

# Measuring Women and Men's Work

Summary of  
Main Findings and  
Recommendations from  
a Joint ILO and World  
Bank Study in Sri Lanka



## Introduction

Between 2017 and 2019, the International Labour Organization (ILO) and World Bank completed a multistage pilot study, in collaboration with the Department of Census and Statistics (DCS) of Sri Lanka, with the objective of developing guidance, for various types of household surveys, on good practices in the measurement of women and men's work, as defined in the standards adopted by the 19th International Conference of Labour Statisticians (ICLS). The Sri Lanka pilot study was designed to enable a comparison of the outcomes of two types of household surveys, namely, the labour force survey (LFS) and the multitopic living standards

survey (MLSS). The study was supported by Data2X and the William and Flora Hewlett Foundation under the umbrella of the Women's Work and Employment Partnership.

The findings and their implications have been described in a joint ILO and World Bank report. A selection of the main findings is highlighted in Box 1 below and further elaborated in the following sections. It focuses on some of the key lessons that should be considered by anyone seeking to design a household survey to measure participation in work in line with the latest standards.

Box 1

## Summary of main findings and recommendations

- The Sri Lanka study highlighted important differences between the LFS and MLSS in the measurement of key labour-related indicators, e.g. participation in various forms of paid and unpaid work. These differences were reduced through relatively minor changes in questionnaire content (e.g. by adding recovery questions) and survey implementation.
- The risks of misclassification or measurement difficulties were concentrated among people engaged in casual, low-hours jobs or people helping on family farms or in businesses. This is highly relevant from a gender perspective because these types of activities were more common among women respondents in the pilot study, which is also likely to be true in many other settings.
- In addition, the measurement of working time in unpaid work appears to be highly sensitive to the measurement approach (e.g., the use of one vs. two questions to capture hours per week). This was especially evident in unpaid care work. Further work is planned on this issue to study the most efficient methods to capture good quality information on time spent doing unpaid household service work.
- All surveys should emphasize good translation and national adaptation, as well as interviewer training and supervision.
- The harmonization of questionnaire content may be a way to improve consistency in measurement, but it cannot be assumed that absolute consistency between the LFS and other household surveys can be achieved or that the need for a national process of adaptation and testing can be ignored.
- The findings presented in this report are being used to update the [guidance, tools and support of the ILO for LFSs](#) and the [World Bank for MLSSs](#). The model questionnaires, guidance and tools are excellent reference points for those facing the task of designing a questionnaire to capture work- and labour-related issues through a household survey in line with the latest standards.

# 1

## Summary of findings: identification of employment

### Measurement of employment – key messages

- The measurement of employment, as defined within the 19th ICLS standards, has been shown to be sensitive to survey design and content. The Sri Lanka study demonstrates that there is a higher risk of the underidentification of certain types of employment activities, such as helping without pay in family businesses or farms or casual jobs and small income-generating activities. Women are more likely to be involved in these types of activities than men. This creates a risk that the scale of gender gaps in employment may be misrepresented. These findings are consistent with earlier pilot studies by both the [ILO](#) and the [World Bank](#).
- Experience, including that gained from the study in Sri Lanka, has shown that these risks can be reduced with well-targeted, well-worded “recovery” questions that more directly reference the types of activities at risk of undercount in the national context.<sup>1</sup>
- Good adaptation to national context and translation to local languages are also

