

## RESEARCH BRIEF

# Monitoring Target 16.2 of the United Nations Sustainable Development Goals: multiple systems estimation of the numbers of presumed victims of trafficking in persons

*Romania*

In partnership with Walk Free Foundation

## Introduction and methodology<sup>1</sup>

The statistical technique to estimate the size of hidden populations, known as capture/recapture analysis, multiple systems analysis or multiple record systems analysis, was originally developed by biologists to estimate animal populations. The quintessential idea to estimate the number of fish in a pond is as follows. One catches a number of fish (say 100), tags them, and then throws them back into the same pond. Sometime later, one takes a new catch (say another 100) from the same pond, and counts how many of the second catch are tagged, as having been part of the returned original first catch. If the overlap between the two catches is zero, or very small, this suggests that the population of fish in the pond is much larger than 100. If the overlap is considerable, say 50, this suggests that the population is smaller. The larger the overlap, the smaller the size of the total fish population. If out of the 100 fish in the second catch, 20 are tagged, it follows that the tagged fish in the pond had a chance of one in five of being caught. Assuming that the non-tagged fish have similar catchment chances, the total number of fish in the pond can be estimated at 500, of which 400 untagged.

The capture-recapture approach of biologists has translated into a well-known method for estimating the size of a hidden human population using two independent recording systems (or registers) which partially list its members. Linking the individuals in the two registers allows for the estimation of the number of individuals that are not recorded in any of the registers. For example, with two registers A and B, linkage gives a count of individuals in A but not in B, a count of individuals in B but not in A, and a count of individuals in both A and B. The counts form a contingency table denoted by  $A \times B$ , with the variable labeled A being short for “inclusion in register A” differentiating between the categories “yes” and “no,” and likewise for register B. The statistical problem is to estimate the value in the cell “no, no”. An estimate of the total population size is obtained by adding the estimated count of doubly missed individuals to the counts of individuals found in at least one of the registers.

The capture-recapture method has been successfully applied to estimate the size of hidden human populations by determining the overlaps between unique individuals appearing in separate recording systems (or lists). Using such capture-recapture analysis, estimates have, for example, been made of the numbers of casualties of human rights violations in Peru and irregular migrants in the Netherlands (Lum, Price & Banks, 2013; Van der Heijden, et al, 2015).

Capture-recapture analysis depends on certain assumptions about the lists and the population from which they are drawn. Arguably the most problematic condition to fulfill when using recording systems of human individuals is the condition that recording systems are independent of each other. In the case of records of persons, this assumption of independence of lists is usually not met. For example, persons identified by law enforcement authorities as possible victims of human trafficking are likely to be referred to social assistance programs, and they consequently have a higher probability of being included in the recording systems of the involved NGOs. In this case the inclusion in the list of a service provider is far from independent from registration by the police. This is an example of positive dependence. Since such positive dependence increases the overlap between the two lists, the number of the unobserved population is consequently underestimated. In practice, negative dependence may

---

<sup>1</sup> UNODC research based on contributions by Jan J. M. van Dijk (University of Tilburg, Netherlands), Maarten Cruyff and Peter G. M. van der Heijden (University of Utrecht, Department of Psychology, Netherlands).

also occur, for example, when inclusion in one register lowers the chance of being registered in another register, which leads to an overestimation.

A promising approach to relax the condition of independence is to include a third register, or multiple additional registers, and to analyse the three-way, or multiple-way contingency tables. With three (or more) lists the independence assumption in the two-list case is replaced by the less severe assumption that three (or more) factor interaction is absent. In official statistics, this extension of the two-list capture-recapture method is known as multiple systems estimation (MSE).<sup>2</sup>

In collaboration with the three countries concerned, and in partnership with the Walk Free Foundation, UNODC carried out MSE studies to estimate the total number of victims of trafficking in persons in Ireland, Romania and Serbia. This Research Brief presents the findings from Romania.

---

<sup>2</sup> See also Van Dijk, J.J.M. and P.G.M. van der Heijden (2016).

