

Jointly produced by
CEDEFOP and
EUROFOUND

Labour market change
European Company Survey 2019:
Quality control report

[European Company Survey 2019: Workplace practices unlocking employee potential](#)



CEDEFOP

European Centre
for the Development
of Vocational Training



Eurofound

European Foundation
for the Improvement
of Living and Working
Conditions

Author: Ipsos

Research managers: Giovanni Russo (Cedefop) and Gijs van Houten (Eurofound)

Acknowledgements: This project could not have been completed without the invaluable contribution of Sophia MacGoris (Eurofound) and Christopher White (Eurofound) to the preparation and implementation of the European Company Survey 2019, which included the checking and editing of this report.

Eurofound reference number: WPEF20015

Related reports: European Company Survey 2019: Technical and fieldwork report; European Company Survey 2019: Translation report; European Company Survey 2019: Sampling and weighting report; European Company Survey 2019: Coding report; European Company Survey 2019: Data editing report; European Company Survey 2019: Data quality assessment

© European Centre for the Development of Vocational Training (Cedefop), 2020/European Foundation for the Improvement of Living and Working Conditions (Eurofound), 2020
Reproduction is authorised provided the source is acknowledged.

For any use or reproduction of third-party photos or other material permission must be sought directly from the copyright holders.

Any queries on copyright must be addressed in writing to: copyright@eurofound.europa.eu

Research carried out prior to the UK's withdrawal from the European Union on 31 January 2020, and published subsequently, may include data relating to the 28 EU Member States. Following this date, research only takes into account the 27 EU Member States (EU28 minus the UK), unless specified otherwise.

This report presents the results of research conducted largely prior to the outbreak of COVID-19 in Europe in February 2020. For this reason, the results do not fully take account of the outbreak.

The European Centre for the Development of Vocational Training (Cedefop) is the European Union's reference centre for vocational education and training. It provides information on and analyses of vocational education and training systems, policies, research and practice. Cedefop was established in 1975 by Council Regulation (EEC) No 337/75.

The European Foundation for the Improvement of Living and Working Conditions (Eurofound) is a tripartite European Union Agency, whose role is to provide knowledge in the area of social, employment and work-related policies. Eurofound was established in 1975 by Council Regulation (EEC) No. 1365/75 to contribute to the planning and design of better living and working conditions in Europe.

European Centre for the Development of Vocational Training

Telephone: +30 2310490111

Email: info@cedefop.europa.eu

Web: www.cedefop.europa.eu

European Foundation for the Improvement of Living and Working Conditions

Telephone: (+353 1) 204 31 00

Email: information@eurofound.europa.eu

Web: www.eurofound.europa.eu

Contents

Introduction	1
1 Sampling	2
1.1 Net sample size	3
1.2 Sampling strategy.....	7
1.3 Sampling plan	7
1.4 Sampling frames	8
1.5 Reference statistics	13
2 Weighting.....	16
2.1 Weighting strategy.....	16
2.2 Weighting implementation.....	17
3 Questionnaire.....	21
3.1 Questionnaire development.....	21
3.2 Advance translation	23
3.3 Cognitive testing	24
3.4 Translation.....	26
3.5 Pilot.....	30
4 Fieldwork	33
4.1 Fieldwork infrastructure.....	33
4.2 Scripting and script checks (consistency checks)	34
4.3 Interviewer training and monitoring.....	36
4.4 Contact strategy.....	38
5 Fieldwork monitoring	41
5.1 Interviewer monitoring	41
5.2 Fieldwork monitoring.....	45
5.3 Data validation.....	46
6 Recoding	48
6.1 Recoding strategy.....	48
6.2 Recoding quality control.....	48
7 Data and reporting	49
7.1 Micro dataset.....	49
7.2 Reporting.....	51
8 Punctuality	51
9 Conclusions	56
9.1 Reflections on quality control results in the 2019 ECS.....	56
9.2 Reflections on the quality control approach.....	61

Introduction

This report describes in detail the approach to **quality assurance and control** for the European Company Survey (ECS) 2019, a survey of establishments in the EU28 conducted by Ipsos on behalf of the European Foundation for the Improvement of Living and Working Conditions (Eurofound) and the European Centre for the Development of Vocational Training (Cedefop).

In accordance with the Terms of Reference, the quality assurance approach for the ECS 2019 was based on the quality concepts of the European Statistical System (ESS), as developed by Eurostat¹, in addition to other quality frameworks such as the Cross-Cultural Survey Guidelines² and the Total Survey Error Approach.³ The **quality assurance plan** for the ECS 2019 was built around the five key quality criteria defined by the ESS but slightly redefined to make them more appropriate to the ECS: Relevance and Timeliness, Accuracy, Accessibility, Coherence and Comparability, and Punctuality. The quality criteria are defined in Table 1.1.

Table 1.1: Eurofound's quality criteria

Dimension	Description
Relevance & Timeliness	Relevance for users of the survey data and survey-based reports, both in terms of substance and timing of publication
Accuracy	Validity and reliability of the survey data
Accessibility	Availability of outputs and transparency of processes
Coherence & Comparability	Consistency with other data sets
Punctuality	Adherence to timeline as set at the start of the project

Although all five criteria have some bearing on all stages of the survey, quality monitoring in each of the stages is focussed on those criteria that require attention at that stage of the survey cycle.

These quality criteria formed the basis of the **quality assurance framework**, which included quality indicators and targets for each quality dimension and (sub-) theme of the survey, specified the evidence to be provided, and included an indication of the role of each party (such as the local partner in different countries, the Ipsos coordination team, or Eurofound and Cedefop).

The quality assurance plan was initially developed by Eurofound, all subsequent changes to the quality assurance plan during the project were made by the designated Delivery Manager at Ipsos, in close cooperation and agreement with both Eurofound and Cedefop. This included adjusting indicators to reflect changes in the methodology. The quality assurance plan was a live document, which was monitored throughout the project and updated monthly by the Ipsos Delivery Manager.

¹ http://ec.europa.eu/eurostat/en/web/products-manuals-and-guidelines/-/QAF_2012

² <http://ccsg.isr.umich.edu/index.php/chapters/survey-quality-chapter>

³ Herbert F. Weisberg (2005). The Total Error Approach. A Guide to the New Science of Survey Research. Chicago: Chicago University Press.

The updated versions of the quality assurance plan were shared with Eurofound and Cedefop and potential deviations from targets were discussed and corrective action taken where possible.

This report documents how the quality assurance plan was implemented and whether or not each of the 131 quality assurance indicators were met. The report is structured around the ECS 2019's cycle, with thematic chapters for each stage of the project, divided into sub-sections based on key sub-themes.

Each section follows the same structure:

1. Short overview describing the specific aspect of the survey process
2. Tables for each sub-section showing the relevant quality assurance targets and indicators along with the rationale underlying the indicators as well as the result specifying whether the target was met or not.

Colour coding is used to denote whether the target was a **'requirement'** that had to be achieved (red shading) or a **'target'** that it was expected could be achieved (yellow shading). Similarly, colour coding for the result indicates whether the target was achieved (green font) or not (red font).

3. Discussion of each table – including a brief account of what was done to ensure the targets were met, what happened when targets were not met, what mitigating actions were taken when it became apparent targets might not be met, and ultimately, an assessment of the severity of the quality implications in case targets were not met

The ECS 2019 consisted of two phases, the first phase focused on preparation including questionnaire design, translation and (pilot) sampling whilst the second phase focused on implementation focusing on mainstage sampling, fieldwork and reporting.

Some quality indicators are applicable only to the preparation phases; others to both the preparation and implementation phases and finally some only to implementation. Initially, 32 countries were included in the preparation contract (EU28 plus four accession countries – Serbia, Montenegro, North Macedonia and Turkey) but for implementation, only the EU28 were included. Consequently, for some indicators the base for 'all countries' is 32, whereas for others it is 28. It will be indicated what the base is where necessary throughout this report.

The conclusions section of this report includes a brief reflection on the results of quality control for the ECS 2019 as well as reflections on the overall approach itself - its effectiveness, efficiency and appropriateness, highlighting shortcomings and potential improvements for future waves.

1 Sampling

The quality indicators and targets relating to sampling reflect the different stages in drawing a sample for a random probability survey of businesses as used for the ECS 2019. The first set of indicators relate to net sample size, followed by indicators and targets for the overall sampling

strategy, sampling plan, sampling frames and reference statistics. Further details of the sampling approach taken for the ECS 2019 can be found in the Sampling and weighting report.⁴

1.1 Net sample size

Requirements

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accuracy	1.29	Percentage of countries where the net sample size >= planned sample size.	Sample size targets are set to enable desired analysis. Targets need to be met to ensure this.	100%	57%	Contract(s) (planned sample); Final dataset (net sample)
Accessibility	1.28	Percentage of countries for which all stratification variables and distributions of universe statistics are made available in interim and final datasets.	Including the universe statistics on the size of the stratifications cells allows for applying finite population correction when carrying out analysis. This should be possible for all countries.	100%	100%	Interim and final datasets

Indicators 1.29 and 1.28 are requirements that are both related to the net sample size. It was only possible to meet the planned sample size in 17 out of 28 countries (57%). The net sample size was smaller than planned in 12 countries – Cyprus, Czechia, France, Germany, Italy, Luxembourg, Malta, Poland, Portugal, Romania, Spain and the United Kingdom as shown in Table 1.2. This table provides the planned sample size for the management interviews as well as the achieved sample sizes for the management and employee representative interviews. Interviews that were dropped due to quality issues have been excluded from the count of those achieved.

The reasons why 12 countries were unable to meet the planned sample size vary across countries but some of the recurring issues were:

- smaller than anticipated gross sample available from the frame
- higher non-contact rate during fieldwork
- lower CAWI conversion rate in mainstage compared to the pilot (in part attributed to higher refusal rates and lower co-operation rates than expected)
- very low CATI yield compared to anticipated (in part attributed to higher refusal rates and lower co-operation rates than expected)
- MM online interviews flagged for low quality post-fieldwork

⁴ Eurofound and Cedefop (2020), *European Company Survey 2019: Sampling and weighting report*, European Company Survey 2019 series, Eurofound working paper, Dublin (<https://www.eurofound.europa.eu/publications/report/2020/european-company-survey-2019-workplace-practices-unlocking-employee-potential#wp-101981>).

