ILO Global Estimate of Forced Labour

Results and methodology

ILO Global Estimate of Forced Labour

Results and methodology

ILO Global Estimate of Forced Labour

Results and methodology

International Labour Office (ILO)

Special Action Programme to Combat Forced Labour (SAP-FL)

2012

Copyright © International Labour Organization 2012
First published 2012

Publications of the International Labour Office enjoy copyright under Protocol 2 of the Universal Copyright Convention. Nevertheless, short excerpts from them may be reproduced without authorization, on condition that the source is indicated. For rights of reproduction or translation, application should be made to ILO Publications (Rights and Permissions), International Labour Office, CH-1211 Geneva 22, Switzerland, or by email: pubdroit@ilo.org. The International Labour Office welcomes such applications.

Libraries, institutions and other users registered with reproduction rights organizations may make copies in accordance with the licences issued to them for this purpose. Visit www.ifrro.org to find the reproduction rights organization in your country.

ILO Cataloguing in Publication Data
ILO global estimate of forced labour: results and methodology / International Labour Office, Special Action Programme to Combat Forced Labour (SAP-FL). - Geneva: ILO, 2012
1 v.
ISBN: 9789221264125; 9789221264132 (web pdf)
International Labour Office
forced labour / role of ILO / data collecting / methodology / evaluation
13.01.2

The designations employed in ILO publications, which are in conformity with United Nations practice, and the presentation of material therein do not imply the expression of any opinion whatsoever on the part of the International Labour Office concerning the legal status of any country, area or territory or of its authorities, or concerning the delimitation of its frontiers.
The responsibility for opinions expressed in signed articles, studies and other contributions rests solely with their authors, and publication does not constitute an endorsement by the International Labour Office of the opinions expressed in them.
Reference to names of firms and commercial products and processes does not imply their endorsement by the International Labour Office, and any failure to mention a particular firm, commercial product or process is not a sign of disapproval.
ILO publications and electronic products can be obtained through major booksellers or ILO local offices in many countries, or direct from ILO Publications, International Labour Office, CH-1211 Geneva 22, Switzerland. Catalogues or lists of new publications are available free of charge from the above address, or by email: pubvente@ilo.org
Visit our web site: www.ilo.org/publns

Printed in Switzerland
Design and cover : Caroline Chaigne-Hope, Aurélie Hauchère Vuong

ACKNOWLEDGEMENTS

The production of this global estimate involved a great many people, in a genuinely collaborative process. ILO wishes gratefully to acknowledge the indispensable contributions made by everyone who was involved in this team effort.

The project was conducted by the Special Action Programme to combat Forced Labour (SAP-FL) of the ILO Programme to Promote the Declaration on Fundamental Principles and Rights at Work.

The contributions of the following people are specifically and gratefully acknowledged:

Consultant statisticians to the ILO

- Mr. Farhad Mehran
- Ms. Michaëlle de Cock

They were technical advisors to ILO for the 2012 global estimate, and principal authors of the revised methodology and of this report. It is thanks to their dedication and hard work that the production of the new estimate was possible.

Members of the independent peer review group

- Ms. Fiona David, Executive Director, Global Research and Policy, Walk Free, Sydney, Australia
- Mr. Siddharth Kara, Carr Center for Human Rights Policy, Kennedy School of Government, Harvard University, Cambridge, Massachusetts, USA
- Mr. Ravi Srivastava, Centre for the Study of Regional Development, School of Social Sciences, Jawaharlal Nehru University, New Delhi, India
- Mr. J.J.M. Van Dijk, Tilburg Law School, The Netherlands.

Sincere thanks are extended for their expert guidance for refinement of the estimation methodology, including their willingness to travel to Geneva for consultations to finalize the methodology.

ILO research assistants

- Natalia Alexandrova
- Anetha Awuku
- Sarah Gamha
- Barbara Granatelli
- Hélène Michalak
- Zhibo Qiu
- Sai Rong
- Kassiyet Tulegenova

These eight graduate researchers were recruited specifically for data collection. Appreciation is extended to them all, for the enthusiasm, diligence and commitment with which they approached this task. Ms. Anetha Awuku deserves special mention for her additional inputs during the data validation, cleaning, estimation and report-writing stages.

ILO technical contributors

- Mr. Patrick Belser (Conditions of Work and Employment Branch, ILO)
- Mr. Frank Hagemann (International Programme for the Elimination of Child Labour (IPEC))
- Ms. Rosinda Silva (International Labour Standards Department)
- Ms. Valentina Stoevska (Department of Statistics)

They provided technical support throughout the estimation process.

Members of the ILO internal advisory group on the global estimate, and all participants, from both the ILO and external organizations, in the expert round-table discussion held at ILO, Geneva on 7 May 2012, during which the draft estimation methodology was presented and discussed.

ILO production team

- Ms. Beate Andrees, Head, SAP-FL. Project manager
- Ms. Delphine Bois, SAP-FL. Administrative assistant
- Ms. Caroline Chaigne-Hope, DECLARATION. Design and lay-out of the final report
- Ms. Aurélie Hauchère, SAP-FL. Training and supervision of data collection
- Ms. Caroline O'Reilly, DECLARATION. Technical advisor and report editor

Table of contents

1. Introduction 11

2. Main results 13

2.1 Results by form of forced labour 13

2.2 Results by sex and age group of victims of forced labour 14

2.3 Results by region 15

2.4 Results by migration of the victims 16

2.5 Can the 2012 and 2005 global estimates of forced labour be compared? 17

3. What is forced labour? 19

4. Estimation of reported forced labour 21

4.1 Choice of methodology 21

4.2 The basic statistical unit: a reported case of forced labour 21

4.3 Capture-recapture sampling 22

4.4 Capture-recapture probabilities 24

4.5 Stratification 26

5. Data collection, entry and validation 29

5.1 The data collection phase 29

5.2 The data entry and cleaning phase 30

5.3 Data validation and matching 31

5.4 Respect of the capture-recapture assumptions 32

6. Estimation of total stock of reported and unreported forced labour 35

6.1 Stock versus flow estimates of forced labour 35

6.2 Duration in forced labour 36

6.3 Proportion of reported forced labour 38

6.4 Breakdown by sex, age group and migration 39

7. Evaluation of the results 41

7.1 Margins of error 41

7.2 Concluding remarks 42

Annex 1 Regional classification

Annex 2 Data entry templates

Introduction

1

In 2005, the International Labour Office (ILO) published its first global estimate of forced labour.[1] The estimate (a minimum of 12.3 million persons in forced labour at any point in time in the period 1995-2004) received considerable attention by governmental and non-governmental organizations and in the media. It has since been widely cited as the most authoritative estimate of the largely hidden, and therefore difficult to measure, phenomenon of forced labour. The estimate served its main purpose – to raise global awareness of the magnitude of the crime of modern day forced labour, and to stimulate action at all levels against it.

The capture-recapture methodology applied was also subject to scrutiny, particularly by the academic community and certain government agencies. A number of issues were raised concerning the underlying assumptions of the methodology and the procedure by which the extrapolation was made.

The purpose of the present document is to describe in detail the revised methodology used to generate the 2012 ILO global estimate of forced labour, covering the period from 2002 to 2011, and the main results obtained.

The revised methodology has been developed by the ILO in close collaboration with a Peer Review Group composed of four members of the academic community who are experts in the subject of forced labour and human trafficking. It follows the same basic two-stage approach used for the 2005 estimate, but incorporates certain improvements derived from ILO's own experience in the period since 2005, the availability of new primary data sources, feedback received from external experts on the 2005 methodology, and suggestions made by the ILO's statistical consultants and the peer reviewers.

Given the changes made in the methodology and the greater quantity and quality of data available for the present round of estimation, the numerical results of the 2012 estimation are not comparable to those derived in 2005. The 2012 estimate is no longer labeled as a minimum estimate, although it is still regarded as somewhat conservative due the nature of the capture-recapture methodology and the still limited national survey data available for extrapolation purposes.

1 ILO : *A Global Alliance against Forced Labour,* Global Report under the follow-up to the ILO Declaration on Fundamental Princinples and Rights at Work, Geneva, 2005; Belser, de Cock, Mehran: *ILO Minimum Estimate of Forced Labour in the World*, ILO, Geneva, 2005.

Main results

2

The ILO estimates that 20.9 million people are victims of forced labour globally, trapped in jobs into which they were coerced or deceived and which they cannot leave. Human trafficking can also be regarded as forced labour, and so this estimate captures the full realm of human trafficking for labour and sexual exploitation, or what some call "modern-day slavery".[2] The data from which the estimate derives cover the study reference period of 2002-2011. The estimate therefore means that some 20.9 million people, or around three out of every 1,000 persons worldwide, were in forced labour at any given point in time over this ten-year period.

This figure represents a conservative estimate, given the strict methodology employed to measure this largely hidden crime. It is not, however, labeled as a minimum estimate. With a standard error of 1,400,000 (7%), the range of the estimated global total is between 19,500,000 and 22,300,000, with a 68% level of confidence.

2.1 Results by form of forced labour

Forced labour is classified into three main categories or forms: forced labour imposed by the State, and forced labour imposed in the private economy either for sexual or for labour exploitation.

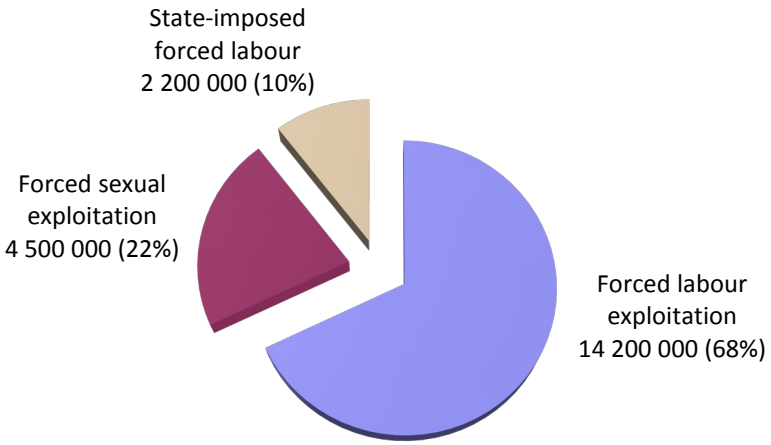
The distribution of the global total number of victims by form of forced labour is shown in Figure 1.