## Romania

### *Introduction*

Romania is mainly a country of origin for victims of trafficking in persons. The total number of identified trafficking victims in the period 2011-2015 was 4,622 (2011: 1,048; 2012: 1,041; 2013: 896; 2014: 757; 2015: 880). In 2016, the number was 756. During 2011-2015 the majority of the identified victims were female (66 per cent). The number of identified child victims was respectively 319 in 2011, 370 in 2012, 300 in 2013, 290 in 2014, and 316 in 2015. The prevailing form of exploitation of the victims in the reporting period was sexual exploitation (54 per cent of the victims), followed by forced labour (34 per cent) and forced begging (6 per cent). The overwhelming majority of the victims were Romanian citizens, with the main countries of destination being Italy, Spain, Germany, France, Greece, Poland and the United Kingdom. The scale of domestic trafficking, i.e. within Romania, has been on the rise, with some 35 per cent of identified victims having been trafficked internally in 2011-2014. In this period, only 15 foreign victims were recorded.

### *Results*

The results of the MSE on the Romanian data from 2015 and 2016 show that the estimated numbers of presumed victims are 1,300 and 1,200 respectively, suggesting a downward trend. The rate of trafficking victims per 100,000 population in 2015 was 6.1. The victimization rates are 9.5 for females and 2.6 for males.

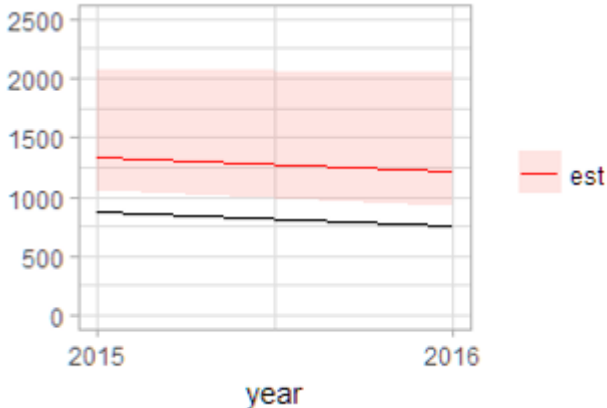
Many of the cases recorded in Romania involved Romanian victims trafficked to other countries. Only 40 per cent of the cases took place on national territory. The trafficking victimization rate on Romanian territory is not 6.1 but 2.5.

The estimated presumed victims are to a large extent females and girls trafficked for sexual exploitation within the country. In 2015, 60 per cent of the estimated cases related to trafficking for sexual exploitation, 10 per cent to trafficking for begging and 30 per cent to other types, mainly trafficking for forced labour.

The estimated number of victims is 50 per cent higher than the observed numbers (a ratio between recorded and estimated numbers of 1.5), implying that a large proportion of all trafficking cases are detected and recorded in Romania.

The ratios between observed and estimated numbers vary across the subgroups distinguished. The numbers are too small to permit the drawing of firm conclusions, however. Ratios for minors are lower than for adults (1.4/1.4 for adults and 1.7/1.9 for minors). In terms of detection, this result suggests that the chances to be detected are significantly lower for minor victims than for adult victims. Moreover, in Romania, female victims appear to be somewhat less likely to be detected than male victims.

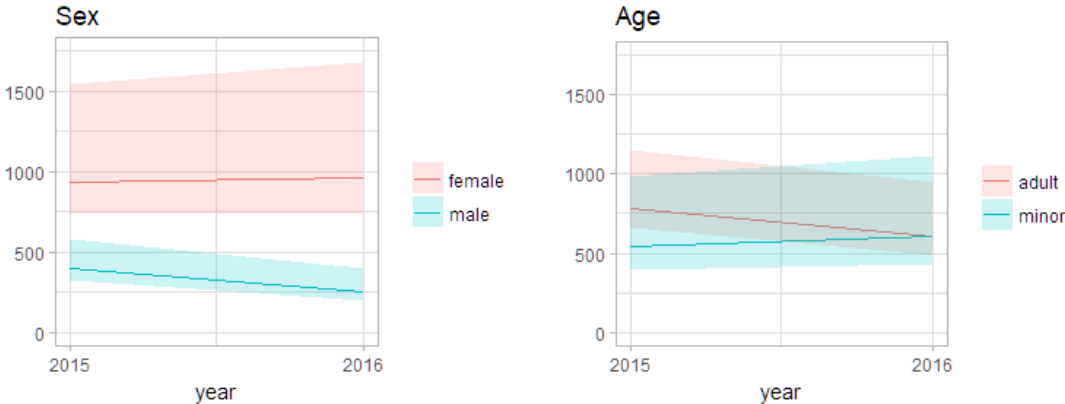
Figure 1: Trends in observed presumed victims (in black) and estimated totals of presumed victims (in red) in Romania, 2015-2016