Table 1.2: Planned and achieved sample sizes – MM and ER online interviews

Country	MM online interviews		ER online interviews
	Planned N	Achieved N	Achieved N
Austria	1000	1010	163
Belgium	1000	1011	100
Bulgaria	1000	1024	100
Croatia	500	560	71
Cyprus	250	122	3
Czechia	1000	904	46
Denmark	1000	1011	134
Estonia	500	501	17
Finland	350	1032	467
France	1500	1360	425
Germany	1000	711	66
Greece	500	501	8
Hungary	1000	1087	34
Ireland	250	300	6
Italy	1500	1498	188
Latvia	500	514	11
Lithuania	500	510	131
Luxembourg	250	237	39
Malta	250	145	5
Netherlands	1000	1030	339
Poland	1150	842	48
Portugal	1000	973	17
Romania	1000	815	76
Slovakia	350	361	27
Slovenia	500	556	23
Spain	1500	1477	196
Sweden	1000	1080	307
United Kingdom	700	697	26

Several solutions were applied to try to resolve the issues detected, again these varied across countries depending on the nature of the problem identified. For example:

- find alternative contact telephone numbers
- carry out extra phone look-ups
- introduce extra CATI reminders
- change the CATI reminder strategy
- introduce a small incentive
- order and use additional sample
- extend the fieldwork period

The issues experienced and the solutions applied are detailed in the qualitative notes prepared for Eurofound/Cedefop by Ipsos. The main impact of not achieving or exceeding the planned sample size is the loss of precision of estimates based on the achieved sample. It should be noted that in some countries, it was decided to focus efforts on achieving smaller increases in sample size across all size/sector categories, rather than on trying to achieve the target sample size by releasing fresh sample which would no longer include entries for some categories. To aid decision-making in this regard, Ipsos ensured that all stratification variables and distributions of universe statistics were made available in both the interim and final datasets for all countries included in the mainstage fieldwork.

Targets

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accuracy	1.27	Percentage of countries where the distributions across stratification categories of the net sample closely approximates the distributions of the universe (sampling plan) (deviations in the proportional size of each of the strata between the two should not exceed 5 percentage points).	To be considered representative (or in the case of the ECS, to strike the right balance between representativeness on the level of establishments and on the level of employees) the sample distribution should closely correspond to the universe distribution.	100%	54%	Sampling and weighting report

The distributions across stratification categories of the **net sample** closely approximates the distributions of the universe (sampling plan; the proportional size of any of the strata in the achieved sample did not differ from the planned sample by more than 5 percentage point) in 15 out of 28 countries (54%). The difference between the achieved sample and the planned sample did exceed 5% in some strata in 13 countries (CY, CZ, DE, EE, ES, HU, IE, LU, MT, NL, PL, SI and SK). These differences ranged between 5.1% and 13%. The highest number of deviations can be observed for the stratification cell ‘Small & NACE G-N, R, S’, which created problems in seven countries. See Table 1.3 for further details. The implication for data quality of not meeting indicator 1.27 is that the unweighted distributions do not match the sampling targets, implying that (more extensive) weighting is required.

Table 1.3: Countries where the distributions across stratification categories exceeded 5 percentage points in final management survey data

Deviations by stratification cell					
Country	Small & NACE B-E	Small & NACE G-N, R, S	Medium & NACE B-E	Medium & NACE G-N, R, S	Large & NACE G-N, R, S
Cyprus		-7.2		13.0	
Czechia			-5.7		
Estonia	-5.8				
Germany		-5.5		5.4	
Hungary	-8.3				
Ireland					-8.4
Luxembourg		10.6		-6.5	
Malta	-12.2		7.8		
Netherlands		-6.8		5.1	
Poland		-6.2			
Slovakia	5.3	-9.7	5.2		
Slovenia	-6.7				
Spain		-6.0			

In Slovenia, a separate problem occurred that is related to indicator/target 1.27 but not directly measured by it. In short, the provider in Slovenia omitted cases for service sectors NACE K-S when it delivered the sample to Ipsos. Upon receipt of the sample, Ipsos checked that the total number of cases sampled from the services stratum was correct but not the distribution across the NACE sectors within the services sector (the implicit stratification levels). This omission was not picked up during fieldwork monitoring either as reporting focused on the stratum level not on categories within stratum cells. The problem was detected during the weighting process, by which time it was too late to rectify the problem. The omission constitutes 19% of the target population in Slovenia, meaning that the overall Slovenian sample is biased as is the Slovenian services sector sample. Although the omission of K-S in Slovenia does also affect the EU estimates, calculations have shown that - due to the size of the country - this bias will not exceed 0.1 percentage point. This issue and its impact are further described in the Sampling and weighting report.

1.2 Sampling strategy

Requirements

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accuracy	1.1	A common sampling strategy is agreed that reflects current academic standards, as illustrated by academic references included in the sampling strategy document.	Adherence to current academic standards ensures that the quality assurance (e.g. peer review) underlying these standards also applies to the ECS 2019.	YES	YES	Sampling strategy

The sampling strategy was developed by the Ipsos Sampling Manager incorporating the quality elements described above. The sampling strategy was agreed in terms of approach with Eurofound prior to the Pilot with a note indicating that the content could be adjusted further as additional information became available or as specific aspects were discussed as the project progressed (indicator 1.1). The sampling strategy was reviewed, and key sections adjusted prior to mainstage implementation, reflecting lessons learnt from the Pilot.

1.3 Sampling plan

Requirements

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accessibility	1.21	Percentage of countries for which distributions of the sample and the population across stratification categories are provided.	Requiring provision of this data ensures that it is collected for and used in the sampling design for the ECS 2019.	100%	100%	Country-level sampling plans

The Ipsos Sampling Manager worked with the local partner in each country to develop a country-specific sampling plan. This included distributions of the sample and the population across stratification categories as required by indicator 1.21. Each of the 32 countries developed a sampling plan and each plan contained the required information, meaning that the information could be used in the sample design for the ECS 2019 and the target of 100% could be met.

Targets

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accuracy	1.20	Percentage of countries where the agreed variables are used for stratification.	Using the same variables – where possible with the same categories – for sampling and weighting enhances cross-national comparability.	100%	88%	Country-level sampling plans

Country-specific sampling plans were developed and agreed for each of the countries participating in the ECS 2019. These included variables relating to size categories for businesses as well as the level of regional data to be used for stratification. In the end, the stratification for the ECS 2019 was based on different size categories or regional data than originally envisaged. In four countries (BE, DK, EL and SE) different size categories needed to be used. In three other countries (BG, MT and PL) cities/towns and post-codes were provided and in five further countries (CY, IE, IT, PT and SE) local administrative areas (LAUs) were provided. Therefore, the stratification for the ECS 2019 was based on different size categories or regions in some countries. In each case, the best available data that could be accessed for sampling was used. For the four countries where different size categories were used, the same categories were available in the reference statistics, meaning the sample design and targets could be set based on the available categories instead. In the case of the regional information using a different level of data for the final stratification variable, this is unlikely to make much difference to the sample given the stratification was proportional.

1.4 Sampling frames

Requirements

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accuracy	1.7	Percentage of countries where specified information on stratification variables (sector and size class) is included in the sampling frame.	To design the sample based on a common set of parameters, information on these parameters needs to be available on all the sampling frames.	100%	100%	Country-level sampling plans
Accessibility	1.11	Percentage of countries for which the characteristics of the sampling frame and procedure are documented in complete accordance with the agreed template (based on Terms of Reference).	Transparency relies on systematic documentation of all phases of the survey preparation and implementation – including the sample frame and sampling procedure.	100%	100%	Sampling strategy and Country-level sampling plans

Prior to selecting a suitable sampling frame for use in the ECS 2019, the Ipsos Sampling Manager and local teams worked together to identify the information available on the frame. Consideration was given to the availability of the variables required for stratification - sector and size class - as required

by quality indicator 1.7. The aim was to design the sample based on a common set of parameters. To do this, information on the parameters needed to be available on all the sampling frames, which was the case for all 32 countries. The Ipsos Sampling Manager documented the characteristics of the sampling frame and procedure for each country as required, meeting indicator 1.11 and ensuring that the process was transparent.

Targets

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accuracy	1.3	Percentage of countries where a sampling frame on the establishment level is used	It is desirable to have a sampling frame that contains the exact unit of analysis (establishment level), as this implies no further sampling steps	44% (pilot); 50% (mainstage)	34% (pilot); 39% (mainstage)	Sampling strategy

A target of 44% was agreed for indicator 1.3 for the Pilot based on Ipsos's proposal. This in turn was based on the availability of sample frames on the establishment level in 14 of the EU28 countries when the proposal was written in 2017⁵. It was desirable to have a sampling frame that contained the exact unit of analysis (establishment level), as this would have meant that no further sampling steps needed to be taken. However, at the time of sampling for the pilot ECS 2019, fewer countries than anticipated (11 out of 32, 34%) could access an establishment level frame of sufficient quality to meet Eurofound and Cedefop's requirements. For the mainstage, this result was 39% (11 out of 28 countries). This meant that additional sampling steps needed to be incorporated in the screener in the countries using company level frames to ensure that establishments could be reached. This added complexity and cost to the sampling process in some countries and generally lower quality samples given subsidiaries in multi-site establishments are under-represented when sampling from a company level frame.

Table 1.4 lists the three countries where a company level frame was used for the mainstage instead of an establishment level frame as offered in Ipsos' proposal. An explanation of why it was necessary to use this level of frame rather than an establishment level frame has also been included.

⁵ Establishment level frames were not available for any of the four non-EU countries.

Table 1.4: Company frame usage by country

Country	Explanation
Belgium	At the time Ipsos submitted the proposal for the survey it was believed that Graydon could be used as an establishment level sampling frame, however, although information about the establishments of multi-site companies could be provided, further investigations showed that their contact details could not be provided, and so the sampling was undertaken at the company level.
Poland	Dun & Bradstreet (provided by the Ipsos central team supplier) provided counts (at the company level) with better coverage than Bisnode so this frame was selected.
Spain	The sampling frame was changed during the project preparation phase, as just before the pilot the previous provider (DataCentric) reported that they had removed a third of their records due to GDPR, reducing coverage significantly. The new source proposed (Informa D&B) had slightly better coverage than DataCentric had prior to the loss of records, and so was selected. This sampling frame was available at the company level (DataCentric was also planned to be used at company level).

Quality dimension	#	Indicator	Rationale	Targ	Ach.	Evidence
Accuracy	1.4	Percentage of countries where the sampling frame covers at least 95% of the population.	Sampling frames should have as high as possible a coverage level, to avoid coverage errors (under-sampling or exclusion).	100%	50% (pilot); 46% mainstage	Country-level sampling plans

The proportion of countries where the sampling frame covered at least 95% of the population is 50%, which is lower than the target of 100% that was aimed for in indicator 1.4. However, most sampling frames (78% from the pilot, 79% from the mainstage) do cover at least 85% of the population. Whilst this does mean that the potential for coverage error is increased, this was the best result that could be achieved given that the most suitable, available sample frames were selected in all countries at the time.