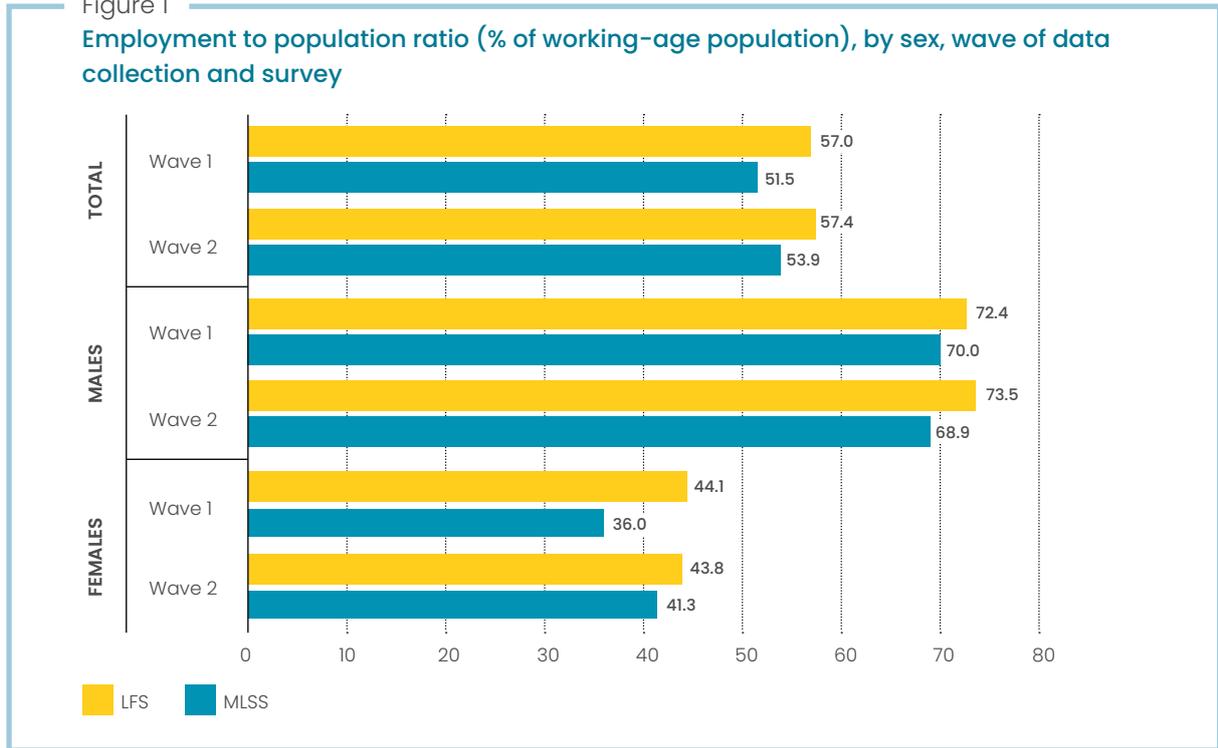
important prerequisites to achieving comprehensive coverage of employment in line with the latest standards.

The 19th ICLS standards (adopted in 2013) established an updated definition of employment that is focused on work done in return for pay or profit, a narrower definition than the previous definition established at the 13th ICLS (1982).

The Sri Lanka study showed that there is a clear risk of undercounting various types of working activities or of misclassifying employment and the own-use production of goods among those engaged in own-account farming. In wave 1 of data collection, the LFS identified 22 per cent more employed women than the MLSS, equivalent to an 8.1 percentage point difference in the measured employment to population ratio. The LFS also identified approximately 3 per cent more employed men, a 2.4 percentage point difference in the employment to population ratio. This resulted in a gap of 10 per cent overall between the surveys, a 5.5 percentage point difference in the employment to population ratio (see **Figure 1**). In-depth analysis of the data led to the conclusion that the gap emanated from the fact that the MLSS, which, unlike the LFS, did not initially include any recovery questions, did not fully capture people engaged in more casual, low-hours jobs, helpers in family businesses and farms or others in informal working activities, primarily women.

<sup>1</sup> Recovery questions are here defined as questions whose purpose is to “recover” persons who were not classified as employed during the core questioning designed to capture employment, even though they were engaged in activities that count as employment.

Figure 1  
**Employment to population ratio (% of working-age population), by sex, wave of data collection and survey**



Source: Joint DCS, ILO, and World Bank pilot study in Sri Lanka, Wave 1 and Wave 2, March–October 2019.

Changes to address these issues were successful in significantly narrowing the gap in wave 2 of data collection (a 6 per cent gap among both men and women, resulting in a 3.5 percentage point difference in the ratios of total employment to population in the two surveys).

This LFS-MLSS comparison highlights the key groups and types of workers at risk of misclassification and supports earlier findings from ILO, World Bank and academic studies that these risks are greater among women than among men.