Of the total number of 20.9 million forced labourers, 18.7 million (90%) are exploited in the private economy, by individuals or enterprises. Out of these, 4.5 million (22% total) are victims of forced sexual exploitation, and 14.2 million (68%) are victims of forced labour exploitation, in economic activities such as agriculture, construction, domestic work and manufacturing. The remaining 2.2 million (10%) are in state-imposed forms of forced labour, for example in prison under conditions which contravene ILO standards on the subject, or in work imposed by the state military or by rebel armed forces.[3]

2 The figures do not include trafficking for the removal of organs or for forced marriage/adoption unless the latter practices lead to a situation of forced labour or service.

3 All percentages and numbers are rounded.

Figure 1. Global estimate by form of forced labour



2.2 Results by sex and age group of victims of forced labour

The results are presented in Figures 2 and 3. Women and girls represent the greater share of total forced labour – 11.4 million victims (55%), as compared to 9.5 million (45%) men and boys. Adults are more affected than children; 74% (15.4 million) of victims fall in the age group of 18 years and above, whereas children aged 17 years and below represent 26% of all forced labour victims (or 5.5 million children).

Figure 2. Global estimate by sex of victims of forced labour

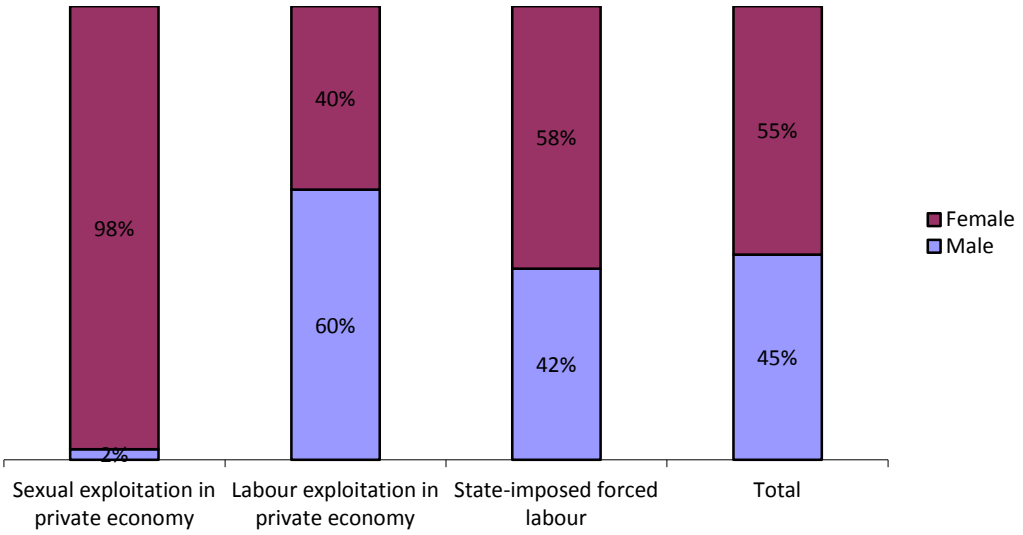
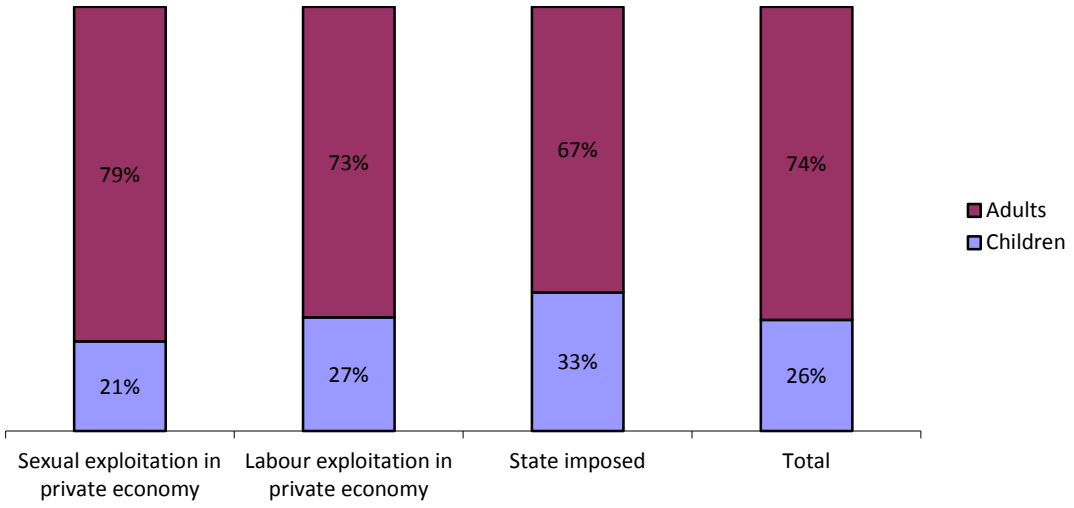


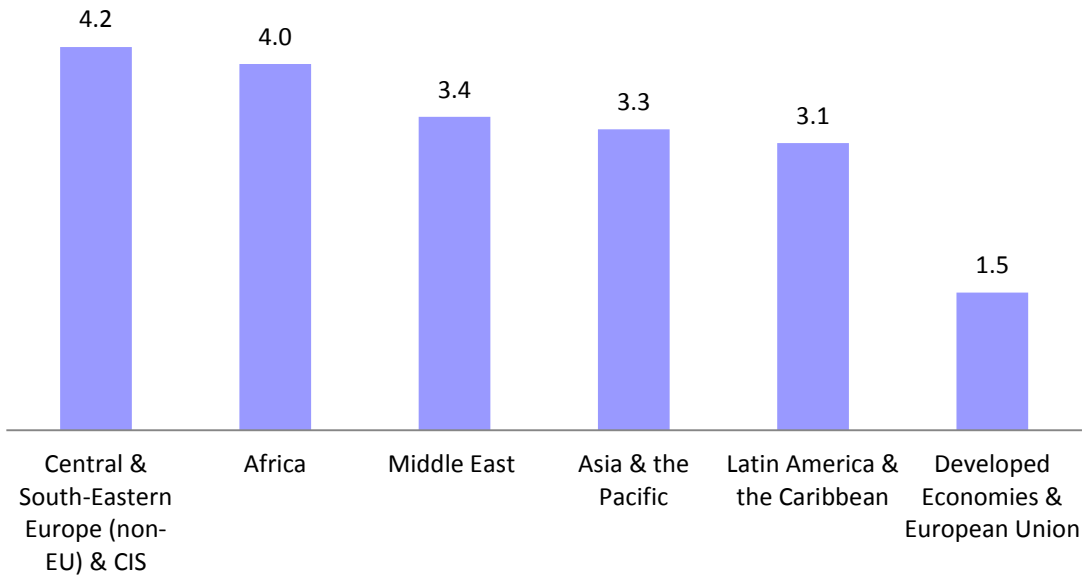
Figure 3. Global estimate by age group of victims of forced labour



2.3 Results by region[4]

When the prevalence of forced labour (number of victims per thousand inhabitants) is examined, the rate is highest in the Central and South-Eastern Europe and Commonweatlth of Independent States (CSEE & CIS) and Africa (AFR) regions at 4.2 and 4.0 per 1,000 inhabitants respectively, and lowest in the Developed Economies & European Union (DE & EU) at 1.5 per 1,000 inhabitants (Figure 4). The Middle East (ME), Asia-Pacific (AP) and Latin America and the Caribbean (LA) regions lie in the middle of the range, at 3.4, 3.3 and 3.1 per 1,000 respectively. The relatively high prevalence in Central and South Eastern Europe and CIS reflects the fact that the population is much lower than for example in Asia, while reports of trafficking for labour and sexual exploitation and of state-imposed forced labour in the region are numerous. The low rate in the Developed Economies and European Union may be attributed to the more effective regulatory mechanisms in place in these countries.

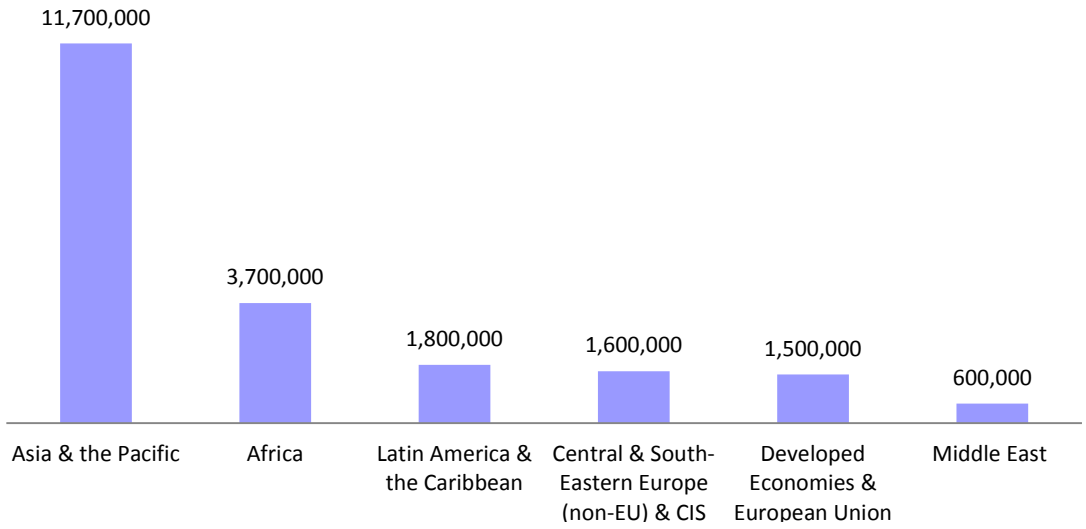
Figure 4. Prevalence of forced labour by region (per 1,000 inhabitants)



4 Regional classification is based on that used in ILO Key Indicators of the Labour Market (2011). The listing of countries by region is presented in Annex 1. It should be noted that the regional classification, and country composition of regions, differ from those used in the 2005 global estimate.