In some countries, relatively many entries in the sample frames lacked information on the size of the company/establishment, these entries were grouped in a “no size” stratum. In seven countries in which the coverage was poor, a “no size” stratum was included in the sampling plan for the pilot to evaluate whether this would be a feasible approach for the main stage. These seven countries were those where coverage was below 80%, based on the sample frame numbers that do include establishment size. In one of these countries, Spain, the sampling frame was changed just before the

pilot to one offering higher coverage (97%), meaning that it was no longer necessary to include ‘no size’ cases in the main survey (as explained in Table 1.4). Another of these countries was not an EU Member State and was not included in the main survey. For two of the five remaining countries (AT and CY), coverage is estimated to have been above the 95% level in the main survey, given the eligibility rates in the ‘no size’ stratum.⁶ In the remaining three countries, Greece, Luxembourg and Malta, the final coverage was approximately 67%, 70% and 85% respectively including ‘no size’.

Quality dimension	#	Indicator	Rationale	Targ	Ach.	Evidence
Accuracy	1.5	Percentage of sampling frame units that refer to non-existent or non-eligible establishments and companies.	Sampling frames should be as accurate as possible to avoid coverage errors (if non-existent or non-eligible cases are not distributed randomly) and to improve fieldwork efficiency.	10%	6%	Sampling and weighting report (based on fieldwork outcome codes)

Indicator 1.5 is a real-world quality target relevant to the sub-topic sampling frames. A target of 10% was set for the proportion of sampling frame units that referred to non-existent or non-eligible establishments and companies. In the end, Ipsos achieved a result of 6%, 4% lower than the threshold set at 10%. The result of 6% is based on the sample dialled and refers to the proportion of cases across all countries that were non-existent or non-eligible establishments and companies. In four countries, more than 10% of cases were non-existent or ineligible – Austria, Greece, Luxembourg and Poland (19%, 17%, 18% and 13% respectively). The calculation considered the following outcomes as ineligible or non-existent: <10 employees; Discontinued business; Not a business and Ineligible activity.

Quality dimension	#	Indicator	Rationale	Targ	Ach.	Evidence
Accuracy	1.6	Percentage of countries where the sampling frame was updated within a year preceding fieldwork.	The more recently a sampling frame is updated the less likely they will suffer from coverage issues.	100%	88% (pilot); 100% (main)	Country-level sampling plans

For indicator 1.6, a target of 100% was set, Ipsos achieved a result of 88% for the pilot (28 out of 32 countries). Most sampling frames were updated within a year preceding the pilot fieldwork meaning that they were more likely to be accurate and less likely to suffer from coverage issues compared to those updated longer ago. If only the year was mentioned in the country-specific sampling plans,

⁶ For these Ipsos looked at the eligibility rate of the no size stratum based on the main stage and applied this to the total size of the no size stratum to estimate the total covered. The other strata were not adjusted, i.e. it was assumed that all other size categories are 100% eligible, as this is the basis of the coverage estimate in all other countries. This allows us to draw the conclusions here.

'2017' was considered to be within a year preceding the pilot / mainstage fieldwork. If the date mentioned was more specific e.g. 'July 2017', this was not considered to be within 1 year. This was the best result that could be achieved for the Pilot given that the most suitable, available sample frames were selected in 32 countries. The corresponding result for the main survey is 100% as in all 28 countries the sampling frame used was updated within a year preceding the mainstage fieldwork that was carried out in 2019. The sample frame used in Cyprus for the pilot was the most up to date version available however telephone numbers were not included meaning that matching was needed to collate this information for selected cases. For the mainstage, it was possible to access an up-to-date frame, which included telephone numbers.

Quality dimension	#	Indicator	Rationale	Targ	Ach.	Evidence
Coherence and comparability	1.8	Common categories are agreed for each of the variables on sector and size class that are to be used to specify the information in the sampling frames in all Countries.	Using the same variables – where possible with the same categories – for sampling and weighting enhances cross-national comparability.	YES	YES	Country-level sampling plans
Coherence and comparability	1.9	Percentage of countries where the specified information on the sampling frame on sector uses the agreed set of categories.		100%	100%	Country-level sampling plans
Coherence and comparability	1.10	Percentage of countries where the specified information on the sampling frame on size class uses the agreed set of categories.		100%	88%	Country-level sampling plans

Indicators 1.8, 1.9 and 1.10 all measure the quality dimension coherence and comparability. Two of the three indicators were met. For indicator 1.8, the target was met as common categories were agreed. Indicator 1.9 was met fully. Indicator 1.10 was not fully met, as different size categories were used in four countries (BE and SE used the size categories: 10-19, 20-99, 200+; DK and EL used the size categories: 10-19, 20-99, 100+; EL has the exact number of employees available but the reference statistics use the size categories 10-19; 20-99; 100+, therefore they used the same categorisation) which produced a result of 88%. The use of different size categories in these countries was unavoidable given no other categories were available from the data sources that were used. For consistency within these countries and to mitigate the potential impact on sample quality, the same categories were used in these countries at all survey stages – reference statistics used to set targets, sample selection, sampling targets and weighting. The use of different size categories in four countries may have led to a slight reduction in the cross-national comparability of the data for these countries against the other 24 countries.

1.5 Reference statistics

Requirements

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accuracy	1.15	Percentage of countries where specified information on stratification variables (sector and size class) is included in the reference statistics.	Using the same variables – where possible with the same categories – for sampling and weighting enhances cross-national comparability.	100%	100%	Country-level sampling plans
Accessibility	1.19	Percentage of countries for which the characteristics of the reference statistics are documented in complete accordance with the template.	Transparency relies on systematic documentation of all phases of the survey preparation and implementation.	100%	100%	Sampling strategy and Country-level sampling plans

Indicators 1.15 and 1.19 focus on the reference statistics used for stratification, in both cases a target of 100% was set and in both cases, this was met. The reference statistics chosen for each of the 32 countries included information on the stratification variables (for sector and size class) and the characteristics of the reference statistics were documented in the country-level sampling plans as well as the sampling strategy.

Targets

Indicator 1.12 measured the proportion of countries where reference statistics on the establishment level were used. The target was set at 44% based on Ipsos's proposal, which in turn was based on the availability of sample frames on the establishment level in 14 of the EU28 countries when the proposal was written in 2017⁷.

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accuracy	1.12	Percentage of countries where reference statistics on the establishment level are used.	It is desirable to use reference statistics that contain the exact unit of analysis, as this implies no further corrective steps need to be taken.	44%	31% (pilot); 36% (main)	Sampling strategy

At the time of sampling for the ECS 2019, fewer countries than anticipated (10 out of 32, 31% - pilot; 10 out of 28, 36% - mainstage) could access reference statistics at the establishment level. It was possible to access reference statistics at the establishment level in AT, DK, FI, FR, DE, IE, IT, NL, SE and the UK.

⁷ Establishment level reference statistics were not available for any of the four non-EU countries.

Generally, the countries with company level statistics were also those where the sampling frame was at the company level, where it made sense to set the sampling based on the company statistics profile. Establishment level statistics are required for weighting however this can be estimated based on the survey data.

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accuracy	1.13	Percentage of countries where the reference statistics fully cover the population.	Reference statistics should preferably not have any coverage issues, to avoid coverage errors in sampling design or weighting.	100%	87%	Country-level sampling plans

Indicator 1.13 measures the extent to which the reference statistics fully cover the population of interest in all 32 countries. Ideally the statistics would not have any coverage issues to avoid generating coverage errors in either the sampling design or weighting phases. A target of 100% was set for this indicator. Unfortunately, the percentage of countries where the reference statistics fully covered the population was only 87%. In four countries coverage was lower than 100% - HR and SK (90% coverage respectively); DE (95%+ coverage) and EL (information not available as the last update of the statistics was in 2015 and the country was unable to provide an estimate for 2019). This result means that for these four countries the sampling targets and weighting might be slightly less accurate. In CY, CZ, PL, RO and UK “almost 100% coverage” was mentioned in the country-level sampling plans, therefore this was regarded as full coverage of the population. The reference statistics in LU do not cover non-profit institutions serving households but this was still counted positively towards the indicator.

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accuracy	1.14	Percentage of countries where the reference statistics used for stratification were updated within a year preceding fieldwork.	The more recently reference statistics are updated the less likely that they suffer from coverage issues.	100%	Pilot: 87%. Main survey: 93%.	Country-level sampling plans

Indicator 1.14 relates to how recently the reference statistics were updated prior to fieldwork. Again, a target of 100% was set – meaning that in all countries, the reference statistics used for stratification should have been updated within a year preceding fieldwork. This would ensure that they were more likely to be accurate and less likely to suffer from coverage issues.

Ipsos reviewed the final sampling templates sent to Eurofound/Cedefop before the mainstage (‘Final templates’, 11 February 2019), and those approved before the pilot (‘Approved’, 16/07/2018). For the Pilot, frames that were updated in 2017 were considered to meet the requirement if no month information was specified. Ipsos achieved a result of 87% for the pilot (28 out of 32 countries), based

on point 2.5 in the sampling templates. The target was missed in four countries – Croatia (data from 2015), Cyprus and Serbia (data from 2016) and Romania (date from April 2017).

For the mainstage, stratification was completed sometime between January-February 2019. A most recent update in '2017' was considered as **not** 'within the last year' for sampling stratification. So, 93% of the countries met the requirement based on point 2.6 in the sampling templates. The target was missed in two countries – Cyprus and Latvia (data from 2017).

Whilst the results from the pilot and the mainstage were both below target, the best available statistics were used for stratification purposes in each country each time. It was not possible to improve on this proportion. The impact on quality is expected to be minimal given that the population distributions change slowly over time.

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Coherence & comparability	1.16	Common categories are agreed for each of the variables on sector and size class that are to be used to specify the information in the reference statistics for all countries.	Using the same variables – where possible with the same categories – for sampling and weighting enhances cross-national comparability.	YES	YES	Country-level sampling plans
Coherence & comparability	1.17	Percentage of countries where the specified information in the reference statistics on sector uses the agreed set of categories.		100%	100%	Country-level sampling plans

Indicators 1.16 and 1.17 both measure the quality dimension coherence and comparability, and both were met. For indicator 1.17, it was considered that a result of 100% was achieved, as the same categories were used in all countries. However, in HR the Structural Business Statistics (the source of the reference statistics) did not cover the R and S sectors. The targets for these sectors for HR were estimated using data from countries with a similar profile.