(est=estimated number of victims)

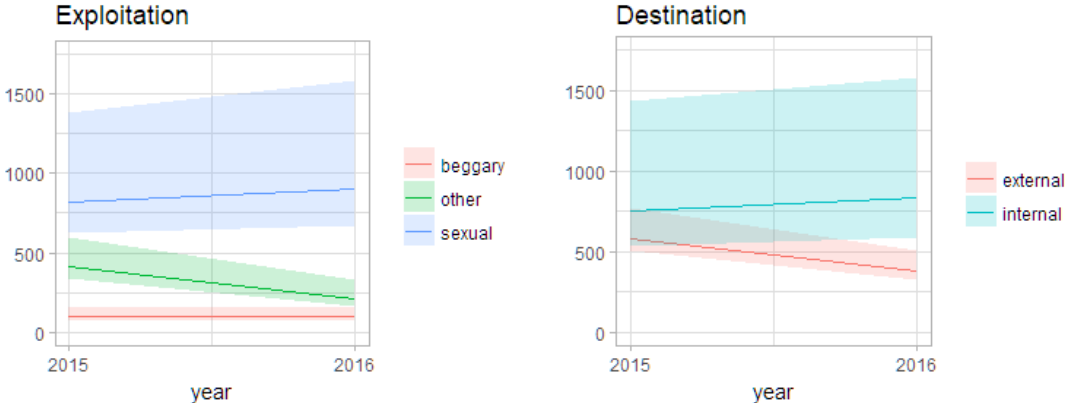
The 95 per cent confidence interval around the estimates is depicted in shaded red. The results show that the estimated numbers are just about fifty per cent higher than the observed numbers (with a maximum of 100 per cent).

Figure 2: Trends in estimated totals of presumed victims in Romania, 2015-2016, by sex and age



The results show that the estimated numbers of female victims are more than twice as high as those of males. The estimated numbers of presumed minor and adult victims are in the same range. The two estimates appear to be converging but the margins of error do not permit for definite conclusions to be drawn.

Figure 3: Trends in estimated totals of presumed victims in Romania, 2015-2016, by type of exploitation and destination (internal/external)



The results show that the estimated numbers of victims trafficked for sexual exploitation are significantly higher than for begging and the ‘other’ category which mainly consists of forced labour. Estimated numbers for domestic trafficking are higher than for transnational trafficking, especially so in 2016. This confirms the ongoing upward trend in domestic trafficking, which is evident in the recorded numbers of presumed victims over the past five years.

**Data**

The National Agency against Trafficking in Persons (ANITP) which is part of the Ministry of the Interior, co-ordinates, evaluates and oversees national anti-trafficking work. It maintains the National Integrated System to Monitor and Assess Trafficking in Persons (SIMEV) on the basis of data provided by relevant institutions, including the Superior Council of Magistracy. The SIMEV database brings together data on presumed victims from five main sources: Police/NATP plus Border Police, IOM, NGOs, foreign authorities (mainly police forces) and other (mainly diplomatic missions). These are designated as R1 to R5. The overlaps between the five lists are given below.

**Overlap registers**

V1	V2	R1xR2	R1xR3	R1xR4	R1xR5	R2xR3	R2xR4	R2xR5	R3xR4	R3xR5	R4xR5
0	0	77	74	65	33	1531	1492	1513	1473	1498	1453
1	0	1493	1498	1432	1509	41	5	29	24	44	89
0	1	44	47	56	88	39	78	57	99	74	44
1	1	22	17	83	6	25	61	37	40	20	50

The database includes information on age, sex, type of exploitation (sexual, ‘other’ which is mainly forced labour, and begging) and the location of exploitation (domestic or transnational). As mentioned, very few foreign victims are being identified. Therefore, the latter variable was not included in the analysis.

The analysis was carried out on data for 2015 and 2016 as these data were deemed to be of superior quality by ANITP.

***Covariates***

Table 2 shows the distribution of all cases over the five covariates mentioned.