By contrast, when the distribution of forced labour (in absolute numbers) is examined, the pattern looks very different (Figure 5). The Asia-Pacific region accounts for by far the highest absolute number of forced labourers – 11.7 million or 56% of the global total. The second highest number is found in Africa at 3.7 million (18%), followed by Latin America and the Caribbean with 1.8 million victims (9%). The Developed Economies and European Union account for 1.5 million (7%) forced labourers, whilst countries of Central, Southeast and Eastern Europe (non-EU) and the Commonwealth of Independent States have 1.6 million (7%). There are an estimated 600,000 (3%) victims in the Middle East.

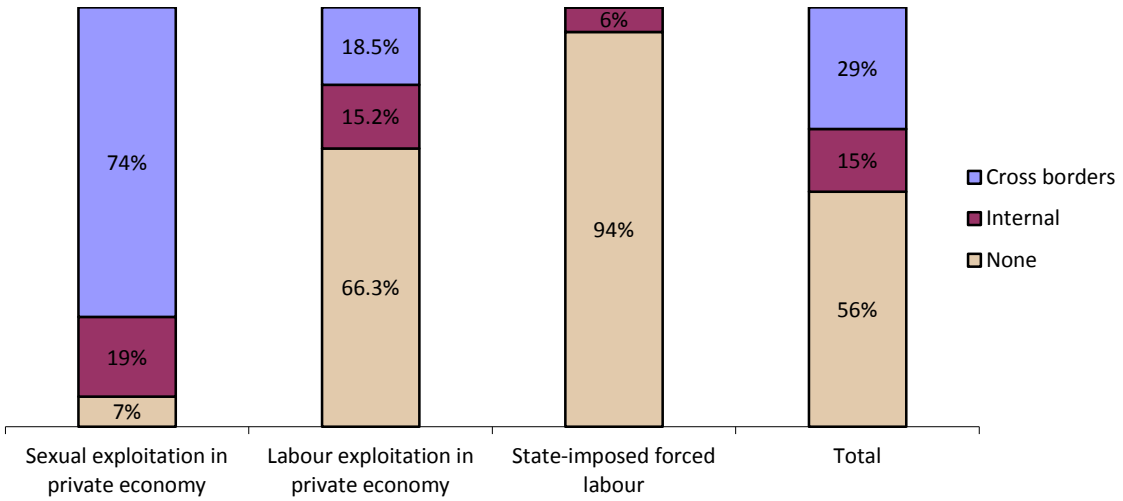
Figure 5. Estimate of forced labour by region



2.4 Results by migration of the victims

The estimates were broken down according to three categories of migration by the forced labour victims: cross-border migration, where the victims have left their country of origin to work in another country where the forced labour took place; internal migration where the persons have left their place of origin or residence and become victims of forced labour in another location within the same country; and no movement, where the forced labour occurs in the location where the victims usually reside. The main results are presented in Figure 6 below.

Figure 6. Global estimate by migration of the victims of forced labour



There are 9.1 million victims (44% of the total) who have moved either internally or internationally, while the majority, 11.8 million (56%), are subjected to forced labour in their place of origin or residence. Cross-border movement is strongly associated with forced sexual exploitation. By contrast, a majority of forced labourers in economic activities, and almost all those in state-imposed forced labour, have not moved away from their home areas. These figures indicate that movement can be an important vulnerability factor for certain groups of workers, but not for others.

2.5 Can the 2012 and 2005 global estimates of forced labour be compared?

The 2012 estimates cannot be compared to those from 2005 for the purpose of discerning trends over time, i.e. whether forced labour has increased or decreased over the period concerned. What can be said is that ILO has produced a more robust estimate in 2012, based on a more sophisticated methodology and more and better data sources.

Nonetheless, it will be noted that this estimate, at 20.9 million victims globally, is considerably higher than ILO's first estimate of a minimum of 12.3 million forced labourers. Another major difference from the 2005 estimate is that state-imposed forced labour represents a lower proportion of the total, at around 10%. This could be due in part to the fact that far fewer data are available on state-imposed forced labour relative to the other forms, pointing to a need for further research in this area.

There is also a divergence in the age distribution of forced labourers between the respective estimates. Children now account for an estimated 26% of all victims, a smaller proportion of the total than was the case in 2005.[5] The new data, however, confirm the ILO's previous conclusion that women and girls are more affected by forced labour, and particularly by forced sexual exploitation. Nonetheless, men and boys account for an overall 45% of all forced labour victims. Finally, while regional comparisons are complicated by the changes made in the country composition of regions between the two estimates, the Asia-Pacific region retains its place as the region harbouring the greatest absolute number of forced labourers in the world, although its proportion of the total has decreased somewhat (to just over one-half of all victims). The share and number of victims in Africa has, by contrast, increased in the current estimate (18% or nearly one-fifth of the total), which it is believed represents a more accurate reflection of reality, thanks to better reporting in the region.

The new estimates on movement, which were not calculated in 2005, demonstrate the fact that cross-border movement is closely allied with forced sexual exploitation, whereas a greater proportion of victims of non-sexual forced labour were exploited in their home area. An interesting new piece of information to emerge from the estimation process is that the average period of time that victims spend in forced labour is approximately 18 months, with significant variation across forms of forced labour and regions.

5 While the estimated proportion of forced child labourers now is less than in the 2005, their absolute number remains more or less unchanged. This may be explained in part by the fact that, in 2005, forced labour of children was better reported on than that of adults. Also, as reported by the ILO in both 2006 and 2010, the number and proportion of children in hazardous work (often used as a proxy for the worst forms of child labour) has progressively declined. See ILO: *The end of child labour: Within reach*, Report to the International Labour Conference, 95th Session, Geneva, 2006; and ILO: *Accelerating action against child labour*, Report to the International Labour Conference, 99th Session, Geneva, 2010.

What is forced labour?

3

The term "forced or compulsory labour" is defined by the ILO Forced Labour Convention, 1930 (No. 29), Article 2.1, as "*all work or service which is exacted from any person under the menace of any penalty and for which the said person has not offered himself voluntarily*". The definition thus contains three main elements: first, some form of work or service must be provided by the individual concerned to a third party; second, the work is performed under the threat of a penalty, which can take various forms, whether physical, psychological, financial or other; and third, the work is undertaken involuntarily, meaning that the person either became engaged in the activity against their free will or, once engaged, finds that he or she[6] cannot leave the job with a reasonable period of notice, and without forgoing payment or other entitlements. Forced labour is thus not defined by the nature of the work being performed (which can be either legal or illegal under national law) but rather by the nature of the relationship between the person performing the work and the person exacting the work. While sometimes the means of coercion used by the exploiter(s) can be overt and observable (e.g. armed guards who prevent workers from leaving, or workers who are confined to locked premises), more often the coercion applied is more subtle and not immediately observable (e.g. confiscation of identity papers, or threats of denunciation to the authorities). Forced labour therefore presents major challenges in terms of detection, for the purposes of both data collection and law enforcement.

Convention No. 29 provides for certain exceptions, with respect to military service for work of a purely military character, normal civic obligations, work as a consequence of a conviction in a court of law and carried out under the control of a public authority, work in emergency situations such as wars or natural calamities, and minor communal services in the direct interest of the community involved (Article 2.2).

A later ILO Convention, the Abolition of Forced Labour Convention, 1957 (No. 105) further specifies that forced labour can never be used as a means of political coercion or education or as punishment for expressing political views or for participating in strike action, of labour discipline, of racial, social, national or religious discrimination, or for mobilising labour for economic development purposes.

Forced labour includes practices such as slavery, practices similar to slavery, debt bondage and serfdom – themselves defined in other international instruments, namely, the League of Nations Slavery Convention (1926)[7] and the United Nations Supplementary Convention on the Abolition of Slavery, the Slave Trade, and Institutions and Practices Similar to Slavery (1956). Further, the definition of forced labour has been interpreted by the ILO Committee of Experts on the Application of Conventions and Recommendations (CEACR) to encompass trafficking

6 While the Convention uses the term "himself", it applies equally to women and men, girls and boys.

7 The Slavery Convention (1926) defines slavery as "the status or condition of a person over whom any or all of the powers attaching to the right of ownership are exercised" (Article 1(1)).

in persons for the purpose of exploitation,[8] as defined by the Palermo Protocol to Prevent, Suppress and Punish Trafficking in Persons, especially Women and Children.[9]

The Palermo Protocol defines trafficking in persons as the recruitment, transportation, harbouring or receipt of persons, by means of coercion, abduction, deception or abuse of power or of vulnerability, for the purpose of exploitation. It goes on to specify that exploitation shall, at a minimum, include sexual exploitation, forced labour, slavery and slavery-like practices.[10] There is therefore a clear link between the Protocol and ILO Convention No. 29. The only type of exploitation specified in the Protocol's definitional article that is not also covered by ILO Convention No. 29 is trafficking for the removal of organs. It should be noted that, as the scope of the estimate is limited to forced labour, no attempt was made to estimate trafficking of adults or children for forced marriage, adoption or organ transplant. Nonetheless, cases in which victims were deceived with false promises of marriage or adoption but were instead put into situations of forced labour were taken into account.

Forced labour affects both adults and children, and both age groups are covered by Convention No. 29. With the adoption of the ILO Worst Forms of Child Labour Convention, 1999 (No. 182)[11], forced labour, slavery, debt bondage and trafficking of children (aged less than 18 years) became, additionally, "worst forms" of child labour, for which immediate action should be taken for their abolition as a matter of urgency. Apart from the explicit inclusion of "*forced or compulsory recruitment of children for use in armed conflict*", there is no specific definition of what constitutes forced labour of children, and therefore the definition specified in Convention No. 29 applies. Any work or service undertaken by a child is considered as forced child labour where some form of coercion or deception is applied by a third party, either directly to the child or to his or her parents, in order to oblige the child to take a job or perform a task, or to prevent him or her from leaving the work. It is clear, therefore, that not all child labour amounts to forced labour, as in many instances children work, frequently under hazardous conditions, in the absence of any third party coercion. Such child labour should of course be eliminated, but does not constitute forced labour of children.[12]

---

8 ILO (2007) *Eradication of forced labour*. General Survey concerning the Forced Labour Convention , 1930 (No.29), and the Abolition of Forced Labour Convention, 1957 (No. 105), Report of the Committee of Experts on the Application of Conventions and Recommendations, Report III (Part 1B), ILC, 96th Session, Geneva, 2007, para. 77.