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Coherence & comparability	1.18	Percentage of countries where the specified information in the reference statistics on size class uses the agreed set of categories.	Using the same variables – where possible with the same categories – for sampling and weighting enhances cross-national comparability.	100%	86%	Country-level sampling plans

Indicator 1.18 measures the quality dimension coherence and comparability and relates to the categories of size class used in the reference statistics for sampling. The reference statistics in BE and SE use size categories: 10-49, 50-99, 100-199, 200-499, 500+ and the reference statistics in DK and EL use size categories: 10-19, 20-49, 50-99, 100+. Therefore, the target was not met and the ability to compare across countries may have been compromised. However, as noted earlier, the use of different categories in these countries was considered beneficial as it ensured that, within these countries, the same categories were used for sampling and weighting. The base for the calculation are the 32 countries included in the preparation phase for the ECS 2019.

2 Weighting

The quality indicators and targets relating to weighting reflect the process developed and implemented for the ECS 2019. The first set of indicators relate to the weighting strategy and the second to implementation. Within this section, sub-sections relate to design weights, post-stratification weights, supra-national weights and the size of weights. Further details of the approach taken for weighting the ECS 2019 data can be found in the Sampling and weighting report.⁸

2.1 Weighting strategy

Requirements

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accuracy	2.1	A common weighting strategy is agreed that reflects current academic standards, as illustrated by academic references included in the weighting strategy document, and integrates all available information on those elements that are foreseen to be included in the weighting procedure, given the sampling plan.	1: adherence to current academic standards ensures that the quality assurance (e.g. peer review) underlying these standards also applies to this project. 2: Transparency relies on systematic documentation of all phases of the survey preparation and implementation.	YES	YES	Weighting strategy

The weighting strategy was developed by the Ipsos Sampling Manager incorporating the quality elements described in indicator 2.1. The strategy was agreed with Eurofound and Cedefop.

⁸ Eurofound and Cedefop (2020), *European Company Survey 2019: Sampling and weighting report*, European Company Survey 2019 series, Eurofound working paper, Dublin (<https://www.eurofound.europa.eu/publications/report/2020/european-company-survey-2019-workplace-practices-unlocking-employee-potential#wp-101981>).

Targets

The weighting strategy was applied to the data for all EU28 countries participating in the ECS 2019, enhancing comparability. The precise application of the strategy in each country is documented in the Sampling and weighting report.

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accuracy	2.3	Percentage of countries where the agreed weighting strategy is used.	Using an equivalent approach for weighting in all countries enhances comparability. The agreed weighting approach will attempt to achieve this equivalence and should therefore be applied in all countries.	100%	100%	Sampling and weighting report

The main point to be emphasised is that the agreed weighting approach aimed to harmonise as much as possible but includes differences necessary to deal with the different availability of sampling frames (establishment vs company) and of additional reference statistics.

2.2 Weighting implementation

Design Weights

Requirements

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accuracy	2.4	Percentage of countries where the design weight is specified in accordance with the sampling design	The design weight should be correct for inequalities in selection probabilities based on the survey design.	100%	100%	Sampling and weighting report
Accessibility	2.5	Design weights included in dataset	Including all the separate weighting components in the dataset allows users to decide to what extent they want to follow the same approach. Also, it improves the transparency of the approach followed by Ipsos/EF/CF and increases future flexibility (reweighting could be limited to specific weighting steps and implications can be assessed at each step).	YES	YES	Final dataset

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accessibility	2.6	Procedure for constructing design weights outlined in weighting report	Transparency relies on systematic documentation of all phases of the survey preparation and implementation and enables data to be reweighted if necessary.	YES	YES	Sampling and weighting report

The percentage of countries where the design weight was specified in accordance with the sampling design was 100% (indicator 2.4). So, in all countries, the design weight can be applied to correct for inequalities in selection probabilities based on the survey design. Ipsos included the design weights in the final dataset (indicator 2.5) and outlined the procedure for constructing these weights in the final Sampling and weighting report (indicator 2.6). By including the design weight (and other weights) in the dataset, users can decide whether to apply it in their analyses. The inclusion and accompanying documentation also improve transparency and accessibility and allows for the possibility of reweighting the data for limited steps if desired.

Post stratification weight

Requirements

Quality indicators 2.8, 2.10 and 2.11 all relate to the post-stratification weights. The post-stratification weight takes all agreed variables into account for 100% of the countries in the mainstage ECS (indicator 2.8).

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accuracy	2.8	Percentage of countries where the post-stratification weight takes all agreed variables into account.	Using the same variables – where possible with the same categories – for sampling and weighting enhances cross-national comparability and transparency.	100%	100%	Sampling and weighting report
Accessibility	2.10	Post-stratification weights included in dataset.	Including all the separate weighting components in the dataset allows users to decide to what extent they want to follow the same approach. Also, it improves the transparency of the approach followed by Ipsos/EF/CF, and increases future flexibility (reweighting could be limited to specific weighting steps and implications can be assessed at each step).	YES	YES	Final dataset

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accessibility	2.11	Procedure for constructing post-stratification weights outlined in weighting report.	Transparency relies on systematic documentation of all phases of the survey preparation and implementation.	YES	YES	Sampling and weighting report

More specifically, sector and size are both used for all countries at the calibration stage of weighting. Indicator 2.10 was also met. The post-stratification (calibration) weights are the final weights, and these are included in the datasets by default. The calibration was not a separate step as it included the weights from the previous weighting steps as input. The Sampling and weighting report documents the procedure for constructing the post-stratification weights (indicator 2.11) ensuring transparency of the process via systematic documentation.

Targets

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accuracy	2.9	Percentage of countries where the same agreed set of variables with the agreed categories are used for weighting.	Using the same variables – where possible with the same categories – for sampling and weighting enhances cross-national comparability.	100%	100%	Sampling and weighting report

The same set of variables with the agreed categories were used for the post-stratification weights for all countries so the result is 100%. Some small tweaks made to the categories for region (NUTS2/NUTS1) and the top-level NACE when the cell count was less than 50.

Supra national weights

Requirements

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accessibility	2.14	Supra-national weights included in dataset.	Including all the separate weighting components in the dataset allows users to decide to what extent they want to follow the same approach. Also, it improves the transparency of the approach followed by Ipsos/EF/CF, and increases future flexibility (reweighting could be limited to specific weighting steps and implications can be assessed at each step).	YES	YES	Final dataset

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accessibility	2.15	Procedure for constructing of, and sources used for, supra-national weights described in weighting report.	Transparency relies on systematic documentation of all phases of the survey preparation and implementation.	YES	YES	Sampling and weighting report

Ipsos included gross weights in the final dataset (indicator 2.14). These perform the same function as cross-national weights or within country weights with the same weight. By including the separate weighting components in the data set users can decide to what extent they want to follow the same approach.

Ipsos included the targets used for these and the calibration weights and outlined the procedure for constructing these weights in the final Sampling and weighting report (indicator 2.15) making the process transparent for Eurofound and Cedefop and other data users.

Target

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accuracy	2.13	Percentage of countries where the weights are based on up-to-date official population statistics collected within two years preceding fieldwork.	The more recently reference statistics are updated the less likely they will suffer from coverage issues. Furthermore, up-to-date statistics improve accuracy by reducing the risk of introducing an error in the weighting.	100%	93%	Sampling and weighting report

The result for indicator 2.13 is 93%. In 26 out of 28 countries, the reference statistics used in the weighting are based on data that is within two years preceding fieldwork. In the other two countries (Greece and Luxembourg) the data is more than two years old, in both cases the most recent available data were from 2016. These data might suffer from coverage issues given that they have not been updated as recently, but they were the best available.

Size of weights

Requirements

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accuracy	2.17	The size of the weights is limited (e.g. through trimming or weight shrinkage), based on a strategy that strikes an appropriate balance between representativeness in terms of the weighting variables and the effective samples size, which is fully documented and replicable.	Overly large weights and overly small weights are undesirable as they increase variance, at the same time, weights do need to be sufficiently correct for under- or over-coverage.	YES	YES	Sampling and weighting report

The size of the weights was limited (through trimming), based on a jointly agreed approach that aimed to strike an appropriate balance between representativeness in terms of the weighting variables and the effective samples size. The approach taken is outlined in the Sampling and weighting report.

3 Questionnaire

The quality indicators and targets relating to the questionnaire reflect the development process for the questionnaire - including advance translation and cognitive testing. This section also includes all phases of the translation process and the pre-test/pilot. Further details of each stage can be found in the respective reports on these activities - the Cognitive interview report, the Advance translation report, the Pilot report,⁹ the Translation report¹⁰ and the Technical and fieldwork report.¹¹

3.1 Questionnaire development

Requirements

A designated Steering Group was set up for consultation on the preparation and implementation of the ECS 2019. The Steering Group was made up of representatives from the Worker, Employer, Government and European Commissions groups on the Management boards of each of the two agencies – Eurofound and Cedefop.

⁹ The Cognitive interview report, the Advance translation report, and the Pilot report have not been published, but can be made available on request.

¹⁰ Eurofound and Cedefop (2020), *European Company Survey 2019: Translation report*, European Company Survey 2019 series, Eurofound working paper, Dublin (<https://www.eurofound.europa.eu/publications/report/2020/european-company-survey-2019-workplace-practices-unlocking-employee-potential#wp-101980>).

¹¹ Eurofound and Cedefop (2020), *European Company Survey 2019: Technical and fieldwork report*, European Company Survey 2019 series, Eurofound working paper, Dublin (<https://www.eurofound.europa.eu/publications/report/2020/european-company-survey-2019-workplace-practices-unlocking-employee-potential#wp-101978>).

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Relevance and timeliness	3.1	Questionnaire has been consulted with Eurofound and Cedefop's stakeholders	Consulting Eurofound and Cedefop's stakeholder ensures the relevance of the questionnaire	YES	YES	Documentation of steering group meetings

Questionnaire development was discussed in three Steering group meetings on 11 May 2017, 21 November 2017, and 20 November 2018 (indicator 3.1). For the meeting in November 2017 the Steering Group Members were provided with full drafts of the Management and Employee Representative questionnaires. Their feedback, provided during the meeting and/or in writing, was implemented in the version that was provided to Ipsos in preparation for the kick-off meeting for the project. The Steering Group were provided – for information – with updated drafts of the questionnaire after the findings from cognitive testing and advance translation were implemented. In the meeting on 20 November 2018 the results from the pilot test were presented to the Steering group, and when making the final changes to the questionnaire the feedback from the Steering group was used for guidance.