These conclusions support the development of guidance on good measurement practices, such as the need for recovery questions, careful wording and translations into local languages, to ensure that people with “small” jobs or people helping in family businesses or farms are identified as employed in the survey. It was also notable that the revisions to the MLSS instrument, while important in the measurement of employment, also improved the measurement of own-use production work in agriculture (described below).<sup>2</sup>

<sup>2</sup> In the MLSS, a common set of questions was used to identify employment in agriculture (that is, agricultural work for pay or profit) and own-use production work in agriculture (that is, for own or family consumption). The distinction between these two concepts is fleshed out in subsequent questions, which seek information on the intended use of the agricultural outputs (for pay or profit or for own or family consumption). Any revisions that improve the ability of the MLSS to capture employment in agriculture will thus enhance the ability of the survey to measure own-use production in agriculture.

## 2

## Summary of findings: measuring other working activities

### Measurement of other working activities – key messages

- As with employment, the measurement of participation in other working activities is highly sensitive to survey content and context. For own-use production of goods, the questions targeting farming work need to be chosen with great care. The experience of the study demonstrates that it is important to use locally familiar terms and to mention a range of relevant activities in the question text.
- A particular sensitivity is required in the measurement of the time spent performing unpaid household service work, most notably the unpaid care of children and adults. Differences in wording and implementation that may be considered minor (such as a greater emphasis on active versus passive caregiving) led to substantial differences in the estimates of working time. Furthermore, to develop guidance on good measurement practices, a dedicated study of methods is needed on the measurement of the time spent in unpaid care and domestic work.

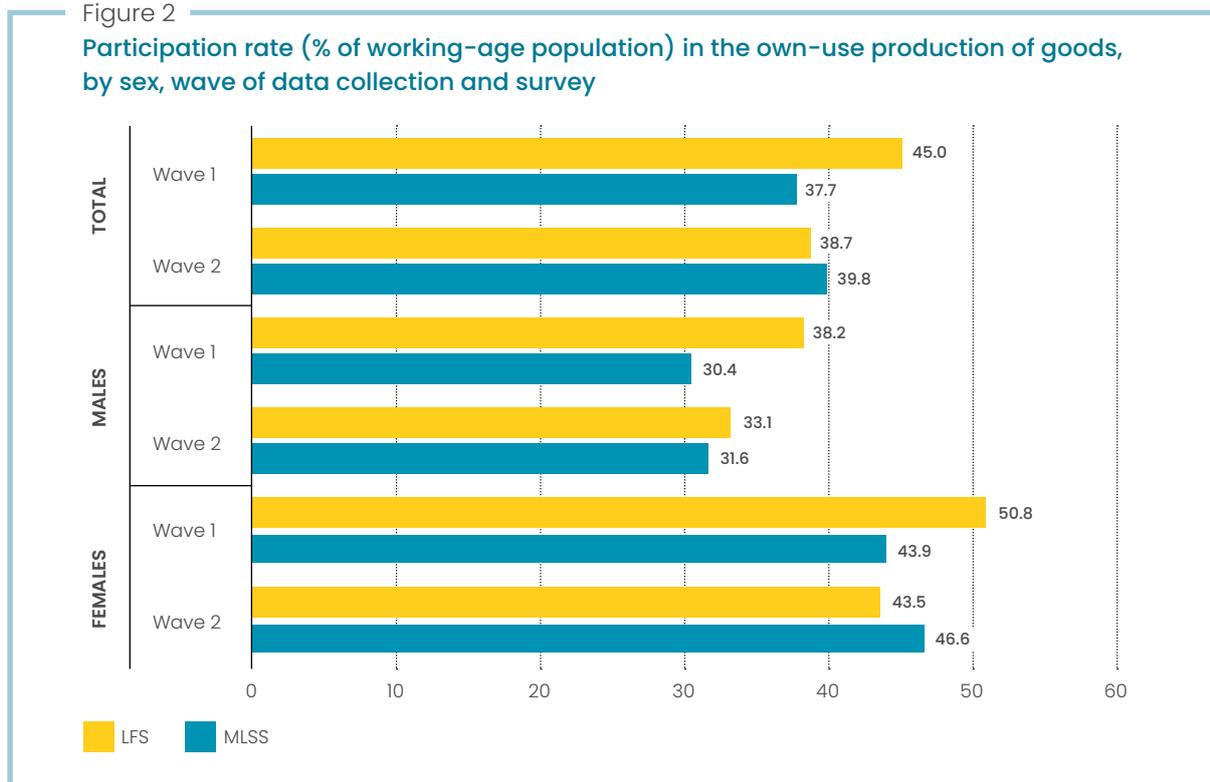
The 19th ICLS standards established a forms of work framework as a basis for identifying

the range of different paid and unpaid working activities in which people engage. This can reveal women and men's full working contribution and close important gender data gaps.