9 The Protocol supplements the United Nations Convention against Transnational Organized Crime (2000). The Protocol criminalizes trafficking in persons, whether this occurs within countries or across borders, and whether or not conducted by organized criminal networks.

10 The full definition of trafficking in persons in Article 3 of the Protocol is: "*the recruitment, transportation, transfer, harbouring or receipt of persons, by means of the threat or use of force or other forms of coercion, of abduction, of fraud, of deception, of the abuse of power or of a position of vulnerability or of the giving or receiving of payments or benefits to achieve the consent of a person having control over another person, for the purpose of exploitation. Exploitation shall include, at a minimum, the exploitation of the prostitution of others or other forms of sexual exploitation, forced labour or services, slavery or practices similar to slavery, servitude or the removal of organs*." (Art. 3 (a)). It also specifies that "*The recruitment, transportation, transfer, harbouring or receipt of a child for the purpose of exploitation shall be considered "trafficking in persons" even if this does not involve any of the means set forth in subparagraph (a) of this article*" (Art. 3 (c)).

11 According to Convention No. 182, "*worst forms of child labour*" shall include "*all forms of slavery or practices similar to slavery, such as the sale and trafficking of children, debt bondage and serfdom and forced or compulsory labour, including forced or compulsory recruitment of children for use in armed conflict*".

12 For further guidance on this issue, refer to ILO: *Hard to see, harder to count. Survey guidelines to estimate forced labour of adults and children*, ILO, Geneva, 2012.

Estimation of reported forced labour

4

4.1 Choice of methodology

The methodology used to generate the 2012 estimate of forced labour is a refinement of that applied by the ILO in 2005, when it made its first global estimate of forced labour.

The reasons invoked by the ILO to justify the selection of its methodology in 2005 unfortunately remain partly true today. The continued lack of reliable national estimates based on specialized data collection instruments, prevents the use of the most usual means to derive global estimates which is to aggregate national estimates into regional and then global figures. The ILO has now developed and published guidelines for the design and implementation of such national surveys.[13] Pilot surveys on forced labour have been successfully implemented in more than ten countries, of which four were national in scope, but this is still an insufficient basis on which to derive a global estimate. More such surveys will have to be implemented in all regions of the world before such an extrapolation method can be applied.

Moreover, a systematic review of all papers published by academic or independent researchers regarding the 2005 estimate, shows that it has been generally well received by experts, who nonetheless identified some weaknesses in the methodology. The ILO therefore resolved to work to rectify these weaknesses to the extent possible, and the results of these efforts are presented in this paper. As was the case in 2005, the revised method relies on a double sampling of "reported cases" of forced labour, from which a global estimate of reported and non-reported cases can then be extrapolated.

4.2 The basic statistical unit: a reported case of forced labour

The unit of information that was sampled was a "reported case" of forced labour. A case is defined as a recorded piece of information about one or more persons who are currently, or have been, victims of forced labour during the reference period (2002-2011). These cases were found in a wide variety of secondary sources, as described in Section 5.

In order to qualify for inclusion in the sample, a "reported case" had to contain details, at a minimum, on the following four elements:

- an activity (work or service) which amounted to forced labour, and which could be classified according to the typology presented in Figure 8;
- a numerical figure indicating the number of victims;
- a geographical location where the activity took place; and
- a date or time period during which the persons were in forced labour, which falls within the reference period of 2002-2011.

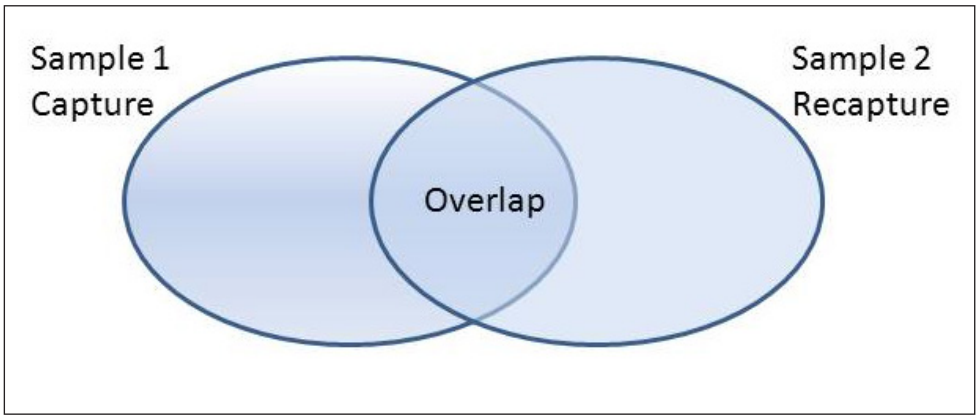
13 ILO (2012). *Op.cit.*

4.3 Capture-recapture sampling

The capture-recapture method of sampling was originally developed for estimating the size of elusive populations such as the number of fish in a lake or the abundance of animal populations in a forest, where there are no sampling frames available. It has since been used in a variety of other applications in social science research such as estimating the number of homeless people in a city.

In the present context, the basic idea of the method is to sample cases of forced labour from the universe of all reported cases of forced labour and then to resample the same universe, so as to find the fraction of cases in the second sample that was also identified in the first sample. The elements of the method are schematically shown in Figure 7.

Figure 7. Schematic representation of capture-recapture sampling



where the rectangle represents the universe of reported cases of forced labour of unknown size, say, N, and the two ellipses (Sample 1 and Sample 2) represent respectively the capture and recapture samples of reported cases of forced labour drawn from the universe. Let

n_1 = Number of cases of forced labour found in sample 1

n_2 = Number of cases of forced labour found in sample 2

n_{12} = Number of cases of forced labour found in both samples (overlap)

Denoting by n the total number of distinct cases found in the two samples and by x the unknown number of cases not found in any of the two samples, the following relationships may be expressed among the various values:

$$N = n + x$$
$$n = n_1 + n_2 - n_{12}$$

The unknown values x and N may be estimated under four assumptions, as follows:

- Closed population, i.e. there is no change in the universe of reported cases of forced labour during the course of the study;
- Cases sampled on both occasions are correctly identified as forced labour and matched with each other without error;
- Each case has an equal chance of being selected in either sample, i.e. there is equal probability of selection in the two samples;
- The selection of a case in the second sample is not influenced by its selection in the first sample, i.e. the two samples are drawn independently of each other.

From these assumptions, particularly the independence assumption, it can be deduced that the probability of selecting a particular case in sample 2 is unchanged whether or not the case was selected in sample 1. In mathematical terms, this means the conditional probabilities of selection in sample 2 should be equal, whether the case was selected in sample 1 or not.

Prob (in sample 2| in sample 1) = Prob (in sample 2|not in sample)

Replacing the conditional probabilities by their values

$$\frac{n_{12}}{n_1} = \frac{n_2 - n_{12}}{N - n_1}$$

and solving the equation for the unknown value N, one finds

$$n_{12}N - n_{12}n_1 = n_1n_2 - n_1n_{12}$$

$$n_{12}N = n_1n_2$$

$$N = \frac{n_1n_2}{n_{12}}$$

Thus, the unknown total number of reported cases of forced labour (N) is estimated by multiplying the number of cases found in sample 1 (size of capture n_1) with the number of cases found in sample 2 (size of recapture n_2), and dividing the result by the number of cases found in both samples (size of overlap n_{12}). The corresponding estimate of the number of cases not found in either sample can be calculated by $x = N-n = (n_1-n_{12})(n_2-n_{12})/n$.

The formula for calculating N and x can lead to overestimates when used with relatively small sample sizes. Thus, generally, the following slightly modified versions are used in practice:[14]

$$N = \frac{(n_1+1)(n_2+1)}{(n_{12}+1)} - 1$$

$$x = \frac{(n_1 - n_{12})(n_2 - n_{12})}{(n_{12}+1)}$$

After deriving the estimate of the total number of reported cases of forced labour (N) based on capture-recapture sampling, the corresponding estimate of the total number of victims of forced labour (M) is obtained by multiplying the total number of reported cases by the average number of victims per case

(1) $$M = N \times m$$

where m is the average number of victims per reported case of forced labour.

Because these estimates are based on random sampling, they are subject to sampling errors. The sampling error or sampling variability of N and m, and therefore that of M, can be calculated from the sample data themselves in terms of their standard deviations expressed as the square root of variances:

$$Var(N) = \frac{(n_1+1)(n_2+1)(n_1 - n_{12})(n_2 - n_{12})}{(n_{12}+1)^2(n_{12}+2)}$$

14 Thompson, Steven, K., *Sampling*, Ch. 18 Capture-Recapture Sampling, A Wiley-Interscience Publication, John Wiley & Sons, Inc., New York, 1992, pp. 212-224.

$$Var(m) = (1 - \frac{n}{N})\frac{\sigma_m^2}{n}$$

where n is the number of distinct cases of forced labour in the two samples as defined earlier and

$$\sigma_m^2 = \frac{\sum_i (m_i - m)^2}{(n-1)}$$

m_i being the number of reported victims in case i. Finally, the variance of the total number of reported victims is obtained by applying the rule for variance calculation for the product to two independent variables[15]

$$Var(M) = m^2 Var(N) + N^2 Var(m) + Var(N)Var(m).$$

4.4 Capture-recapture probabilities

The basic capture-recapture model assumes a binomial probability distribution of the sample cases. According to this model, a forced labour report is either "captured" or "not captured" by a given team with respective probabilities p and 1-p. The values of p are the same for all reports but may differ between the teams, say $p=p_1$ for team 1 and $p=p_2$ for team 2.