*Table 2 Overview of covariates in Romanian dataset and numbers of cases*

- S (sex): female (1169), male (467)
- A (age): adults (1007), minors (629)
- E (form of exploitation): sexual (1028), begging (137), other (471)
- D (destination): transnational (external) (813) and domestic (internal) (823)
- Y (year): 2015 (880), 2016 (756)

The model selection procedure was basically the same as in the analysis of the Dutch data, presented in Annex V. The model search started with a simple model and included interaction terms until the fit of the model to the data became adequate. As explained, this time the additional variables sex (S), age (A), exploitation (E), destination (D), and year (Y) were included. In the STEP procedure, the BIC was preferred over the AIC as the selection criterion, because of its protection against overfitting when the sample is relatively large.

***Model selection***

	DEV	par	df	AIC	BIC	Nhat
Independence model	3601	12	1476	3995	4060	2330
Step model (BIC)	566	37	1451	1010	1210	2541
Step model (AIC)	497	48	1440	963	1223	2112

The linear model (BIC) that was selected as the most parsimonious one is presented below.

***Model effects***

A, AD, AE, D, DY, E, ED, EY, S, SE, Y

R1,R1D,R1R2,R1R3,R1R4,R1R5,R2,R2E,R2R4,R3,R3R4,R3Y,R4,R4D,R4R5,R5,R5D,R5E

## Literature

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

Lum, K. M. E. Price & D. Banks (2013), 'Applications of Multiple Systems Estimation in Human Rights Research,' *The American Statistician*, 67:4, 191-200.

Silverman, B. (2014), *Modern slavery: an application of multiple systems estimation*. Gov.UK.

Van der Heijden, P.G.M., M. Cruyff & G.H.C. van Gils (2015), *Schattingen illegaal in Nederland verblijvende vreemdelingen 2012-2013*. WODC. (Estimations of illegal residents in the Netherlands 2012-2013).

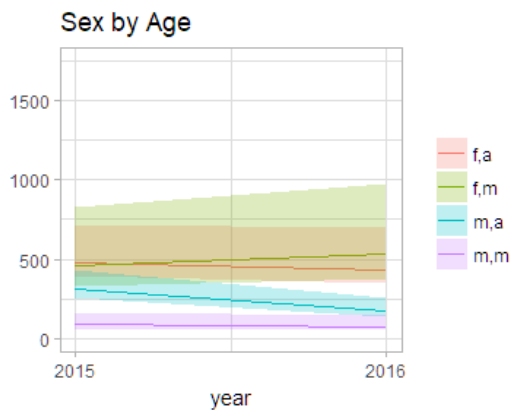
UNODC (2017), *Monitoring Target 16.2 of the United Nations Sustainable Development Goals; A multiple systems estimation of the numbers of presumed human trafficking victims in the Netherlands in 2010-2015 by year, age, gender, form of exploitation and nationality*; Research Brief. Vienna: UNODC.

Van Dijk, J.J.M. and P.G.M. van der Heijden (2016), *Research Brief. Multiple Systems Estimation for Estimating the number of victims of human trafficking across the world*. Vienna: UNODC.

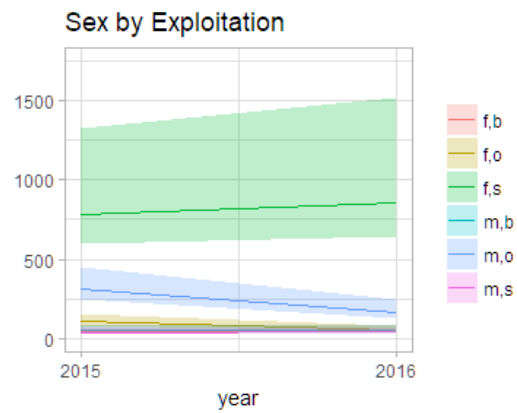


Annex I: Further details - Romania

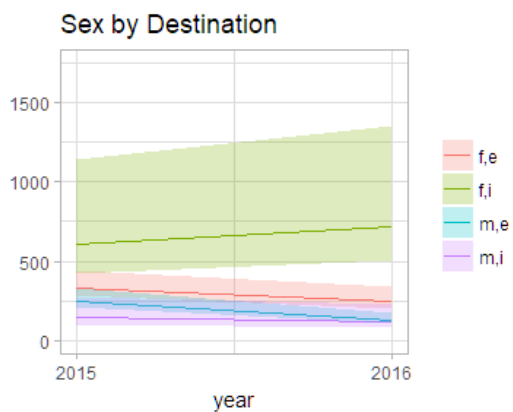
**I A: Two-dimensional plots of disaggregated findings**