The Sri Lanka study also included questions on a selection of unpaid working activities. Specifically, work to produce goods for own-consumption was covered (called own-use production of goods in the standards), which includes subsistence farming, as well as unpaid work to provide services to the household (called own-use provision of services in the standards), such as housework, childcare and other activities that are predominantly carried out by women because of gendered social norms. The other forms of work defined in the standards, namely, unpaid trainee work and volunteer work, were not covered in the Sri Lanka pilot study.

In wave 1 of data collection, the LFS recorded a higher prevalence of both own-use production of goods and own-use provision of services than the MLSS. The difference was concentrated in crop farming, with relatively lower differences in other types of activities. This reflects the fact that (as described above) the MLSS identified fewer family helpers and fewer other marginal workers in farming. The updates undertaken after wave 1 reduced the recorded gap, with a relatively small difference observed in wave 2 (see **Figure 2**). This suggests that the additional

Figure 2 Participation rate (% of working-age population) in the own-use production of goods, by sex, wave of data collection and survey



Source: Joint DCS, ILO, and World Bank pilot study in Sri Lanka, Wave 1 and Wave 2, March–October 2019.

questions and wording updates (mentioned above) were successful in narrowing the gap between the LFS and the MLSS.

Even more notable was the sensitivity of data on hours worked in the own-use provision of services. While the MLSS identified fewer people engaged in these activities in wave 1, it did show a substantially higher average number of hours worked per week (34.2 versus 24.8 – see **Figure 3**). Analysis narrowed this down to a substantial difference between the two surveys in the reported hours spent on childcare and on care for dependent adults. A review of practices identified the source of the disparity to be a difference in implementation between the two surveys. While the two surveys used similar questions to identify individuals engaged in care work among adults and children, the LFS emphasized active caregiving (and included

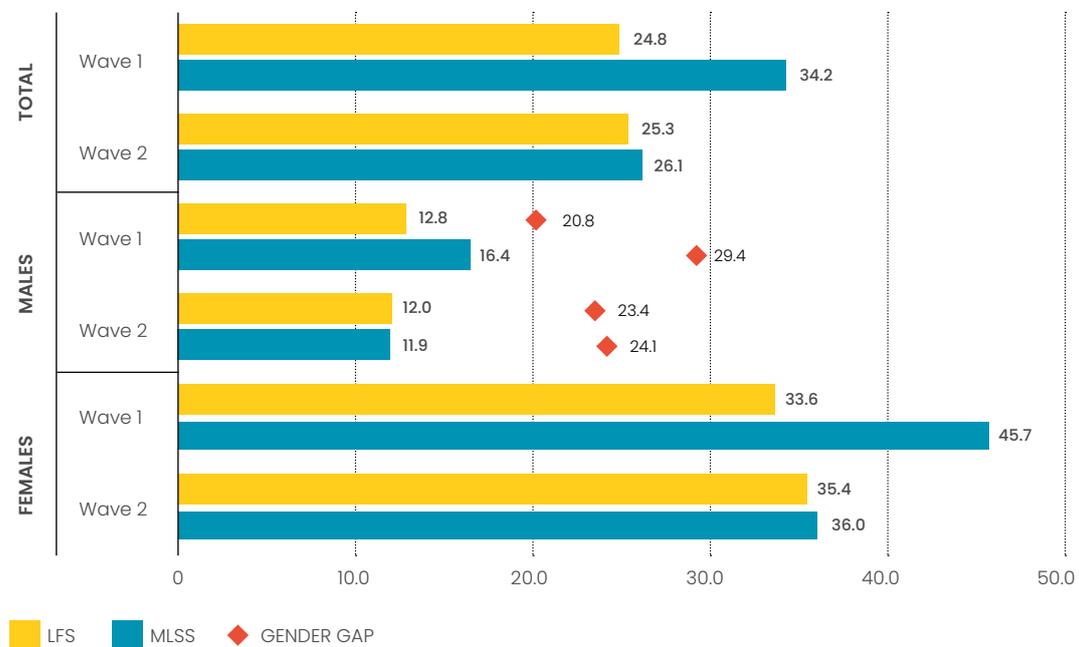
a descriptive text to be read aloud by LFS interviewers to respondents). Conversely, there was no explicit emphasis on active caregiving in the MLSS. As a consequence, the MLSS estimate for caregiving activities in wave 1 was nearly three times the LFS estimate (43.8 versus 16.1). During the wave 2 training, both sets of interviewers were instructed to read the additional text; the impact on the results was clear. The LFS result in wave 2 was relatively consistent with the results in wave 1, while the MLSS estimate fell by half, leaving a much smaller gap in the overall estimate of the time spent in the own-use provision of services in wave 2 (26.1 hours per week in the MLSS and 25.3 hours in the LFS).