With an expanding universe of reported cases of forced labour and heterogeneity of sample selection, this simple binomial model may no longer be appropriate. It may be more suitable to consider more general models for describing the probabilities of occurrences of forced labour cases in the sample, such as the Poisson distribution.

Consider an extension of the capture-recapture model where there is more than one recapture.[16] Let

$$f_o, f_1, f_2, \ldots, f_k$$

represent the frequencies of cases of forced labour identified by the ILO teams 0, 1, 2, …, k times, respectively (f_o is unobserved). The unknown total number of reported cases may be expressed as

$$N = f_o + f_1 + f_2 + \ldots + f_k$$

and the combined sample size of distinct cases found by the teams is

$$n = f_1 + f_2 + \ldots + f_k$$

Assuming that the frequencies f_i, i=0,1, … ,k are generated by a Poisson distribution with parameter λ, the unknown total number of reported cases of forced labour may be estimated by

$$N = \frac{n}{1 - p_o}$$

15 Goodman, Leo A.: "On the Exact Variance of Products," in *Journal of the American Statistical Association*, December 1960, pp. 708-713.
16 Chao, A.: "Estimating the population size for capture-recapture data with unequal catchability," in *Biometrics,* Vol. 43, 1987, pp. 783-791, and Chao, A.: "Estimating population size for sparse data in capture-recapture experiments", in *Biometrics,* Vol. 45, 1989, pp. 427-438. Other proposed methods to deal with heterogeneous capture-recapture probabilities include: Huggins, R. M. and P. S. F. Yip: "A note on nonparametric inference for capture-recapture experiments with heterogeneous capture probabilities", in *Statistica Sinica*, 2001, Vol. 11, pp. 843-853; Cormack, R.M.: "Log-linear models for capture-recapture," in *Biometrics*, Vol. 45, 1989, pp. 395-413, and Lavallée, P. and L.-P. Rivest: "Capture-Recapture Sampling and Indirect Sampling," in *Journal of Official Statistics*, Vol. 28, No. 1, 2012, pp. 1-27.

where p_o is the probability of zero occurrence of a reported case in the combined sample, p_o=exp(-λ), with λ estimated by $2f_2/f_1$. Another estimate of N is given by

$$N = n + \frac{f_1^2}{2f_2}$$

which in terms of the capture-recapture notations introduced earlier may be expressed as $f_1=(n_1-n_{12})+(n_2-n_{12})$ and $f_2=n_{12}$ where the elements n_1, n_2 and n_{12} are calculated based on the frequency of the reported cases without elimination of duplicates within teams.

The corresponding estimate of the global number of reported victims of forced labour is derived by multiplying the estimate of the total number cases N by the average number of victims per case m as described in the previous section. The sampling variance of the final estimate is calculated by the formula for the variance of products of random variables mentioned earlier, except that the variance of N is now obtained from

$$\mathrm{var}(N) = \frac{f_1^2}{2f_2}\left[1 + 2\frac{f_1}{f_2} + \frac{1}{2}\left(\frac{f_1}{f_2}\right)^2\right]$$

In the implementation of this methodology for the 2012 global estimate, out of the total of 5,491 reported cases of forced labour found in the capture-recapture exercise, 4,069 were found only once and 1,422 were found multiple times. Cases found more than once tended to contain higher numbers of victims than those found only once. For example, the average number of victims in cases found just once was 223. In cases found twice, it was 253 and in cases found more than twice, it was 812.

The frequency distribution of the reported cases by the number of times they were found is shown in Table 1. Sometimes, the same report was found by different members of the same team, and in some cases, even by the same person at different times, for example, in different languages.

Table 1. Frequency distribution of reported cases of forced labour by number of occurrences in the capture-recapture sample

Number occurrences	*Reported cases*	*%*
Total	**5491**	**100**
1	4069	74.1
2	1186	21.6
3	167	3.0
4	46	0.8
5	10	0.2
6	7	0.1
7	3	0.1
8	1	0.0
9	0	0.0
10	1	0.0
11	1	0.0

No formal goodness-of-fit test was carried out due to the missing number of reported cases with no occurrence in the capture-recapture sample. But the Poisson distribution seems to fit the data reasonably well, at least for

the first three frequencies. The ratio of the first to the second frequencies (4069/1186= 3.43) is almost correctly proportional to the ratio of the second to the third frequency (1187/167=7.11). Under the Poisson distribution, the ratio of these ratios (3.43/7.11=0.48) should be equal to the ratio of the respective number of occurrences (2/3=0.67).

4.5 Stratification

In order to improve the accuracy of the results and to provide estimates broken down by region and by form of forced labour, the data were stratified in terms of three categories:

i. Geographical region

The geographical stratification is based on the regional classification used in the 2011 edition of *ILO Key Indicators of the Labour Market*, as follows:

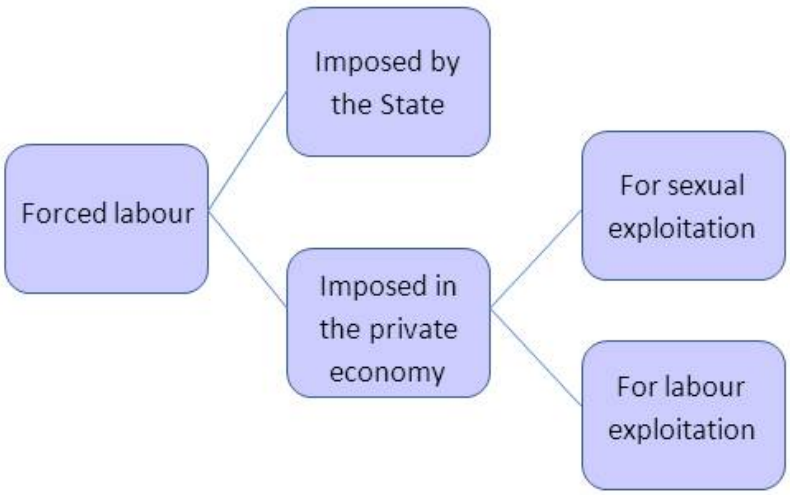
- Developed Economies and European Union
- Central and South-Eastern Europe (non-EU) and the Commonwealth of Independent States
- Asia-Pacific
- Latin America and the Caribbean
- Middle East
- Africa

The list of countries under each region is reproduced in Annex 1. It differs from that used in the 2005 global estimate, as certain new regions have since been designated, and some countries have been shifted from one region to another.

ii. Form of forced labour

The stratification by form of forced labour follows the basic typology used in 2005, with one main modification - no breakdown is given for trafficking. The revised typology is presented in Figure 8. The "mixed" category that was included in 2005, covering victims of both sexual and labour exploitation, has also been eliminated.

Figure 8. Typology of forced labour by form



iii. Type of data

The data were also stratified by type, distinguishing between two categories: first, data describing actual cases in which one or more persons are, or were, in forced labour at the same time, place etc (called “incidents”) and second, data providing information on forced labour cases compiled over a period of time (called “aggregates”). Typically, aggregate data refer to a larger number of victims and their probability of selection in the capture-recapture samples is higher than that of data on incidents. No estimated data of any kind were included in the database.

The statistical treatment of aggregate data differed from that of incidents. The data on incidents were subject to capture-recapture sampling by geographical region and by form of forced labour, while the data from aggregate data were subject to capture-recapture without stratification. The resulting estimate obtained from the aggregate data was then distributed by geographical region and form of forced labour according to the distribution of the incident data. This differential treatment was adopted to improve the robustness of the results. It was noted that the incident data demonstrated a much higher degree of detail and precision than the data from the aggregate data.

Data collection, entry and validation 5

5.1 The data collection phase

The double sampling, capture and recapture, was conducted simultaneously during thirteen consecutive weeks by two teams, based in the ILO headquarters in Geneva: Team 1 (which could be associated with Capture) and Team 2 (associated with Recapture). Each team was composed of four graduate research assistants of different nationalities. The researchers all had the same level of education, having recently completed or being about to complete a Master's Degree in Political Sciences, International Relations or Economics. The researchers were selected on the basis of their knowledge of human rights, labour issues and international migration, and also on their languages skills. The requirement was for members of each team, collectively, to be able to read and communicate in the same eight languages: Arabic, Chinese, English, French, Italian, Portuguese, Russian and Spanish.

All researchers followed an intensive two-day training course on forced labour and human trafficking at the start of their assignment, during which the relevant concepts and definitions were explained in detail. This was followed by practical training in data collection, during which the various research tools were presented, including the database into which the researchers would code the cases they found. They read some "dummy" reports prepared by the ILO, analyzed them and decided whether or not they constituted cases of forced labour, through group and individual exercises, with feedback provided by the trainers. The proposed estimation methodology and the statistical treatment of the data were presented, including the four assumptions underlying capture-recapture. The necessity for the teams not to exchange any data or source of information was emphasized. The teams were located in different parts of the ILO building.

During the first two weeks of data collection, all researchers kept aside the documents containing cases they had rejected, so that the ILO expert could validate their decision, explaining in cases of incorrect assessment, why the case should in fact have been retained. Approximately 2,500 different sources were located containing reported cases of forced labour. Reports were mainly identified through internet searches but also, to a lesser extent, by telephone, email, visits to libraries and face-to-face interviews.

The sources of information used, in decreasing order of frequency, were:

- media reports (including newspapers, radio, TV, internet media)
- local, national, regional, international or thematic NGO
- government documents, from Ministries of Justice, Labour, Social Affairs, Migration, Foreign Affairs and Interior or from special police or other units dedicated to combating trafficking and forced labour
- other international organizations, through their national offices or headquarters
- academic reports
- ILO reports, including those of the Committee of Experts on the Application of Conventions and Recommendations (CEARC)
- trade union reports and employers' organization reports.

