental organisations, charities
GO	Government organisations e.g. UK Border Force, Gangmasters and Labour Abuse Authority
PF	Police forces, NCA
GP	General public (various routes)

LA	X					X	X	X								X	X	**X**	
NG		X				X			X	X	X				X	X	X	**X**	
PF			X				X		X			X	**X**		X	X		**X**	
GO				**X**				X		X		X		X	X		X	**X**	
GP					X						X		**X**	X					
	54	463	995	**695**	316	15	19	3	62	19	1	76	**11**	8	4	1	1	**1**	??

For example: 695 cases appear on list GO only. 11 cases on PF and GP but not others. One case on all four of LA, NG, PF and GO, but not GP. The "dark figure" is the number which should appear in the last column—those cases which do not appear in any list.

A lot of work went into this table! This is all we are allowed to know.

10,000 to 13,000 victims

- The estimated confidence interval is from 10K to 13K, so this suggests that the Strategic Assessment was aware of 20% to 25% of all the potential victims of trafficking in the UK in 2013
- A tentative conclusion: the model is based on assumptions that (while sensible) can't be easily verified and it inevitably uses data that have some limitations
- Analysis also reveals correlation between various sources: e.g. cases reported to NGOs are more likely also to be known to police; cases reported by general public are less likely to be in other lists
- Robustness studies involving leaving out less reliable lists and merging smaller ones were encouraging
- Choice of model is important

UK is home to 13,000 slaves: Home Office says number is four times higher than previously thought

- The shocking figures include women trafficked for prostitution or service
- Others are brought in to work in factories, fields or even fishing boats
- The Home Office confirmed it underestimated the scale of the problem
- The most common victims are from Romania, Poland, Albania and Nigeria

By MAIL ON SUNDAY REPORTER

PUBLISHED: 01:55, 30 November 2014 | **UPDATED:** 15:56, 30 November 2014

Share 189 shares 442 View comments

Up to 13,000 people in Britain are being held in conditions of slavery, four times the number previously thought, the Home Office has said.

In what is said to be the first scientific estimate of the scale of modern slavery in the UK, the Home Office has said the number of victims last year was between 10,000 and 13,000.

They include women forced into prostitution, domestic staff and workers in fields, factories and fishing boats.

Scroll down for video

Home Secretary Theresa May, pictured, described the scale of human trafficking as 'shocking'

Data from the National Crime Agency's Human Trafficking Centre had previously put the number of slavery victims in 2013 at 2,744.

The new estimate is based on a statistical analysis by the Home Office chief scientific adviser, Professor Bernard Silverman, which aims for the first time to calculate the 'dark figure' of victim numbers who are not reported to the law enforcement agencies.

'Modern slavery is very often deeply hidden and so it is a great challenge to assess its scale,' he said.

'The data collected is inevitably incomplete and, in addition, has to be very carefully handled because of its sensitivity.'

Read more:
http://www.dailymail.co.uk/news/article-2854601/UK-home-13-000-slaves-Home-Office-says-number-four-times-higher-previously-thought.html#ixzz3L2PU4VAD

Data from New Orleans

Lists	Number	Lists	Number	Lists	Number	Lists	Number
A only	25	F only	6	A & E	1	D & E	2
B only	5	G only	6	B & F	1	E & H	1
C only	70	H only	21	C & D	1		
D only	33	A & C	1	C & E	1	A & C & G	1
E only	6	A & D	2	C & G	1	A & D & E	1

- Much smaller numbers
- Eight lists; very many combinations with zero count
- Extremely sensitive considerations about confidentiality, especially for federal agencies
- Standard MSE programs don't really work
- Overall estimate is 650 to 1600

What we need to do

- ✓Put data into public domain, both for transparency and for future research. Can we have a single public repository?
- ✓Socialise data providers about need to do this, but reassure them about security and confidentiality.
- ✓Socialise policy makers and others about uncertainty. Do not make exaggerated claims about accuracy; they will eventually rebound.
- ✓Rigour is needed. Use the standards of repeatability and transparency of code now standard in other fields.
- ✓Develop specific methodology for sparse tables.
- ✓Aim for standard method(s).

Investigating the methodology

- There's never any ground truth
- One approach is to look at some standard data sets and see if you get stable results
- Unfortunately the social science culture on open data and open research is a long way behind what's now standard in most physical and life science areas.
- In addition there are genuine sensitivities about privacy
- There are several data sets out there and a variety of different methods of model choice and fitting. The estimates based on real data can vary drastically according to which method is used. A Bayesian method is promising.

The Poisson log-linear model

Suppose there are *t* lists. A *capture history* is a collection $\boldsymbol{\omega}$ of lists, giving the lists on which an individual is observed. Here $\boldsymbol{\omega} \subseteq \{1, 2, \ldots, t\}$.

Let $N_{\boldsymbol{\omega}}$ be the number of cases that appear on the lists in $\boldsymbol{\omega}$ and no others. Model $N_{\boldsymbol{\omega}}$ as Poisson($\lambda_{\boldsymbol{\omega}}$) where

$$\log(\lambda_{\boldsymbol{\omega}}) = \alpha + \sum_{i \in \boldsymbol{\omega}} \alpha_i + \sum_{\{i,j\} \subseteq \boldsymbol{\omega}} \alpha_{ij}$$

The expected value of the dark figure is e^{α} since this corresponds to the case where the collection $\boldsymbol{\omega}$ is empty, so estimating α gives the estimate (and confidence interval) for the dark figure.

To choose a model, specify which interactions α_{ij} are to be included.

Sparse tables

- Identify all non-overlapping pairs $\{i, j\}$ in the model. The estimates of the corresponding parameters are $-\infty$.
- Remove all $N_{\boldsymbol{\omega}}$ for which $\boldsymbol{\omega}$ contains one of these non-overlapping pairs. Estimate remaining parameters from the remaining data points using standard GLM algorithms.
- However, this isn't the whole story...you have to check further for the existence and identifiability of the estimates. This is considered very abstractly by Fienberg and Rinaldo (2012) and is not dealt with in any standard GLM software.
- We have implemented their criteria for the MSE model. In addition, the conditions for AIC and BIC don't apply, so we have developed a different fitting approach.

Prevalence estimation for the UK Modern Slavery Strategy and for US Cities

Sir Bernard Silverman

Professor of Modern Slavery Statistics

University of Nottingham, UK

www.bernardsilverman.co.uk