The study also shows that the measured weekly hours spent on the own-use provision of services are significantly lower if the survey relies on only one question (seeking information on the number



Figure 3

Average hours actually worked in the own-use provision of services, by sex, wave of data collection and survey



Source: Joint DCS, ILO, and World Bank pilot study in Sri Lanka, Wave 1 and Wave 2, March–October 2019.

Note: Averages were calculated only for those respondents who reported that they had carried out some own-use provision of services during the reference period. The red diamond indicates the gender gap in working time in the activities covered. The diamond is included on the bar of the gender with lower working time. If it is included on the bar for women, it shows the amount by which the average working time of women in the activity was less relative to men and vice versa if it is shown on the bar for men.

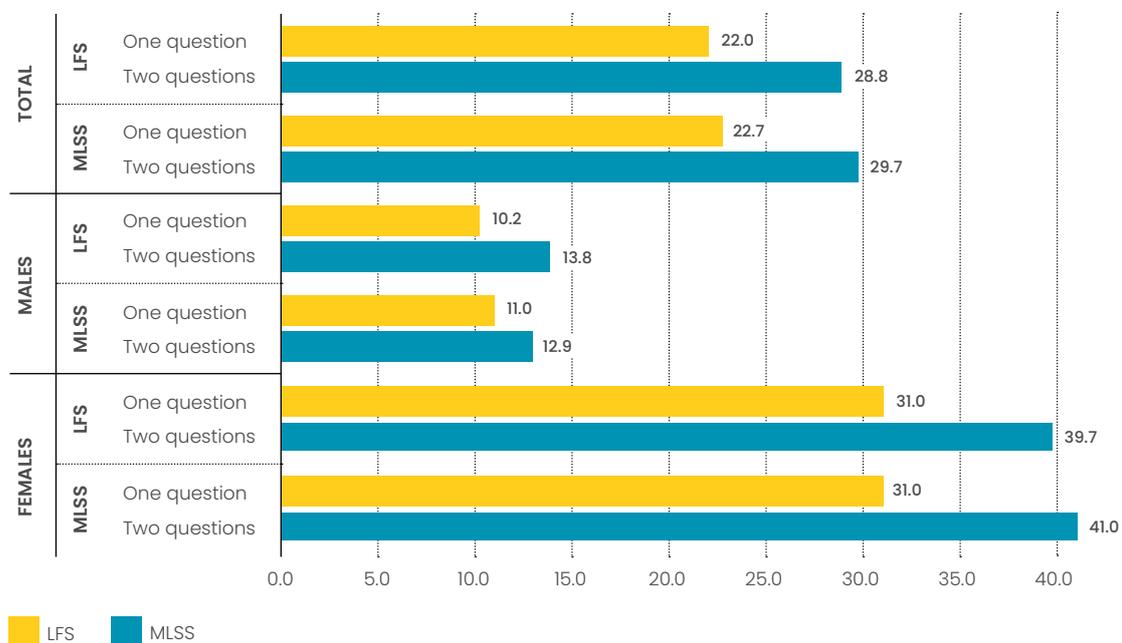
of hours worked during the previous week) rather than two questions (seeking information on the number of days worked during the previous week and the average number of hours per day). In wave 2, both the LFS and MLSS surveys administered half the samples the one-question approach and the other half the two-question approach. The results of both surveys were highly consistent. The two-question approach yielded a number of weekly hours spent on the own-use production of services that was approximately 30 per cent higher than the number of weekly hours captured by the one-question approach (see **Figure 4**). This pattern was repeated among both men and women albeit with slightly different gaps. A possible explanation for this last outcome is that the rounding of daily averages in using