Targets

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accuracy	3.2	Percentage of questionnaire items in the final source questionnaire that meet international methodological standards of question design (such as outlined in Saris & Gallhofer (2007)).	Adherence to current academic standards ensures that the quality assurance (e.g. peer review) underlying these standards also applies to this project.	100%	100%	Documentation of expert consultation

Indicator 3.2 refers to the accuracy of the development of the final source questionnaires for the management respondents and employee representative respondents. To ensure this, the authors of the source questionnaires familiarised themselves with the literature on questionnaire design in cross-national surveys. Additionally, Eurofound and Cedefop organised a methodological review of the draft source questionnaires by five experts with varying expertise – all in survey research but some with a more substantive focus and others with a stronger focus on (online) survey methodology. Feedback from these experts, as well as feedback received from experts in Ipsos, was incorporated in the final versions of the source questionnaires.

3.2 Advance translation

Requirements

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Coherence and comparability	3.4	Percentage of the questions where substantive ambiguities are spotted, for which elaborate documentation of the consideration for translation is provided.	Transparency relies on systematic documentation of all phases of the survey preparation and implementation. Making the reasoning underlying any adaptations of the questionnaires following the advance translation explicit is important, as it can guide translation instructions, as well as any future adaptations of the items affected or of similar items.	100%	100%	Translation instructions / Documentation of advance translation (advance translation report)
Accessibility	3.5	Comprehensive documentation of the process of advance translation.	Transparency relies on systematic documentation of all phases of the survey preparation and implementation.	YES	YES	Documentation of advance translation (advance translation report)
Accuracy	3.6	Percentage of questionnaire items where substantive ambiguities are spotted for which either the source questionnaire is adjusted, or a translation instruction is drafted.	Substantive ambiguities in the source questionnaire should be eliminated as much as possible, as they might result in difference in the translated versions across languages. If they cannot be addressed by changing the question, they should be addressed by drafting clear and comprehensive translation instructions.	100%	100%	Translation instructions / Documentation of advance translation
Accessibility	3.7	Clear translation instructions.	Clear translation instructions decrease the likelihood of differences in meaning between translated versions.	YES	YES	Translation instructions

Indicators 3.4, 3.5 and 3.6 refer to the identification, addressing and documentation of the process of advance translation and of any substantive ambiguities that were observed during this process. As part of the advance translation, 13 items from the management questionnaire and seven from the employee representative questionnaire were marked as potentially ambiguous. All items were

corrected following the suggestions provided. Overall amendments or instructions were provided for all items where ambiguities were spotted (indicator 3.6), and the process and all observed ambiguities were discussed in the Advance translation report (indicators 3.4 and 3.5). The original Excel version of the questionnaires as well as the translation files produced by Ipsos provided clear translation instructions on an item-by-item level, including notes for translators, definitions of complex concepts and a description of the objective of the question, meaning that indicator 3.7 was met.

3.3 Cognitive testing

Requirements

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accuracy	3.10	A strategy for cognitive testing is agreed that that reflects current academic standards, as illustrated by academic references included in a document outlining the strategy for cognitive testing.	Adherence to current academic standards ensures that the quality assurance (e.g. peer review) underlying these standards also applies to this project.	YES	YES	Cognitive test strategy
Accessibility	3.12	Percentage of items included in the cognitive test for which systematic documentation is provided about the extent to which answers in the cognitive interviews correspond with the concepts that are intended to be captured by the questions (as indicated in the glossary).	Transparency relies on systematic documentation of all phases of the survey preparation and implementation. Furthermore, the extent to which the results of the cognitive interviews support that the questions capture the concepts as intended determines the validity (accuracy) of the questionnaire.	100%	100%	Cognitive interview report
Accuracy	3.13	Percentage of countries in which the selection of the respondents and composition of the sample corresponds with the agreed approach and design.	To ensure the reliability and generalisability of the cognitive test results the respondents need to be distributed meaningfully across a selection of characteristics. The more the achieved sample corresponds with the desired sample the better.	100%	0%	Cognitive interview report

Indicators 3.10, 3.12 and 3.13 refer to the development of a strategy, documentation of the process and adherence to the sampling approach for the cognitive testing that was undertaken as part of the questionnaire development process. A strategy was written by Ipsos incorporating the quality

elements described above and reflecting best practice in cross-national cognitive testing (indicator 3.10). This ensured that the QA underlying the academic standards in the strategy also applied to the cognitive testing that was undertaken on the ECS 2019. The results from the cognitive testing were systematically documented for all items in the cognitive interview testing report (indicator 3.12). The comprehensive report was divided into sections on methodology and findings documenting the approach taken for the cognitive testing as well as the results from each protocol (management representative, employee representative and materials) and for each of the four countries where testing was undertaken (UK, France, Germany and Poland).¹²

Quota sampling was used for the cognitive testing, with quotas set on six characteristics – type of respondent, size of establishment, sector, number of sites, number of years established and number of years in role. While many of the quotas were met, some flexibility was needed given the short timescale for recruitment and interviewing. Despite the deviations from the agreed quotas, the composition of the achieved sample meant that a range of respondents were recruited across almost all the quotas cells and that the questions were tested on those with characteristics closely resembling the target population, thus Ipsos was confident that missing this target did not impact the quality of the process itself. The quality target (3.13) for sampling focuses on the ‘percentage of countries’ and the result is 0% - since none of the countries were able to meet all of the quotas and each country failed on a different one. On reflection, the target (relating to the proportion of countries) is not particularly meaningful in relation to assessing the quality of the sample achieved compared to the agreed quotas set. It might be more appropriate to calculate the percentage of quota cells that have been met (rather than the % of countries that have met all the quotas). If this was applied for the ECS 2019, the result would be 73% - based on 84 cells in the quota grid being met in three out of the four countries. The UK was excluded from the calculation since specific quotas were not agreed for this country.

Targets

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accuracy	3.9	Number of questions for which 'major' issues are detected that are kept.	If major issues are observed for an item, it cannot be assumed that the item is valid, and should therefore not be retained. It should also be avoided that items are drastically changed after the cognitive test, as no further testing would be possible, so there would be no guarantee that the changed items perform better than the original flawed items.	0	4	Cognitive interview report

The outcome for indicator 3.9 was documented by Eurofound/Cedefop via an annex to the cognitive testing report. Of the six items¹³ that were considered ‘very problematic’, even after revision during

¹² Cognitive interviews in the UK were not required in the tender specifications, however, Ipsos carried out eight interviews in the UK to ensure a test of the source language version.

¹³ Five questions asked of management respondents; one of employee representatives.

cognitive testing, two items were cut, and four items were replaced. Out of the four items that were replaced, two were replaced by questions that were fielded in the previous wave of the ECS in 2013. For the remaining two items, alternative questions were formulated that were considered less complex but were included in the ECS 2019 untested, which was not ideal. So, the 'number of questions for which 'major' issues are detected that are kept' is four.

3.4 Translation

Translator training

Requirements

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Coherence and comparability	4.2	Percentage of translators and adjudicators that take part in translator training.	Appropriate training is crucial to ensure the translation process is carried out consistently across countries.	100%	100%	Attendance lists

Ipsos ensured that all translators and adjudicators involved in the translation process for the ECS 2019 took part in a translator training session as stipulated by indicator 4.2. The evidence that was retained in proof of this includes emails from the linguists confirming their preferred training session and a list of attendance for those sessions; for later sessions screen shots from a WebEx platform showing attendees were collected.

Translation materials

Requirements

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accessibility	4.3	Translation materials (e.g. translator instructions, and translator and adjudicator training materials) are constructed using input from the cognitive test and advance translation and are provided to the translators.	The quality of the training relies on the quality of the training materials, which should include all the implications of observations collected in the process of questionnaire development.	YES	YES	Translation report (for 'input & provision') / Translation instructions and training materials, training sessions (for 'provision to translators')

For indicator 4.3, the translation materials (e.g. translation file and translator/adjudicator instructions) were created by Ipsos from the revised source questionnaire after cognitive testing. The translation materials were sent in advance to the linguists for the translator training. The

translation file contains a column with translator instructions that are based on the outcome of the cognitive testing and the advance translation.¹⁴

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accessibility	4.4	Translation materials are made publicly available.	Transparency depends on the availability of documentation on all stages of the survey process.	YES	NO	Availability on Eurofound website

Contrary, to the target set for indicator 4.4, the translation materials for the ECS have not been made public. At this point, it is only foreseen that the translations themselves will be made publicly available although further documentation will be available on request. The implication of not meeting this target is a slight reduction of transparency, however this is partially off-set by the fact that documentation is made available on request.

Translation languages

Requirements

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Coherence and comparability	4.5	Percentage of languages for which translation or adaptation follows the agreed process.	The translation and adaptation processes are designed to ensure equivalence across languages and countries, and therefore should be followed closely in each language.	100%	100%	Translation report

Ipsos, Eurofound and Cedefop agreed that the translation process for 31 countries would follow TRAPD procedures and that in the remaining country - Cyprus – an adaptation process would be implemented instead (indicator 4.5). The percentage of language versions for which the translation or adaptation follows the agreed process was 100%.

Target

Ipsos translated the ECS 2019 materials into one additional language – German in Luxembourg – meeting indicator 4.6. This ensured that potential respondents in Luxembourg could be addressed and respond in German if this was the language in which they were most comfortable responding in.

¹⁴ For example, clarifying key terms to ensure that the correct wording/phrasing is chosen to ensure comparability.

Quality Dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Coherence and comparability	4.6	Number of additional languages - beyond those for which translation is required - for which translation is carried out following the agreed process for translation.	It is assumed that potential respondents are more likely to respond, and are more engaged when addressed in the language in which they are most comfortable. Adding additional languages increases the likelihood that the survey is available in the language that a (potential) respondent is most comfortable with.	1	1	Translation report; questionnaire(s)

Initial translation

Requirements

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accessibility	4.7	Percentage of languages, out of those for which translation is required, for which systematic documentation of results of initial translation is provided.	Transparency relies on systematic documentation of all phases of the survey preparation and implementation.	100%	100%	Translation documentation (Translation log file)

Ipsos systematically documented the results of the initial translations for all languages being used in the ECS 2019, meaning that the result for indicator 4.7 was 100%. The documentation ensures that the initial translation process is accessible and transparent.

Within country adjudication

Requirements

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accessibility	4.9	Percentage of languages, out of those for which translation is required, for which systematic documentation in English is provided about the process and results of adjudication.	Transparency relies on systematic documentation of all phases of the survey preparation and implementation.	100%	100%	Translation documentation

Ipsos systematically documented the results of the adjudication process for all languages being used in the ECS 2019, meaning that the result for this indicator 4.9 was 100%. The documentation ensures that the adjudication process is accessible and transparent.