As briefly explained above, the researchers looked for two main types of data:

- "Incidents" of forced labour, which are cases in which one or more persons are described as being or having been in forced labour together, whether at the same location, trafficked by the same person, working for the same employer, or freed during the same operation (police or NGO raid, for example). Incidents are usually reported in considerable detail, describing aspects such as the recruitment process, working conditions, means of coercion, etc.
- "Aggregate" data, which are statistics published on a periodic basis, usually by organizations working on some aspect of forced labour. Typical examples are police reports of the number of victims identified over a given period of time, and NGO reports of the number of trafficking victims sheltered. While some reports give information on the nationality of the victims, or the industry in which they were working, it is quite rare to find other details such as the length of time that victims spent in forced labour.

A third type of data was reference data, which provided contextual or other information relevant to the forced labour situation. Researchers entered such sources in the database if they encountered ones of interest, but did not deliberately look for them.

5.2 The data entry and cleaning phase

Once classified by the researcher as being a case of forced labour, a description of the case and its source was entered in an Access database designed for this exercise. Two databases were compiled, one by each team, with an identical structure. The two databases were located on different computers, and there was no link between them.

The structure of the database allowed for entering data other than the variables necessary for the computation of the global estimate. This additional information will be exploited by ILO in the future to produce thematic reports, for example, on legal responses to forced labour (court cases, sentences, fines) and on the economics of forced labour. Two data entry forms were used: a first to describe the source of information, and a second to describe details of the cases reported in the source. The data entry forms are presented in Annex 2.

Each report was first coded with ten variables: unique identifier for report, title, organization, source, authors, type of source (International Organization, NGO, government, etc.), website, ISBN, date of publication and language.

Each forced labour case was described by 72 variables:

- a unique identifier, generated by the system (1 variable)
- source of information (4 variables)
- coding of the type of data: incident, aggregate or reference (2 variables)
- description of the victims (number, age, sex, ethnic or religious group) (11 variables)
- geographical details of the case, including of migration if relevant (6 variables)
- indicators of threats and penalty, forced or deceptive recruitment, working and living conditions imposed by the employer and impossibility to change employer (29 variables)
- dates and duration in forced labour (6 variables)
- economic data (6 variables)
- exit process (1 variables)
- judicial data (2 variables)
- form of forced labour (State/non-State; sexual/labour exploitation) and migration (2 variables)
- summary and comment (2 variables).

A hidden control was created to record the day the case was entered, to allow for subsequent verification.

A printed copy of the source was filed for the archives and, in the case of a non-written source (such as video footage or a phone call), a short summary was written and printed. In addition, the (internet) translations into English of reports written in Arabic, Chinese and Russian were filed, which were needed for validation of the data by the ILO expert.

The data cleaning phase aimed to detect and correct typographical and other errors or missing values. The consistency between the case summary (50 to 100 words) and the main variables was checked, with particular emphasis on those concerning the indicators of forced labour. The original source of information was often consulted to rectify the error or to enter missing variables. The most common missing values were on the dates and duration of exploitation.

5.3 Data validation and matching

The aim of the data validation phase was to identify and mark all data (both incident and aggregate) which fulfilled four conditions for their retention in the database for the purpose of calculation of the estimate:

- the working situation could be qualified as one of forced labour;
- the number of victims is specified;
- the exploitation took place between 2002 and 2011; and
- the place where the exploitation took place is indicated.

For the validation of **incidents**, the first and most important assessment was the determination of the presence or absence of the main forced labour indicators, involuntariness and penalty (or menace of a penalty), using the 28 variables in the database relating to these indicators. In order to be validated, the incident had to show one or more indicators of involuntariness and penalty. The only exception to this was in cases reported by the judiciary. If the judicial report was of a person condemned in court for a forced labour offence, the case was assumed to be genuine even if no details were given, and the stated number of victims was included in the database.

The second element was to know in which type of activity the victim was working; for this, the minimum requirement was to have sufficient information to classify the incident as one of either state-imposed forced labour or of forced labour or sexual exploitation in the private economy. The third element for validation was the number of victims, ideally by their age and sex, or at least by the total number. A case described as involving "many" or "dozens of" women, for example, was not validated. Many cases, especially in the media, were somewhat surprisingly reported without any indication of when and where they actually took place. "Victim testimonies" included in reports, such as *"I was ten years old when I was sold to a master and became his slave"*, often fell into this category. Such cases were discarded as there was no evidence to show that they were not invented simply for communication purposes.

Regarding **aggregate** data, many are published with very little detail included. When the information needed to validate the data was missing, the source of the information was evaluated instead. Thus, reports coming from sources considered as reliable, such as national police reports or other governmental sources, international organizations and established international NGOs, the data were assumed to be valid and were retained. Other sources were only validated if the data were presented in sufficient detail to indicate their authenticity. When no information was given on the type of activity performed by the victims, as was often the case with aggregate data on trafficking, the cases were coded as "sexual and labour exploitation".

The results of the data collection and validation process are presented in Table 2.

Table 2: Number of reported cases of forced labour collected and validated

	No. of cases collected	No. of cases validated
Team 1	3,959	3,644
Team 2	4,171	3,875
Total	8,132	7,519

The next stage was to match the cases common to both teams, i.e. cases which have been both captured and recaptured. It also involved identifying duplicates within a single team. Even though the teams developed their own procedures for minimizing duplicates, such as by allocating different countries to each team member according to

their language skills, duplication still occurred. For example, a case of international trafficking might be reported in both the country of origin and of destination of the victim and therefore be captured by two researchers working in the same team.

"Matching" of incidents (whether within one team or common cases between the two teams) was done by comparing the following elements in each case:

- the number of persons involved
- the location
- the date of exploitation or release
- the description of the form of exploitation
- and, on occasion, the names of the exploiters.

When incidents were described in detail, matching could be quite straightforward. But in other cases, especially where the incident was reported in different languages or the data were not entirely consistent between different sources, the matching process was longer and more difficult.

Matching was also often challenging for "aggregate" data. Official statistics were generally the easiest to match in this category. For example, aggregate data on the number of victims identified by the police in a given country in a given year could be quite easily matched, even when the same data were published in a report by another organization or even in a different language, as the original data source was most often cited.

However, when different sources reported information possibly regarding the same persons, matching became much more problematic. For example, were trafficking victims sheltered by the International Organization for Migration (IOM) in a given country the same persons as those identified and reported by that country's government in the same year? The following rule was applied in such cases: If the number of victims, place, type of exploitation and date (or period) were the same in both cases, they were treated as duplicates. If one of these elements was different, and there was no means to prove that both cases related to the same victims, the cases were assumed to be different.

At the end of this process, a new table was created of validated cases, indicating whether each case was found by only one team or by both. A unique identifier was then assigned to identify each separate case for the purpose of calculation.

5.4 Respect of the capture-recapture assumptions

In order to be valid, the capture-recapture procedure must respect certain assumptions as detailed below.

Capture-recapture must take place in a closed population

The universe for capture-recapture is the set of all reports on forced labour available worldwide during the three months of data collection. Because capture and recapture took place simultaneously, it can be assumed that no report entered or left the universe between the two sampling exercises. In reality, it is likely that some reports disappeared and others appeared during this period (e.g. all new reports published and some websites or reports which closed or disappeared from the internet during the three months of data collection). But, as the three-month period of the study may be considered to be sufficiently short, it is not believed that any departure from this assumption had a significant impact on the results.

Cases sampled must be correctly identified as forced labour and matched with each other without error

The identification of cases of forced labour was done using the lists of indicators for incident data and verifying the reliability of the source of information for aggregate data. In case of doubt, a conservative judgement was made. For example, a newspaper report about exploitative working conditions, even if judged by the journalist to be forced

labour or slavery, would not be validated as such if there was no evidence of deception or coercion. Matching of cases was the most time-consuming phase of the procedure. It was done by a meticulous sequence of checks, successively refining the criteria whenever there was a suspicion of duplicates. Quite often, the original sources were retrieved, or even new sources located which contained all the elements needed to make the assessment.

In the basic capture-recapture model, all cases must have an equal probability of being sampled

The underlying assumption of equal probability in the capture-recapture methodology is reflected in the binomial distribution, which may not be appropriate given the heterogeneity of the data sources. It is clear that a case reported in the US Department of State's annual Trafficking in Persons report had a higher probability of being sampled than a case reported by a local NGO working in an African community. One way to allow for unequal probability of selection in the capture-recapture methodology is the use of the Poisson distribution. Selecting a report of forced labour is a "rare event" and the frequency of its selection follows the probability distribution of "rare events", which is generally formulated by a Poisson distribution. For the calculation of the 2012 estimate, the Poisson distribution was used to describe the probabilities of occurrences of forced labour cases in the sample.

The two phases (capture and recapture) must be independent of each other

What would foster interaction between the data collection teams and thus decrease their level of independence? The main reason would be direct leaks of information, either because of friendship between researchers and their willingness to help each other, or by having access to cases by "stealing" from the other team without first looking for them. It cannot be proved that this did not happen, but it is believed to be highly unlikely. The researchers had fully understood the importance of this rule, which was explained to them in detail during the training. During the data collection period, several opportunities were taken to re-emphasize to the researchers that they were personally responsible for the quality of the data and respect of the rules. Some random checks were also made on the dates on which common reports were found by the two teams, and there was no evidence that "sharing" had occurred.

Estimation of total stock of reported and unreported forced labour

6

The result of the capture-recapture sampling is an estimate of the total number of persons reported to have been victim of forced labour at some time during the ten-year reference period, called here "flow estimate of reported forced labour." There are two important aspects to this estimate. First, it counts equally all persons, irrespective of the length of time they spent (duration) in forced labour; thus victims of forced labour for a few days are counted in exactly the same way as others who were victims during the entire ten-year period. Second, it includes only those cases and victims of forced labour that were reported, and does not account for those that went unreported.