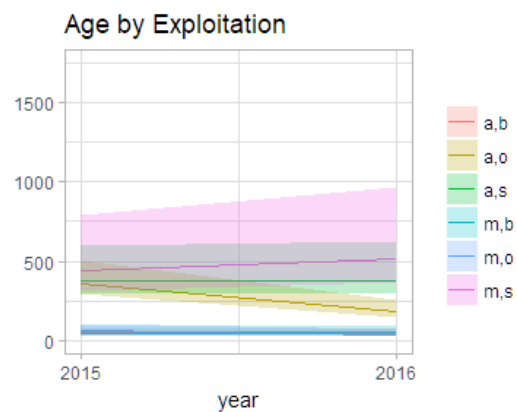
Legend: f=female; m(1<sup>st</sup> column)=male; a=adult; m(2<sup>nd</sup> column)=minor.



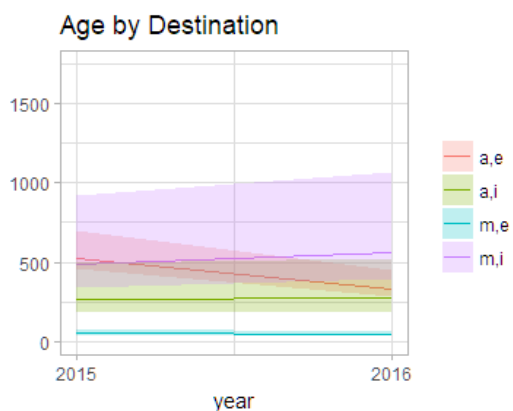
Legend: f=female; m=male; b=begging; s=sexual exploitation; o=other forms of exploitation.



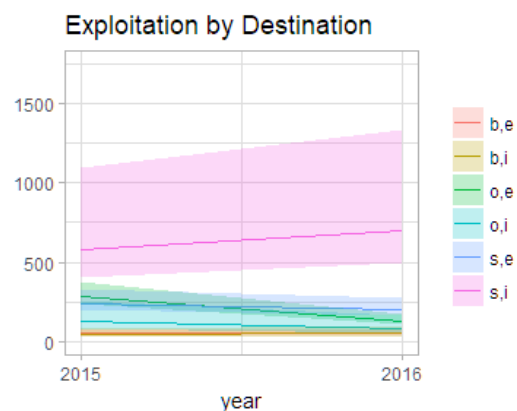
Legend: f=female; m=male; e=external (out of Romania); i=internal (within Romania).



Legend: a=adult; m=minor; b=begging; s=sexual exploitation; o=other forms of exploitation.



Legend: a=adult; m=minor; e=external (out of Romania); i=internal (within Romania).



Legend: b=begging; s=sexual exploitation; o=other forms of exploitation; e=external (out of Romania); i=internal (within Romania).

**I B: Observed, estimated, min and max of 95% confidence interval and ratio estimated/observed, according to year (2015 and 2016), sex, adult/minor, type of exploitation and internal or external**

Y			obs	est	min95%	max95%	ratio
1			880	1328	1062	2083	1.5
2			756	1211	931	2057	1.6
Y	S		obs	est	min95%	max95%	ratio
1	female		583	934	733	1546	1.6
2	female		586	960	731	1682	1.6
1	male		297	395	325	581	1.3
2	male		170	251	199	399	1.5
Y	A		obs	est	min95%	max95%	ratio
1	adult		564	784	660	1149	1.4
2	adult		443	605	490	952	1.4
1	minor		316	544	395	984	1.7
2	minor		313	606	434	1113	1.9
Y	E		obs	est	min95%	max95%	ratio
1	beggary		69	97	72	157	1.4
2	beggary		68	99	71	165	1.5
1	other		313	413	343	599	1.3
2	other		158	214	169	331	1.4
1	sexual		498	818	626	1382	1.6
2	sexual		530	898	674	1589	1.7
Y	D		obs	est	min95%	max95%	ratio
1	external		492	578	506	768	1.2
2	external		321	377	322	508	1.2
1	internal		388	751	529	1431	1.9
2	internal		435	835	588	1576	1.9
Y	S	A	obs	est	min95%	max95%	ratio
1	female	adult	307.0	477	392	719	1.6
2	female	adult	313.0	431	344	701	1.4