Cross country adjudication

Requirements

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Coherence and comparability	4.11	Percentage of cross-national review sessions, in which adjudicators from each of the countries sharing the particular language participate.	The translation and adaptation processes are designed to ensure equivalence across languages and countries, and therefore should be followed closely in each language.	100%	100%	Attendance lists
Accessibility	4.12	Percentage of languages, for which a cross-national review is required, for which systematic documentation in English is provided about the process and results of the cross-national review.	Transparency relies on systematic documentation of all phases of the survey preparation and implementation.	100%	100%	Translation documentation

To achieve indicator 4.11, Ipsos ensured that adjudicators from each of the countries that shared a language participated at all (100% of) the cross-national review sessions. This included an additional cross-national review session that was organised between Croatia, Montenegro and Serbia at Eurofound and Cedefop's request. Ipsos retained attendance lists showing the names of participants at each of the sessions.

For indicator 4.12, Ipsos provided Eurofound and Cedefop with comprehensive documentation in English about the process and results of the cross-national review for all (100% of) languages. The translation documentation ensures that the adjudication process is accessible and transparent.

Edit final translated questionnaires

Targets

A very small proportion of final translated questionnaire items required editing – 0.01% (indicator 4.14). This illustrates that the translation process for the ECS 2019 was implemented and

documented to a high standard. Of all the question items, in all language versions (8,208 items¹⁵), only 94 were edited.

Quality Dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Coherence and comparability	4.14	Percentage of final translated questionnaire items that required editing (e.g. correcting typo's, copying & paste errors, etc.).	If the translation process is implemented to a high standard, the final review of the translated questionnaire items should not reveal any items that require further editing.	<5%	0.01%	Translation documentation

As part of the standard fieldwork preparation processes for all surveys Ipsos carries out, translated versions of the script are overlaid in all countries. For the ECS 2019, a mistake was made in the language overlaying process for the management online survey in Sweden. In the local language script, the NACE sectors were not correctly sorted. This affected both the pilot and mainstage surveys. Ipsos regard this as a translation issue rather than a scripting error since the problem occurred in the translated version of the script. This has been documented further in the Technical and fieldwork report.

3.5 Pilot

Requirements

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accuracy	8.1	Percentage of countries where pilot interviews are carried out with at least 30 management respondents and at least 30 employee representative respondents (except in the smaller countries, where lower targets were agreed - 15 and 5 in Cyprus, Luxembourg and Montenegro, and 7 and 5 in Malta). In the larger countries that require a 'no size' stratum (Austria, Greece, North Macedonia and Spain) the MM target was reduced to 25.	Sampling target are sets to enable desired analysis. Targets need to be met to ensure this.	100%	0% for ER and 39% for MM (11 out of 28 countries)	Pilot report

The results for indicator 8.1 fell short of target. Ipsos calculated the results based on the final data from the end of pilot fieldwork. It was not possible to achieve the target number of pilot interviews with employee representative respondents in any of the 32 countries. However, in 11 out of 28 countries (39%), interviews were achieved with the target number of management respondents (AT,

¹⁵ 36 language versions in total multiplied by 228 items = 8,208 items (across all languages).

HR, LV, LT, MT, NL, PT, RO, SI, ES, MK). The low number of completes for employee representative respondents was due to a combination of factors: incomplete knowledge about presence of employee representation among respondents in screener interviews, refusals to give contact details of employee representatives during screener interviews, and low CAWI conversion overall (both for management respondents and employee representative respondents). These reasons are documented in more detail in the Pilot report for the ECS 2019.

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accuracy	8.2	Pilot interviews are carried out with at least 15 management respondents and at least 5 employee representative respondents in Cyprus, Luxemburg and Montenegro and at least 7 management and 5 employee respondents in Malta).	Sampling target are sets to enable desired analysis. Targets need to be met to ensure this.	100%	0% for ER and 25% for MM (1 out of 4 countries)	Pilot report

The results for indicator 8.2 also fell short of target. Ipsos calculated the results based on the final data from the end of pilot fieldwork. These were compared against the targets agreed in the Pilot sampling strategy. Ipsos fully worked the gross samples that were agreed with Eurofound and Cedefop prior to the start of pilot fieldwork. The only exception being in countries where it was agreed *during* pilot fieldwork that remaining sample should be saved for the mainstage. Despite this, Ipsos was unable to achieve either of the targets in CY, LU or ME – resulting in 0% of the target for employee representative respondents. For MT, the target for interviews with employee representative respondents was not met but the target relating to management representatives was – resulting in 25% (one out of four countries). As mentioned in relation to indicator 8.1, the low number of completes for employee representative respondents was due to a combination of factors: incomplete knowledge about presence of employee representation among respondents in screener interviews, refusals to give contact details of employee representatives during screener interviews, and low CAWI conversion overall (both for management respondents and employee representative respondents). These reasons are documented in more detail in the Pilot report for the ECS 2019.

By missing the target number of interviews set at indicator 8.1 and 8.2 limited the possibility to assess the functioning of the questionnaire in each language and to test the gross sample requirements. However, it did not compromise the test of the contact strategies.

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accuracy	8.3	Percentage of countries where pilot interviews are carried out in all local languages.	The pilot is supposed to be a full dress-rehearsal, as well as a test of questionnaire functioning, so should cover all languages that are to be used in mainstage.	100%	71%	Pilot report; Pilot data set

Ipsos completed interviews in all the local languages but did not reach the target number of interviews for LU and MT - so the target was met in five out of seven countries (71%). The results for each country can be seen in Table 1.5 below.

Table 1.5: Pilot interviews by language

Country	Language	MM online interviews	ER online interviews	Total interviews (MM & ER)
Belgium	Dutch	24	5	29
	French	1	0	1
Estonia	Estonian	21	1	22
	Russian	3	0	3
Latvia	Latvian	30	0	30
	Russian	0	0	0
Luxembourg	French	11	0	11
	German	2	0	2
Malta	Maltese	3	0	3
	English	20	1	21
Spain	Spanish	24	5	29
	Catalan	5	0	5
North Macedonia	Macedonian	25	5	30
	Albanian	3	0	3

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accuracy	8.4	Percentage of issues detected in the pilot test for which a solution is agreed.	The pilot is aimed to detect issues with the aim to address them, not to leave them unresolved.	100%	100%	Email(s) exchanges between Ipsos and client; Pilot report
Accuracy	8.5	Percentage of translation issues detected in the pilot, for which the solution was based on the input of two independent translators.	Translation issues should be addressed following approach that meets a similar standard as the general translation process. Therefore, issues should be addressed by involving at least two independent translators.	100%	100%	Pilot report

The targets for indicators 8.4 and 8.5 relating to the resolution of issues and translation issues arising from the pilot test were both achieved with results of 100% for each indicator. For 8.4, a solution

was agreed on for each of the issues that were encountered in the pilot test and these described in the Pilot Report. For 8.5, a translator and adjudicator were consulted for each of the proposed amendments to the translated questionnaires from the pilot where their input was required.

4 Fieldwork

The quality indicators and targets relating to fieldwork reflect several stages to prepare for fieldwork in multiple countries for the ECS 2019. The first set of indicators relate to the fieldwork infrastructure and the second to scripting and script checks. This is followed by dedicated indicators on interviewer training and monitoring followed by indicators dedicated to assessing the quality of the contact strategy. Further details of these areas can be found in the Technical and fieldwork report.¹⁶

4.1 Fieldwork infrastructure

Target

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Coherence and comparability	5.1	A common integrated system for sampling management and all stages of interview administration is used in all countries.	A common integrated system for sample management avoids issues with harmonising fieldwork monitoring data, delays in reporting, delays in sample release, etc. Anomalies can be detected and QA resource prioritised.	YES	NO	Technical and fieldwork report

Dimensions software was used in all countries for the online and CATI fieldwork undertaken for the ECS 2019. The Dimensions CATI scripts were accessible directly in 16 countries and indirectly via web links in 16 countries.¹⁷ Those using the CATI links used a different system that was not fully integrated, thus indicator 5.1 was not met for either the pilot or mainstage fieldwork. By using a common integrated system in all countries issues with harmonising fieldwork monitoring data, delays in reporting, delays in sample release, etc. can be avoided. It also enables quality checking to be planned and scheduled, since the data are available at the same time every week for all countries. This should allow developing problems to be detected early and solutions found which decreases risk. In practice, for the ECS 2019, there were occasional lags in reporting, which meant

¹⁶ Eurofound and Cedefop (2020), *European Company Survey 2019: Technical and fieldwork report*, European Company Survey 2019 series, Eurofound working paper, Dublin (<https://www.eurofound.europa.eu/publications/report/2020/european-company-survey-2019-workplace-practices-unlocking-employee-potential#wp-101978>).

¹⁷ Direct access: BE, BG, DE, ES, FI, FR, HR, IE, IT, LU, NL, PT, RO, SI, SE and UK.
Web links: AT, CY, CZ, DK, EE, GR, HU, LT, LV, MT, PL, SK, TU, RS, ME and MK.

that there were gaps in the weekly fieldwork monitoring information, although no major problems occurred.

4.2 Scripting and script checks (consistency checks)

Requirements

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Coherence and comparability	6.2	Same script used in all countries.	A common script avoids errors, due to differences between countries in questionnaire administration.	YES	YES	Technical and fieldwork report / Questionnaire script

Ipsos developed a common script that was used in all countries for both the pilot and the mainstage meeting the requirements of indicator 6.2.

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accuracy	6.4	Number of consistency rules (soft and hard) identified.	Consistency rules can improve data quality, because issues are flagged at the time of interview administration, at which point they can still be addressed. A more precise target could not be set, as the questionnaire had not been finalised when the QAP was compiled.	>0	YES	Technical and fieldwork report / Questionnaire script
Accessibility	6.6	Comprehensive documentation of all consistency rules.	Transparency relies on systematic documentation of all phases of the survey preparation and implementation.	YES	YES	Technical and fieldwork report

To meet quality indicator 6.4 consistency rules were identified for the management representative and employee representative surveys and for the screener questionnaire – for both the pilot and mainstage. There were no consistency checks included in the reminder script. The consistency rules consisted of ‘soft checks’. Respondents were notified if inconsistencies were identified in their answers, but they were permitted to ignore the warnings.

For the pilot, in total 54 consistency rules were identified. For the management survey, there were 36 - nine of which were for the skips where respondents click ‘next’ without answering a question. For the employee representative survey, there were 12 consistency checks, four of which were for the skips where respondents click ‘next’ without answering a question. For the screener, six consistency checks were included. All checks have been incorporated into the script. For the mainstage, the total number of rules remained the same (N=54) but some changes were made after the pilot. There was one fewer check for both the management survey (N=35) and the employee

representative survey (N=11) and two additional checks were introduced to the screener (N=8). All checks were incorporated into the script. All consistency rules that were applied in the pilot were documented in the Pilot report, similarly those applied in the mainstage are documented in the Fieldwork report (indicator 6.6).