The next step in the estimation process is therefore to extrapolate the capture-recapture estimate to cover both reported and unreported cases of forced labour, and also to make an adjustment for the differences in duration in forced labour among the victims. The result of this step provides a stock estimate of total forced labour, which may be interpreted as being the "full-time equivalent" number of victims of forced labour during the entire ten-year period, or alternatively as the average stock of persons in forced labour at any given point of time.[17]

6.1 Stock versus flow estimates of forced labour

The first element of the estimation method is to convert the capture-recapture estimate of reported forced labour over time into a stock estimate at a given point of time. This involves multiplying the capture-recapture estimate (called here the reported flow estimate) by the average duration in forced labour measured as a fraction of the length of the reference period. Thus, denoting by μ_h the average duration in forced labour in stratum h, the total reported stock of forced labour at a point of time during the reference period is expressed by

(2) $$T_{reportedstock} = \sum_h T_{reportedflow_h} \mu_h$$

where $$T_{reportedflow_h} = M_h$$

M_h being the capture-recapture estimate of forced labour measured in stratum h as described in section 4.

The second element of the estimation method involves the decomposition of the total stock of forced labour into the reported stock and the unreported stock as follows

(3) $$T_{stockforcedlabour} = T_{reportedstock} + T_{unreportedstock}$$

17 Kutnick, Belser, and Danailova-Trainor: *Methodologies for global and national estimation of human trafficking victims: current and future approaches*, Working Paper No. 29, Programme on Promoting the Declaration on Fundamental Principles and Rights at Work, Special Action Programme to combat Forced Labour, ILO, Geneva, 2007.

Denoting by r_h the ratio of reported to unreported forced labour in stratum h, the combination of (2) and (3) leads to

$$T_{stockforcedlabour} = \sum_h T_{reportedflow_h} \mu_h + \sum_h T_{reportedflow_h} \frac{\mu_h}{r_h}$$

$$= \sum_h T_{reportedflow_h} (1 + \frac{\mu_h}{r_h})$$

$$= \sum_h T_{reportedflow_h} \frac{\mu_h}{p_h} \quad (4)$$

where p_h is the proportion of reported forced labour in total forced labour.

The result (4) indicates that the total stock of forced labour at a point of time is a weighted sum of the capture-recapture estimates in the different strata, where the weight in each stratum is the ratio of the average duration in forced labour to the proportion of reported forced labour in that stratum.

In the 2005 estimate, the weights were assumed all equal to 1, resulting in the simplified equation

total stock of reported and unreported forced labour at a point of time

= total flow of reported forced labour over time

The justification for this simplified equation was the assumption of a positive relationship between the probability of reported cases of forced labour and the duration in forced labour, i.e. forced labour of short duration is less likely to be reported than forced labour of long duration. Assuming that the relationship is continuous and monotone (not necessarily linear), it was argued that the curve describing the relationship intersects the line of equality somewhere in the mid-range of the duration in forced labour. At the point of intersection, the duration in forced labour is equal to the reporting probability. It further assumed that the amount of forced labour to the left side of the point of intersection is essentially the same as the amount to the right in each stratum, then the assumption that (μ_h/p_h=1) should lead to a reasonable approximation of the total stock of forced labour in that stratum.[18]

6.2 Duration in forced labour

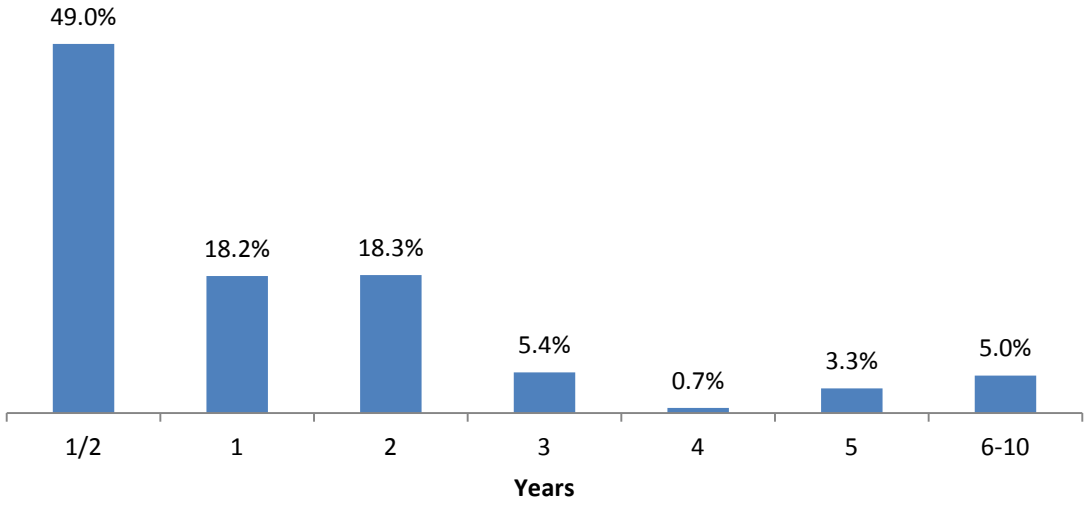
A refinement introduced in the 2012 methodology was the use of data on duration in forced labour, as recorded in the database. Fields were established in the database to record the dates on which forced labour started and ended. In practice, however, the duration data were found in a variety of formats, such as “one week”, “1 or 2 years”, “over two months”, “during the harvest season”, “in 2008”, “a couple of months”, etc.

For analysis, it was decided to use the month as the unit of measurement, and the year for reporting purposes. Rules were established to convert the reported information on duration into a number of months or fraction of a month. Thus, for example, “one week” was converted to “0.25 months”, “over two months” or “a couple of months” to “2 months”, “during the harvest season” to “3 months”, “in 2008” to “12 months” and “1 or 2 years” to “18 months”.

18 The argument would also be valid if the relationship between duration in forced labour and the probability of reporting is in fact monotonously decreasing, as long as the duration-reporting curve crosses at some point the line of equality and the mass of forced labour on the two sides of the point of intersection is equal.

Overall, 797 reported cases contained data on the duration in forced labour. The size distribution of duration of forced labour is presented graphically below. Reported durations of longer than 120 months were truncated to 120 months, corresponding to the ten-year reference period.

Figure 9: Duration in Forced Labour



According to these results, nearly half of all reported spells of forced labour were six months or less. More than one third were between one and two years. Some 8 % of cases were, however, of forced labour episodes lasting 5 years or more. Overall, the average duration in forced labour, across all forms, was 17.7 months, or about 1.5 years.

The range of duration in forced labour varied among geographic regions and forms of forced labour. Duration in state-imposed forced labour was found to be somewhat lower (at about 7 months) than forced labour either for commercial sexual exploitation (about 17 months) or for other exploitation (about 19 months) in the private economy.

However, the duration data obtained from the reported cases are unlikely to be representative of the duration in forced labour of unreported cases. Almost all reported cases of forced labour follow some sort of intervention resulting in the termination of the forced labour episode. Hence, the reported duration of forced labour can be expected to be shorter than would be the case in the absence of such an intervention, as in an unreported case.

This issue is similar to the difference between interrupted and completed spells of unemployment. The measured duration of unemployment in household surveys is the duration of unemployment up to the time of the survey, generally referred to as the interrupted spell of unemployment. The completed spell of unemployment is the total span of unemployment experienced by the individual before obtaining employment or leaving the labour force, irrespective of the timing of the survey.

Drawing from this parallel, a distinction was made in the 2012 methodology between interrupted and completed spells of forced labour. The data obtained from the reported cases were considered to be estimates of the duration of interrupted spells of forced labour.

By comparison, the victims of forced labour in unreported cases would experience a longer duration in forced labour, called here the completed spell. The duration of completed spells of forced labour is taken to be twice as long as the duration of interrupted spells. This is because the timing of the intervention may be assumed to be independent of the duration of forced labour and therefore to fall, on average, just in the middle of the forced labour episode.

Given the truncation of 120 months covering the maximum span of the ten-year reference period, the average duration of completed spells of forced labour measured in months is derived by the following relationship:

Completed spell of forced labour = min (120, 2*Interrupted spell)

The estimated average duration of completed spells of forced labour was found to be 29.4 months, or less than twice the estimated average duration of interrupted (reported) forced labour of 17.7 months.

A related issue concerns repeated spells of forced labour. In practice, many victims of forced labour experience discontinuous, though often linked, episodes of forced labour. For example, some seasonal migrants in brick kilns or in quarries may be treated as being in forced labour for the duration of a single season, but the episodes may recur every season, linked to each other through continued indebtedness to the same employer. Such repeated spells of forced labour are accounted for through the duration estimates. As the unit of measurement is a reported case of forced labour, and each forced labour spell has a chance of being reported, repeated spells are in principle represented in the capture-recapture sample.

6.3 Proportion of reported forced labour

The estimation of the global stock of reported and unreported forced labour requires, in addition to data on duration in forced labour, information on the share of reported to total forced labour. This corresponds to the values of p in the expression (4) of the total stock of forced labour. In 2005, the need for this information was addressed by assuming that the ratio of the average duration in forced labour to the probability of reporting is equal to one.

With the availability of new data from four national surveys on forced labour conducted recently by the ILO, the 2012 methodology attempted to directly estimate the proportion of reported cases.

Because three of these surveys addressed returned migrants residing in their home areas, the respondents were no longer in a situation of forced labour. Given this, it can reasonably be assumed that they answered the survey questions accurately, and the survey results can thus be considered to cover all cases, both reported and unreported, of forced labour in these countries.

In general, let T_i be the estimate of forced labour in country i based on a national survey with reference period t_i. The estimate refers to the total number of persons who experienced forced labour at some time during the t_i-year reference period of the survey. The corresponding number for a ten-year reference period can be obtained proportionally as $(120/\mu_i)T_i$, where μ_i is the average duration in forced labour, measured in months, for country i.

The comparison of the adjusted survey result with the capture-recapture estimate for the corresponding country provides an estimate of the share of reported number of victims in total forced labour in that country,

$$p_i = \frac{M_i}{(\frac{120}{\mu_i})T_i}$$

where M_i is the capture-recapture estimate of reported forced labour in the same country.