the two-question approach leads to a relative overestimation relative to the one-question approach. However, while the direction and scale of the impact are quite consistent, which of the two sets of results is more valid is not certain.

The study covered many other issues, the analysis of which enhances our understanding of good practices in the measurement of work, employment and labour underutilization, as framed by the 19th ICLS standards. A variety of additional findings is highlighted in the main report. Perhaps a general summary is that, as highlighted above, the measurement of work may be sensitive to questionnaire design and implementation, and the study has allowed us to identify where misclassification risks appear greatest.

Figure 4

**Average hours actually worked in the reference week by own-use providers of services, by sex, survey and type of questions used to capture working time**



Source: Joint DCS, ILO, and World Bank pilot study in Sri Lanka, Wave 1 and Wave 2, March–October 2019.

Note: Each of the two survey samples was divided into two random groups. The questions about number of hours worked were asked using only one question to one of the groups (that is, “How many hours did you spend doing this last week?”) and with two questions among the other group (that is, “Last week, on how many days did you do this work?” and “And, on average, how many hours per day did (you/NAME) spend doing this last week?”).

## 3

## Other general points of note

The findings of the studies have been or will be incorporated into tools and guidance provided by the [ILO for the LFS](#) and by the [World Bank for the MLSS](#). The various tools available are a useful reference point for anyone engaged in the design of a household survey that seeks to measure work and labour market engagement in line with the 19th ICLS standards. However, in deciding on the most appropriate approach to adopt in any given household survey, various issues will need to be considered such as the following:

- What is the objective of the measurement of work and labour market engagement within the survey? For example, is the motivation a desire to generate estimates of the prevalence of participation and time spent in work, or is it the generation of a variable that can be used as an explanatory variable in the analysis for poverty or other issues? Another consideration is the desired range of the indicators on work (for example, the characteristics of jobs, the types of work to be covered) and the levels of disaggregation targeted. Dedicated LFSs are typically designed to achieve a wide coverage of a range of work-related indicators and often rely on larger sample sizes than other household surveys, thereby enabling greater precision in the labour market indicators and allowing more extensive disaggregation (within the limits of the sample size and design). By contrast, other household surveys with a less central focus on labour, but nonetheless seeking to cover selected labour-related issues, typically dedicate fewer questions to the topic and often (though not always) rely on smaller samples that reflect these objectives and design choices. Surveys with substantially different overall measurement objectives and designs should not be expected to generate comparable estimates of key labour market- and work-related concepts. However, the experience of the Sri Lanka study shows that careful design can improve the quality of the estimates generated by distinct types of household surveys and reduce the level of the differences observed.
- What is the socio-economic context of the country? The tools provided by the ILO include multiple approaches that can be adopted to a country's context. For instance, differences in the approach to the measurement of labour-related indicators may be justified in countries with a high prevalence of subsistence farming versus countries with smaller agricultural sectors. In settings characterized by substantial subsistence farming, the coverage of both employment and the own-use production of goods in each household survey with a focus on labour will be critical, while this may be less important in countries with less prevalence of these activities. However, even these countries may wish to measure such indicators periodically. Such differences in context are reflected in the various versions of the tools provided by the ILO for the [LFS](#) as well as the short sequences proposed for the [population census](#). Similar considerations are relevant in other types of household surveys seeking to cover work and labour market engagement. For instance,



MLSSs conducted in settings with substantial subsistence agriculture need to ensure that the own-use production of agricultural goods is covered along with employment by the survey instrument. Otherwise, such surveys risk underestimating the importance of agriculture in rural livelihoods, undercounting women's work in agriculture and potentially leading to biases in the estimation of sectoral labour productivity. These considerations have informed the new [Living Standards Measurement Study \(LSMS\) Guidebook for measuring labor in MLSS-type surveys](#).