1	male	adult	257.0	308	259	434	1.2
2	male	adult	130.0	174	141	256	1.3
1	female	minor	276.0	457	330	829	1.7
2	female	minor	273.0	529	379	973	1.9
1	male	minor	40.0	87	61	161	2.2
2	male	minor	40.0	77	54	147	1.9
Y	S	E	obs	est	min95%	max95%	ratio
1	female	beggary	36.0	48	33	77	1.3
2	female	beggary	31.0	49	33	82	1.6
1	male	beggary	33.0	50	35	83	1.5
2	male	beggary	37.0	51	35	87	1.4
1	female	other	72.0	104	81	155	1.5
2	female	other	48.0	54	40	85	1.1
1	male	other	241.0	308	252	450	1.3
2	male	other	110.0	160	126	248	1.5
1	female	sexual	475.0	782	599	1331	1.6
2	female	sexual	507.0	858	646	1519	1.7
1	male	sexual	23.0	37	24	64	1.6
2	male	sexual	23.0	40	26	73	1.7
Y	S	D	obs	est	min95%	max95%	ratio
1	female	external	256.0	327	280	444	1.3
2	female	external	231.0	245	206	337	1.1
1	male	external	236.0	251	215	335	1.1
2	male	external	90.0	132	109	174	1.5
1	female	internal	327.0	607	421	1144	1.9
2	female	internal	355.0	715	501	1356	2
1	male	internal	61.0	144	98	279	2.4
2	male	internal	80.0	119	81	233	1.5
Y	A	E	obs	est	min95%	max95%	ratio
1	adult	beggary	37.0	51	37	76	1.4
2	adult	beggary	41.0	47	33	72	1.2
1	minor	beggary	32.0	46	30	87	1.4

2	minor	beggary	27.0	52	33	99	1.9
1	adult	other	277.0	354	297	503	1.3
2	adult	other	139.0	178	142	261	1.3
1	minor	other	36.0	59	38	111	1.6
2	minor	other	19.0	36	23	70	1.9
1	adult	sexual	250.0	379	304	597	1.5
2	adult	sexual	263.0	380	298	623	1.4
1	minor	sexual	248.0	439	314	792	1.8
2	minor	sexual	267.0	518	367	967	1.9

Y	A	D	obs	est	min95%	max95%	ratio
1	adult	external	444.0	521	455	699	1.2
2	adult	external	285.0	334	286	449	1.2
1	minor	external	48.0	56	43	79	1.2
2	minor	external	36.0	43	33	61	1.2
1	adult	internal	120.0	263	180	499	2.2
2	adult	internal	158.0	272	185	521	1.7
1	minor	internal	268.0	488	343	925	1.8
2	minor	internal	277.0	563	393	1065	2

Y	E	D	obs	est	min95%	max95%	ratio
1	beggary	external	45.0	57	43	82	1.3
2	beggary	external	45.0	49	36	71	1.1
1	other	external	253.0	283	241	375	1.1
2	other	external	106.0	131	107	173	1.2
1	sexual	external	194.0	238	200	331	1.2
2	sexual	external	170.0	197	163	277	1.2
1	beggary	internal	24.0	40	24	84	1.7
2	beggary	internal	23.0	50	31	102	2.2
1	other	internal	60.0	130	85	258	2.2
2	other	internal	52.0	84	54	166	1.6
1	sexual	internal	304.0	581	401	1094	1.9
2	sexual	internal	360.0	701	492	1333	1.9

Annex II: Past MSE studies - United Kingdom

In the United Kingdom, the obligation to identify presumed victims of human trafficking is discharged by the NRM, a framework for identifying victims and ensuring they receive appropriate protection and support. Its datasets are managed by the United Kingdom Human Trafficking Centre (UKHTC) of the Home Office. The National Crime Agency (NCA) of the Home Office collates data from various sources to produce Strategic Assessments of presumed victims. In 2013, 2,744 unique presumed trafficking victims were identified. The information about presumed victims came from a large number of separate source organizations. This information can be summarized into five lists based on the source type:

- LA: Local Authority
- NG: Non-governmental organization
- PF: Police force/National Crime Agency
- GO: Government Organization (mostly Home Office agencies e.g. UK Border Force, Gangmasters Licensing Authority)
- GP: The general public, through various routes

Of the 2,744 victims included in the 2013 database some appeared on two and a few on three or four of the five lists. Table 1 shows the distribution of the identified victims over the five lists.