Target

Quality dimension	#	Indicator	Rationale	Targ.	Ach	Evidence
Accuracy	6.1	Number of scripting errors encountered in the pilot test.	If the scripting process is of high quality, the pilot test should not reveal any scripting errors.	0	1	Pilot report

One scripting error was detected in the pilot test (indicator 6.1). This was found in the routing for the random assignment of scenarios for additional establishments in the screener questionnaire. In the pilot, four scenarios were used for selecting the respondents for the screener questionnaire. Additional establishments (i.e. those added to the sample via the recruitment interview with the originally sampled establishment) should have been randomly assigned to one of the four scenarios, but due to the scripting error, they were assigned to just two scenarios. The scripting error had an impact on the size of the groups assigned to each scenario – but the impact was minor because few records were added to the samples via this procedure.

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accuracy	6.5	Percentage of identified consistency rules integrated in the script.	Based on the questionnaire it is possible to identify all possible inconsistencies that could occur, checks for which can subsequently be integrated in the script.	100%	100% - pilot and mainstage	Technical and fieldwork report / Questionnaire script

The consistency rules described in relation to indicator 6.4 were all integrated into the respective scripts – management representative, employee representative and the screener, meaning that this indicator (6.5) was fully met for both the pilot and the mainstage of the ECS 2019.

4.3 Interviewer training and monitoring

Requirements

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Coherence and comparability	7.1	Percentage of local project leaders/country coordinators attending the fieldwork manager instruction meeting.	The quality of the work relies on clear instruction of the local agencies, as well as on the extent to which they feel ownership of the project, which is supported by exchanging views with them at several points during the process.	100%	100% - Pilot; 79% - Mainstage	List of participants

Indicator 7.1 relates to attendance of local project leaders/country coordinators at the instruction meetings that were held prior to the pilot and mainstage fieldwork. For the pilot, Eurofound, Cedefop and Ipsos agreed that the indicator referred to the training of the ‘person who has the day-to-day responsibility for managing and monitoring pilot fieldwork’ rather than the ‘local project leaders/country coordinators’ as specified. The target was met in all countries for the pilot. In Belgium, Sweden, Montenegro and Turkey, the local project leader could not participate, and was replaced by another member of the local team. In each of these countries, the person taking part in the training was the person who had day-to-day responsibility for managing and monitoring pilot fieldwork, and who reported back to the local project leader.

For the mainstage, the result (79%) is based on 22 out of 28 project leaders / country coordinators attending the meeting. In Belgium, Cyprus, France, Malta and Slovenia, the local project leader could not participate, and was replaced by another member of the local team. In each of these countries, the person taking part in the training was the person who had day-to-day responsibility for managing and monitoring fieldwork and reported back to the local project leader. For Estonia, the local project leader was replaced by the CEO of the company. Eurofound and Cedefop approved the proposed replacements before the meeting.

As mentioned in the rationale - the quality of the work by the local agencies relies on clear instruction and addressing any errors, ambiguities or misunderstandings in advance of fieldwork. Involving the local leaders/coordinators gives a sense of ownership and an additional layer of quality assurance at the country level. Whilst the target could not be fully met for the mainstage all those attending the training were responsible for managing and monitoring fieldwork on a day-to-day basis and had a key role in the delivery of the ECS 2019. Therefore, Ipsos did not think that the quality of the fieldwork was adversely affected by another member of the team attending the training compared to the project leader/country coordinator since those that attended were responsible for the fieldwork in their respective countries.

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accuracy	7.3	Interviewer training materials are comprehensive, fully covering the process for respondent selection and recruitment as well as for questionnaire administration.	The quality of the training relies on the quality of the training materials.	YES	YES	Interviewer training materials
Coherence and comparability	7.4	Interviewer training materials are translated into the working languages of all the country teams.	Training is more effective if it is received in a language the trainee is completely comfortable in.	YES	YES	Interviewer training materials
Accuracy	7.6	Percentage of interviewers that take part in the training.	All interviewers should receive project specific training.	100%	100%	Attendance lists
Punctuality	7.7	Percentage of interviewers that are trained before they participate in fieldwork.	Training should take place before the work is carried out.	100%	100%	Attendance lists; email confirming the fact to client

Indicators 7.3 and 7.4 both relate to the quality of the interviewer training materials – specifically that they are fully comprehensive and translated into the working languages of all country teams. These indicators were met for both the pilot and mainstage surveys. Indicators 7.6 and 7.7 relate to the percentage of interviewers taking part in project-specific training and being trained before participating in fieldwork (regardless of whether they start work at the beginning of the fieldwork or join the project later). Running the training sessions provides an additional layer of quality assurance to detect and address issues before they become problems during fieldwork. These indicators were met for both the pilot and the mainstage survey (with results of 100% recorded for each survey). On reflection, there is some overlap in indicators 7.6 and 7.7 which seem to measure the same thing. In future, indicator 7.7 would suffice and indicator 7.6 could be dropped in future. This is discussed further in Section 9 of this report.

4.4 Contact strategy

Requirements

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accuracy	9.2	Percentage of gross sample entries that are discarded before the net sample is realised, for which a final outcome has not been realised (i.e. no cases to be lost).	Fieldwork rules are designed to ensure that all sample entries have a reasonable chance of being included in the survey. Consequently, sample entries should not be discarded before a final outcome is achieved, in accordance with these fieldwork rules.	0% ¹⁸	0%	Technical and fieldwork report; Weekly reporting; Contact data
Accuracy	9.4	Percentage of sample entries to which a final status of 'non-contact' was assigned that were not called at least six times.	It is a requirement that six contact attempts are made to a sample entry before it can be marked a 'non-contact'. It is possible that the phonenumber is busy, or that there is nobody available to answer the call, but it is unlikely that this is the case on six separate occasions. The fewer the number of contact attempts, the more likely viable respondents are excluded from the study.	0%	0%	Technical and fieldwork report; Weekly reporting; Contact data
Accuracy	9.5	Percentage of respondents that agreed to participate but for which final non-response is accepted that were not reminded with at least three emails and at least one phone call.	The reminder strategy is designed to maximise response rate. Cases cannot be discarded as 'non-response' without all reminders having been issued.	0%	0%	Technical and fieldwork report; Weekly reporting; Contact data

Indicators 9.2, 9.4 and 9.5 are all requirements related to the contact strategy and are intended to maximise response rates and minimise non-contact and non-response. Indicator 9.2 refers to the proportion of gross sample entries that are discarded before the net sample is realised that do not have a final outcome. The result is 0%. It was possible to achieve this result through close monitoring

¹⁸ The original target for this indicator was stated as 100%, however during reporting this was identified as a mistake and amended to 0%. The only reason permitted for abandoning cases is the finalisation of fieldwork – as complete efficiency is not feasible.

of sample usage and checking the assignment of outcome codes during fieldwork monitoring and reporting.

Indicator 9.4 focuses on the proportion of sample entries to which a final status of 'non-contact' was assigned before the finalisation of fieldwork that were not called at least six times. The result is also 0%. After excluding the cases abandoned upon finalizing the fieldwork, the percentage of sample entries that were not called at least 6 times across all countries was 0%. Looking at countries individually, there were two countries with a percentage higher than 0 of gross sample entries that were not called at least 6 times - Italy (2.8%) and Estonia (8.9%).

Indicator 9.5 refers to the percentage of respondents that agree to participate [in follow-up] but for which final non-response is accepted that were not reminded with at least three emails and at least one phone call. The target was 0% and the result was also 0%.

Targets

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accuracy	9.1	Percentage of countries where the used gross sample is smaller or equal to the planned gross sample.	Sampling and sample release should be planned such, that it is not necessary to order additional sample during fieldwork, as this might cause delays and runs the risk of issues with selection probabilities.	100%	56%	Technical and fieldwork report; Weekly reporting

For nine out of sixteen countries who reached their target (56%), the gross sample was smaller or equal to the planned gross sample (see Table 1.6). The table includes the 16 countries where the target number of interviews was met and highlights (in blue) the nine in which the gross sample was smaller than the planned gross sample.

Table 1.6: Gross sample usage

Country	Gross sample	Sample dialled
Austria	15611	17961
Belgium	18721	13569
Bulgaria	14859	13449
Croatia	3512	5068
Denmark	15959	11275
Estonia	5294	4572
Finland	8614	7307
Greece	9672	8966
Hungary	25294	21018
Ireland	3576	8158
Latvia	4129	4819

Country	Gross sample	Sample dialled
Lithuania	3872	3242
Netherlands	7440	13634
Slovakia	8200	8594
Slovenia	3564	3734
Sweden	20346	12834

Given that this was a new methodology, Ipsos did not have a lot of information to start from to calculate the gross sample size. The main information used was the pilot CATI and CAWI yield – but these were very small samples so did not offer an ideal starting point. At the start of the main stage fieldwork, **MM targets** were set across the nine sampling strata (three size categories by three sector categories) for each country. The MM targets were then extrapolated to completed **CATI screener fieldwork targets**, by multiplying the MM targets by an overall predicted CAWI yield ratio for each country. The ratio was based on the pilot yield rates, with consideration given to the potential for improvements in the main stage. **Gross sample sizes** were then set on the nine strata, by extrapolating to the total sample expected to be needed to reach the CATI screener targets. This was set to the overall expected country CATI yield level based on the pilot, with variation by strata based on the yield rates from the ECS 2013. Whilst additional sample was ordered before fieldwork started, it was still necessary to order additional sample during fieldwork in some countries. This has implications for calculating selection probabilities and for the timing of fieldwork as sample introduced later still needs to follow the same contacting rules as those sampled earlier.

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accuracy	9.3	Percentage of gross sample entries that are contacted at least once, but are discarded upon realising the net sample (i.e. cases that were opened but not finalised).	Similar to 9.2, however, in this case it refers to the planning of sampling release, which should ensure that sample entries are contacted gradually, to avoid contacting a lot of cases very late in fieldwork or leaving many cases open until very late in fieldwork.	<10%	5%	Technical and fieldwork report; Weekly reporting; Contact data

The proportion of gross sample entries with fewer than six call attempts and no final outcome upon realising the net sample was 5% across all countries. Only Italy had more than 10% discarded open sample by the end of fieldwork: 13% of the cases were left in the “recall” queue and had fewer than six call attempts. Of these cases left in the “recall” queue, 28% had erroneously been assigned the call outcome of “possible wrong number” instead of “wrong number”. The CATI system is set up in such a way that records coded as “wrong number” are closed from further contacting, while records coded as “possible wrong number” stay available in case an alternative phone number can be found. At the end of the fieldwork period, phone lookups were stopped and roughly 900 records had “possible wrong number” (no alternative number had been found) as a final call outcome. The

percentage of discarded open sample by the end of fieldwork, after excluding the records with final call outcome “possible wrong number”, was 9%.