A general point is the critical need for good questionnaire development and testing practices to put surveys on a sound footing. This is true

both at the international level in the activities of international agencies and at the national level among national statistical compilers. In the absence of appropriate testing, the degree of sensitivity of measurements may never become apparent. This leaves open the possibility that the statistics generated may not be representative of the true situation among the underlying population, for example, by obscuring important differences between women and men's working lives. Activities at the international level can provide a major support to countries, but not entirely replace the need for the effective translation and adaptation of questionnaires to national contexts, a process that needs to be supported by qualitative and quantitative testing at the national level to the extent possible.

# Summary Conclusions

- The first key conclusion is that the field experiment undertaken in Sri Lanka has generated a wealth of rich data that may be used to identify good practices in questionnaire design and apply the latest standards in the domain of work statistics. The findings of the study are enabling existing guidance on good practices in the measurement of women and men's work to be extended to surveys other than the LFS, including MLSS-type surveys. The depth and breadth of the conclusions generated by the study could not have realistically been generated through another mechanism.
- From a gender perspective, it is difficult to overstate the value of the study. As highlighted in the report on the findings, a much larger part of women's work relative to men's work tends to be invisible, at risk of underreporting, or not measured at all in official statistics. The value of the data generated as the new standards are applied is demonstrated in recently published [reports](#). It is clear that pilot studies such as this one are crucial to improving the measurement of both paid and unpaid work, thereby playing a key role in ensuring that the potential of the standards is achieved. These enhancements have been the outcome of a long, ongoing process, which gained significant momentum with the adoption of the 19th ICLS standards. This has been the main focus of the ILO and World Bank agenda to operationalize these standards and refine survey methods in the measurement of work and labour market engagement under the Women's Work and Employment Partnership.
- The Sri Lanka study highlighted important cross-survey differences in the measurement of key labour-related variables, particularly participation in various forms of paid and unpaid work. The impact of the changes undertaken in the MLSS before wave 2 of data collection suggests that some of the differences can be reduced, if not removed entirely, through relatively minor changes in questionnaire content or survey implementation.
- A variety of other sensitivities could be identified, such as the sensitivity of the measurement of working time in unpaid work to the measurement approach (for example, the use of the one-question method or the two-question method or other differences). This was especially evident in unpaid care work. In addition, all surveys should emphasize good translation and national adaptation, as well as interviewer training and supervision to promote consistency in measurement.
- The risks of misclassification or measurement difficulties were concentrated among people engaged in certain types of activities. For example, higher risks of misclassification or undercounting were seen in the case of people engaged in casual, low-hours jobs or people helping on family farms or in businesses. This is highly relevant in gender studies because these types of activities were more common among women respondents in the pilot study, which is also likely to be true in many other settings.

- The questionnaires were successful in capturing a range of paid and unpaid working activities. This unlocks a wide range of analytical potential, such as in obtaining a deeper understanding of gender gaps in working activities and labour market engagement. An important future objective arising from this study and related guidance on measurement practices is the promotion of the mainstreaming of the measurement of unpaid working activities, alongside labour market engagement, to enable this type of analysis on a regular, wide-scale basis.
- The harmonization of questionnaire content may be a way to improve consistency in measurement, but it cannot be assumed that absolute consistency between the LFS and other household surveys can be achieved or that the need for a national process of adaptation and testing can be ignored. The

differences in surveys and across countries mean that questionnaires should be adapted to context and fully tested to improve the quality of the statistics generated. A starting point of any survey process should be clear discussion and clarification of the key objectives of the survey. This can form the basis of the choices that need to be made about questionnaire content, matching the survey objectives and country context.

- The findings presented in this report are a subset of the many findings possible from the studies. The findings are being used to update the guidance, tools and support of the ILO for LFSs and the World Bank for MLSSs. The model questionnaires, guidance and tools are excellent reference points for those facing the task of designing a questionnaire to capture work- and labour-related issues through a household survey in line with the latest standards.