Table 1: Contingency table for the National Crime Agency Strategic Assessment data, 2013\*

LA	X					X	X	X							X	X	
NG		X				X			X	X	X				X	X	X
PF			X				X		X			X	X		X	X	
GO				X				X		X		X		X	X		X
GP					X						X		X	X			

Source: Silverman (2014).

\*Each column shows the number of cases which fall in the combination of lists indicated by the cells marked.

Columns corresponding to patterns which do not occur in the observed data are omitted. The bottom row of Table 1 gives the numbers of presumed victims falling under each of the possible categories. MSE allows an estimation of the number of individuals not appearing on any of the lists, given the distribution of individuals in the contingency table. This is done by assuming that each of the counts is derived from a Poisson distribution, a distribution for the occurrence of rare events. A restrictive Poisson log-linear model is estimated for each of the cells and the parameter estimates are projected on the cell with the non-appearing (or hidden) individuals. With log-linear modeling, it is possible to assess how much being on one particular list affects a person’s chances of being on another. Possible interactions between lists can be detected, and controlled for in the estimates. The condition of independence can therefore be relaxed.

Bales, Hesketh and Silverman (2015) fitted a log-linear model to the data presented in Table 1 which allows for individual list effects, and also for interaction between lists. The estimated number of victims was 11,304. The 95 per cent confidence interval for the actual population size was estimated between 10,000 to 13,000, including the 2,744 victims already known.

This suggests that the Strategic Assessment was aware of roughly 20 per cent to 30 per cent of all possible victims in the UK in 2013. In round numbers, therefore, the dark figure is around 7,000 to 10,000.

There is a positive correlation between list LA and each of lists NG and PF, so that being known to the local authority apparently increases the chance of being known to NGOs or the police. This may reflect the existence of referral pathways for potential victims between these agencies, in particular in relation to minors who, unlike adults, do not need to consent to referral to the NRM, or joint operations between the local authorities and other agencies. The upshot of the exploratory MSE carried out on the lists of the National Crime Agency is that the true number of victims during 2013 is estimated at 11,300, or four times the numbers of detected victims (2,744).

Annex III: Past MSE studies - Netherlands

A multiple systems estimation (MSE) was carried out using the statistics on possible victims identified by different groups of organizations reporting to CoMensha1 on behalf of the Dutch National Rapporteur on Trafficking in Human Beings and Sexual Violence against Children over a period of six consecutive years (2010 - 2015) (UNODC, 2017). In total six different groups of organizations (lists) reported to CoMensha, among which the Border Police. The presumed victims reported by the Border Police concern presumed victims of a particular type of trafficking that is not informed by the Palermo Protocol. Moreover, this type of trafficking is no longer upheld as human trafficking by the Supreme Court in the Netherlands. For these reasons, two log-linear models were fitted: one including those reported by the Border Police (based on six lists, concerning 8,234 presumed victims between 2010 and 2015), and one excluding those exclusively reported by the Border Police (based on five lists, concerning 6,935 presumed victims between 2010 and 2015). To enhance the robustness of the estimates, they were stratified by four covariates, namely age (minor/adult), gender (female/male), form of exploitation (sexual/non-sexual) and nationality (Dutch/non-Dutch).

A model search was carried out using the stepwise selection procedure of the R-package STEP. This procedure is similar to well-known stepwise regression analyses, that is, it starts with a simple model and includes significant and deletes non-significant interaction terms between the lists until the fit of the model to the data is deemed adequate. The criteria for model selection are the Aikake Information Criterion (AIC) and/or the Bayesian Information Criterion (BIC); the model with the lowest AIC and/or BIC is selected. In the context of MSE models, we prefer the BIC because it offers better protection against overfitting.

The key finding of the analysis is that in 2014 and 2015, the most recent years for which records are available, the total number of presumed victims of human trafficking in the Netherlands was approximately 6,500 (six lists) / 6,250 (five lists) per year. This means that the estimated numbers are four to five times higher than the recorded numbers of victims that come to the attention of the authorities.

Figure 1 shows in black the trend in the total presumed victims recorded by CoMensha and in red the total estimated numbers of victims. The 95 per cent confidence interval around the estimates is depicted in shaded red.

Figure 1: Trends in observed presumed victims and estimated totals of presumed victims in the Netherlands, 2010-2015 (a: based on model with six lists, b: model with five lists)