A pooled estimate of p is obtained by smoothing the four individual country estimates using the duration data as auxiliary variable. Thus the pooled estimate is given by the expression

$$p = \frac{e^{a+b\mu}}{1+e^{a+b\mu}}$$

where a=-4.774 and b=6.107 are the estimated parameters of the logistic regression of the survey data on p and μ (measured as fraction of the 120 month reference period), and $\mu = p\mu_1 + (1-p)\mu_2$ is the overall average duration in forced labour calculated as a weighted average of duration of interrupted spells of reported cases $(\mu_1 = 17.7/120 = 0.146)$ and completed spells of unreported cases $(\mu_2 = 29.4/120 = 0.245)$. The resulting estimate of p obtained by iteration is

$$p = 3.6\%$$

According to this result, on average, for every reported case of forced labour, about 27 cases go unreported. This estimate of p is based on the four national surveys on forced labour currently available to ILO, and can be improved upon in the future when more national survey data become available.

Because of the fragility of the estimate of p, no attempt has been made to use separate estimates of p for the different strata as expression (4) of the global estimate would require. The same value of p=3.6% is therefore used for all regions and all forms of forced labour.

6.4 Breakdown by sex, age group and migration

In the final step of the estimation process, the global estimate was disaggregated by sex of the victim and by age group (adult and child) for each form of forced labour. In addition, a breakdown by "migration" was computed, for the categories of internal migration, cross-border migration and no migration.

The breakdown was based on information obtained from the reported cases. Data on sex of the victim were available in 1,860 cases, on age group in 2,184 cases and on movement in 2,208 cases. The unknown values were distributed in proportion to the known values within each form of forced labour. For age group, additional information from the aggregate data was also used.

The observed proportions obtained from the reported cases were first applied to the global estimate for all forms of forced labour combined. The estimates by form of forced labour were then derived by iterative proportional fitting. This procedure ensured consistency between the breakdown estimates and the global estimates obtained by capture-recapture sampling.

Evaluation of the results

7

7.1 Margins of error

Given the elusive nature of the target population and the limited direct measurement of the phenomenon, it would be unrealistic to expect global estimation of forced labour with a high degree of accuracy. Like all estimates based on sampling, the global estimates obtained here are subject to both sampling and non-sampling errors.

Sampling errors

Sampling errors arise from the fact that the estimate is based on observations from a sample of cases and generalization to the population at large. Had different samples been obtained on different occasions by the research teams, the resulting global estimate would also have been somewhat different. Because capture-recapture sampling produces a probability sample, the information obtained from the sample gives not only a global estimate of forced labour, but also an estimate of the sampling error. This sampling error, called standard error, is calculated based on variance of the estimate and also taking into account the additional variability due to the estimation of number of victims per case and of average duration in forced labour.

The resulting estimate of the standard error is found to be about 1,400,000. The standard errors of the different regional and component estimates of forced labour are presented in Table 3.

Table 3. Standard errors of global estimate of forced labour

	Estimate	Standard error
World	20,900,000	1,400,000
Asia-Pacific	11,700,000	800,000
Latin America and the Caribbean	1,800,000	600,000
Africa	3,700,000	300,000
Middle East	600,000	70,000
Central and South-Eastern Europe and CIS	1,600,000	200,000
Developed Economies and European Union	1,500,000	200,000
State-imposed forced labour	2,200,000	200,000
Forced sexual exploitation in the private economy	4,500,000	1,000,000
Forced labour exploitation in the private economy	14,200,000	500,000

The sampling errors may be interpreted in terms of confidence intervals. Thus, the unknown global number of victims of forced labour, estimated from this sample, is likely to be within an estimated range of roughly one standard error, which in the present context is

$$20{,}900{,}000 +/- 1{,}400{,}000$$

or in the range

$$19{,}500{,}000 - 22{,}300{,}000.$$

The level of confidence associated with this confidence interval is about 68%. Broader confidence intervals with higher confidence levels (95% or 99%) can be calculated using two or three standard errors.

Corresponding standard errors can also be calculated for the global estimates of forced labour by sex, age group and movement. For this purpose, the generalized variance procedure may be used based on the formula

$$\text{relative standard error} = \sqrt{(a+b/x)}$$

where the values of a and b are obtained by fitting an appropriate linear regression.

Non-sampling errors

An assessment of non-sampling errors such as coverage errors, classification errors, editing errors and data processing errors is extremely difficult in the present context. The detailed description of the data collection and validation procedures provided in section 5 demonstrates the extent of the efforts made to ensure high quality of the underlying data.

7.2 Concluding remarks

While acknowledging the limitations of the estimation methodology, ILO believes it to be the best possible given the current status of data on forced labour, and therefore the resulting estimates to represent the most reliable available at this time. As more and better information is produced, especially through primary surveys at national level, it will become possible to generate progressively more accurate estimates in the future. This will further strengthen the empirical basis for the design and implementation of more effective policy responses so as to end the crime of modern day forced labour.

–

Annex 1
Regional classification

Developed Economies & European Union

European Union
- Austria
- Belgium
- Bulgaria
- Cyprus
- Czech Republic
- Denmark
- Estonia
- Finland
- France
- Germany
- Greece
- Hungary
- Ireland
- Italy
- Latvia
- Lithuania
- Luxembourg
- Malta
- Netherlands
- Poland
- Portugal
- Romania
- Slovakia
- Slovenia
- Spain
- Sweden
- United Kingdom

North America
- Canada
- United States

Other Developed Economies
- Australia
- Israel
- Japan
- New Zealand

Western Europe (non-EU)
- Iceland
- Norway
- Switzerland

Central & South-Eastern Europe (non-EU) & CIS

Central & South-Eastern Europe
- Albania
- Bosnia and Herzegovina
- Croatia
- Serbia and Montenegro
- The former Yugoslav Republic of Macedonia
- Turkey

Commonwealth of Independent States
- Armenia
- Azerbaijan
- Belarus
- Georgia
- Kazakhstan
- Kyrgyzstan
- Republic of Moldova
- Russian Federation
- Tajikistan
- Turkmenistan
- Ukraine
- Uzbekistan

Asia

South Asia
- Afghanistan
- Bangladesh
- Bhutan
- India
- Maldives
- Nepal
- Pakistan
- Sri Lanka

South-East Asia
- Brunei Darussalam
- Cambodia
- East Timor
- Indonesia
- Lao People's Democratic Republic
- Malaysia
- Myanmar
- Philippines
- Singapore
- Thailand
- Viet Nam

Pacific Islands
- Fiji
- Papua New Guinea
- Solomon Islands

East Asia
- China
- Hong Kong, China
- Korea, Democratic People's Republic of
- Korea, Republic of
- Macau, China
- Mongolia
- Taiwan, China

Latin America & the Caribbean

Caribbean
- Bahamas
- Barbados
- Cuba
- Dominican Republic
- Guadeloupe
- Guyana
- Haiti
- Jamaica
- Martinique
- Netherlands Antilles
- Puerto Rico
- Suriname
- Trinidad and Tobago

Central America
- Belize
- Costa Rica
- El Salvador
- Guatemala
- Honduras
- Mexico
- Nicaragua
- Panama

South America
- Argentina
- Bolivia
- Brazil
- Chile
- Colombia
- Ecuador
- Paraguay
- Peru
- Uruguay
- Venezuela, Bolivarian Republic of

Middle East
- Bahrain
- Iran, Islamic Republic of
- Iraq
- Jordan
- Kuwait
- Lebanon
- Oman
- Qatar
- Saudi Arabia
- Syrian Arab Republic
- United Arab Emirates
- Occupied Palestinian Territory
- Yemen

Africa

North Africa
- Algeria
- Egypt
- Libyan Arab Jamahiriya
- Morocco
- Sudan
- Tunisia

Eastern Africa
- Burundi
- Comoros
- Eritrea
- Ethiopia
- Kenya
- Madagascar
- Malawi
- Mauritius
- Mozambique
- Réunion
- Rwanda
- Somalia
- Tanzania, United Republic of
- Uganda
- Zambia
- Zimbabwe

Middle Africa
- Angola
- Cameroon
- Central African Republic
- Chad
- Congo
- Congo, Democratic Republic of
- Equatorial Guinea
- Gabon

Southern Africa
- Botswana
- Lesotho
- Namibia
- South Africa
- Swaziland

Western Africa
- Benin
- Burkina Faso
- Cape Verde
- Côte d'Ivoire
- Gambia
- Ghana
- Guinea
- Guinea-Bissau
- Liberia
- Mali
- Mauritania
- Niger
- Nigeria
- Senegal
- Sierra Leone
- Togo

Annex 2
Data entry templates

Form 1 - Data source

No report | Title | Organisation | Source | Authors | Type of source

Website | ISBN | Date publication | Language:

Form 2 – Reported cases

Source
Ref
No Report
Page number
Source quoted
Quoting:

Reference data Type info:

Victims
Total no of victims
Nb adults
Nb men
Nb women
Nb children
Nb boys
Nb girls
Estimate
Status:
Population:
Details on population

Geography
Country FL
Region
Internal migratio
Country origin
Country transit
Other details

Recruitment
Forced recruitment
Deceptive recruitment
Recruitment by abuse of vulnerability
Details
Trafficker:

Working or living conditions
Forced overtime
Forced tasks
Living conditions imposed
Details

Impossibility to leave employer freely
Impossibility to leave employer
Details

Coercion
Threats of violence
Violence
Sexual violence
Debt bondage
Debt linked to travel
Debt linked to advance
Retention of wages
Retention of Id docs
Isolation
Confinement
Abuse of vulnerability
Threats of denunciation
Other punishment
Employer
Intervention of an intermediary
Intermediary
Case qualified as:
FL took place between
Beginning
End
Duration FL:
Date interview:
Details
Identified or released between
Date 1:
Date 2:

Economic data
Industry:
Occupation:
Salary:
Recruitment fee:
Economic data:
Urban/rural:

Exit process
Exit process:

Justice
Court case
Judgement:

Summary
Migration:
Typology

Comment
Validated:
Reason for no:

ILO Special Action Programme to combat Forced Labour (SAP-FL)
Programme for the Promotion of the Declaration on Fundamental Principles and Rights at Work

International Labour Office
4 route des Morillons
CH-1211 Geneva 22
Switzerland

Tel. + 41 22 799 63 29
Email: forcedlabour@ilo.org
www.ilo.org/forcedlabour