5 Fieldwork monitoring

5.1 Interviewer monitoring

The quality indicators and targets relating to fieldwork monitoring for the ECS 2019 focus on monitoring interviewers, monitoring fieldwork progress against targets as well as data validation following fieldwork closure. Further details of can be found in the Technical and fieldwork report.¹⁹

Requirements

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accuracy	10.1	An explicit, comprehensive and discriminative interviewer monitoring strategy is outlined.	Interviewer behaviour is a key factor for fieldwork success. A monitoring strategy needs to be in place to ensure that ‘bad behaviour’ can be addressed, and good practices can be promoted and shared.	YES	YES	Fieldwork strategy
Punctuality	10.2	The interviewer monitoring strategy is sent to the Client for approval prior to fieldwork.	Client reviews are intended to ensure a comprehensive document is produced and any problems rectified before implementation.	YES	YES	Email to client delivering fieldwork strategy

Indicators 10.1 and 10.2 refer to the development of an interviewer monitoring strategy. Ipsos developed a fieldwork monitoring strategy that included a section on the supervision of interviewers (10.1). The strategy ensured that problems and issues could be identified and addressed, and good practices could be recognised and promoted to encourage other interviewers. The interviewer monitoring strategy was sent to Eurofound and Cedefop for approval prior to implementation (10.2).

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accuracy	10.15	% of countries where at least 90% of [interview] target was achieved.	Sampling targets are set to enable desired analysis. Targets need to	100%	82%	Weekly reporting; Technical and fieldwork report

¹⁹ Eurofound and Cedefop (2020), *European Company Survey 2019: Technical and fieldwork report*, European Company Survey 2019 series, Eurofound working paper, Dublin (<https://www.eurofound.europa.eu/publications/report/2020/european-company-survey-2019-workplace-practices-unlocking-employee-potential#wp-101978>).

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accuracy	10.16	At least 95% of [interview] target across countries reached.	be met to ensure this.	YES	YES	Weekly reporting; Technical and fieldwork report

Indicators 10.15 and 10.16 relate to the target number of interviews – these were set to enable the desired analyses to be completed with enough respondents. Over 95% of the overall interview target was reached (indicator 10.16). However, only in 82% of countries, at least 90% of the interview target was achieved (10.15). Five countries (CY, DE, MT, PL, and RO) did not achieve the 90% target, out of the 28 countries that participated in the mainstage. The reasons for this varied.

For Cyprus and Malta, the frames were exhausted before the targets were met. Before the start of the main stage fieldwork, Ipsos' partner agency in Cyprus suggested that an application for a new sampling frame might be successful (this had been declined at the pilot stage). Eurofound/Cedefop agreed to proceed with this and the outcome was that a new sampling frame was received, containing only cases that the statistical authority confirmed as having 10 or more employees. This new register indicated that the target population was smaller than anticipated. After a matching exercise to add phone numbers (and remove the companies that had been contacted in the pilot phase), a total of 3,772 companies was left and this final gross sample was much smaller than anticipated. The sampling frame in Malta included 'no size' cases to boost coverage and sample size achievement. At the pilot stage, the total number of cases on the Dun & Bradstreet (D&B) sampling frame, including those without size (69% out of the total number of cases), was similar to the population (2,346 companies compared with 2,178 according to the available reference statistics), meaning many of the cases were expected to be eligible, which was confirmed by the pilot. In the sample for mainstage fieldwork that was received from D&B, the available numbers was considerably smaller: a total of 1,870, against 2,346 expected (less the cases used in the pilot). The full reduction was due to cases without size. The number of cases 'with size' had only reduced by two cases. D&B explained the reduction of the cases without a size indication as due to ongoing cleaning of the data.

In both countries additional efforts were made to improve the yield rate, but these did efforts did not sufficiently increase the yield rate to realise the target samples.

In Germany, the target was not met due to low CATI yield coupled with low CAWI conversion. The CAWI yield in Germany was the lowest observed across the countries in ECS 2019. This very low level of cooperation in the online survey was also observed at all other stages of the fieldwork (refusals to participate in the CATI screener, refusals to provide email address during screener interview, refusals during the CATI reminder to receive further reminder emails etc.). Several additional measures were taken which led to some improvement, but this was not enough to reach the required target, within the agreed budget and timeframe.

For Poland, the issue was low CATI yield. The CATI screener fieldwork was very challenging, due to a high ineligibility rate, a high non-contact rate and a high refusal rate. Interviewers reported a low level of interest in the study, partially due to the choice for an online mode.

Finally, in Romania, the CATI yield in the mainstage was lower than that of the pilot. Despite ordering and calling additional sample cases, the target was not reached. The projections had shown that the target would be achieved but the CAWI conversion was lower than anticipated in the last few weeks of fieldwork - so it was not possible to reach it in the end. Since the extra sample order did not contain large companies²⁰, the CATI yield was lower than for the original sample;²¹ and, as such, fewer screeners than foreseen were completed. 92% of the target number of screeners was completed; for the CAWI completes, however, there was a larger gap – 83% of the target number of MM online interviews was completed. This is due to the large difference in CAWI yield between the original sample and the additional sample, a difference that could only be observed at the end of the fieldwork period (due to the time that passes between sending out invitations and online completes), when no further actions could be taken. A number of factors can explain this lower CAWI yield: (1) the additional sample contained a larger share of establishments in the services sector, which was characterised by a lower CAWI yield; (2) a larger, more mixed, pool of interviewers completed screeners for the additional sample order, of whom some were less good at getting cooperation from MM respondents.

Further details on the problems experienced in CY, DE, MT, PL, and RO and the measures taken to try to resolve these issues can be found in the Technical and fieldwork report.

Targets

For indicator 10.3, Ipsos, Eurofound and Cedefop agreed to define 'issues' as any problems relating to slow fieldwork progress or where there was a concern that a weekly target might not have been met.

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accuracy	10.3	Percentage of issues identified based on information in weekly monitoring data for which a solution is provided.	Weekly monitoring aims to detect issues and address them, not leave them unresolved.	100%	88%	Technical and fieldwork report; Weekly reporting; Contact data
Accuracy	10.4	Percentage of countries where at least 10% of successful screening interviews are monitored.	Interviewer behaviour is a key factor for fieldwork success. A monitoring strategy	100%	100%	Technical and fieldwork report; Weekly reporting; Contact data

²⁰ All large companies (250+ employees) available on the sample frame had already been included in the original sample order.

²¹ This because the CATI yield was better for large companies.

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accuracy	10.5	Percentage of countries where at least 10% of follow-up calls for CAWI completion are monitored.	needs to be in place to ensure that 'bad behaviour' can be addressed, and good practices can be promoted and shared.	100%	100%	Technical and fieldwork report; Weekly reporting; Contact data

The RAG status of the weekly reports was reviewed by Ipsos to identify the countries in which issues arose as well as those that were resolved and those that were not. Ipsos assessed whether an issue was 'resolved' by looking at if the situation improved, regardless of whether the improvement was sufficient to meet the overall targets.²² Ipsos identified 33 issues - 29 of which were resolved (88%). Four issues were not resolved by the end of fieldwork – two of these related to the sample frame issues in CY and MT (mentioned above related to indicator 10.15). The remaining two issues were in Germany and Poland.

In Germany, low CAWI conversion was an ongoing issue that was not completely resolved by the end of the fieldwork period. Many extra measures were agreed with Eurofound and Cedefop in an attempt to improve the CAWI yield:

- (1) From 24 April (week 14) until the end of the CATI screener fieldwork, the reminder schedule for MM was changed, and the CATI reminder was conducted after the 1st email reminder, instead of after the 2nd email reminder. This allowed interviewers to follow up with MM respondents more quickly after the invitation email.
- (2) From 15 May (week 17) until the end of the CATI screener fieldwork, to keep interviewers motivated during the last weeks of fieldwork, interviewers were offered an additional incentive if respondents that they had recruited went on to complete the online survey.
- (3) From week 17 onwards, interviewers referred to a charity donation during the screener interviews, and a reference to this donation was also added to the email reminders. The reference to this donation was made in the same way as used in another large-scale survey.²³
- (4) Between 29 May and 5 June 2019, 165 MM respondents who were recruited during the period before Easter (between 2 and 11 April 2019), and who received reminder emails during this holiday period, were contacted a second time by phone to remind them of the survey (and explain that it was not yet too late to participate). All these respondents had indicated during the 1st CATI reminder that they would complete the survey. However, just 25 of these respondents were reached again and only three went on to complete the online survey.

²² Ipsos did not focus on whether the overall interview targets were met in the end, as this was already covered by indicators 10.15 and 10.16.

²³ Für die Teilnahme an der Online-Befragung spendet Ipsos einen Gesamtbetrag in Höhe von 2.000 EURO an das Kinder- und Jugendhilfswerk „Die Arche“. English for info: For participating in the online survey, Ipsos donates a total amount of 2,000 EUR to the children's and youth project "Die Arche".

Unfortunately, the extra measures did not make a significant difference to the fieldwork outcomes.

Similarly, in Poland, very low CATI yield was observed across all fieldwork weeks. During week 5, a number of measures were taken in an attempt to improve the CATI yield:

- (1) Adjustments in the selection of interviewers based on their experience, CATI yield and CAWI yield achieved;
- (2) Additional listening-in, focusing on identifying each interviewer's strengths and weaknesses in light of achieved response rates and CAWI yield rates;
- (3) Refresher training on B2B interviewing, followed by individual conversations about strengths and weaknesses of each interviewer;
- (4) Additional training on ECS 2019, followed by role playing exercises focused on getting past gatekeepers; and
- (5) During this training, further explanation was also provided on using the full functionality of Dimensions, so that more reassurance emails would be sent as part of the recruitment process.

The implication of the issues that were not resolved is that the target of 'at least 90% of interview targets to be achieved' in 100% of countries (indicator 10.15) was not met.

The targets for 10.4 and 10.5 were both met – all countries monitored at least 10% of successful screening interviews and at least 10% of follow-up calls (CATI reminders) for CAWI completion were monitored. This is documented in the Technical and fieldwork report.

5.2 Fieldwork monitoring

Requirements

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accessibility	10.6	Percentage of countries covered in weekly monitoring data (in accordance with template).	Monitoring needs to be comprehensive to be effective.	100%	96%	Weekly reporting

All countries were covered in the weekly monitoring data as required by indicator 10.6, however it was not possible to include Estonia every week. So, the result of 96% is based on 27 out of 28 countries. Whilst the quantitative data from Estonia was missing from the reporting, partial information was in the RAG report and the quantitative data was collected and checked by Ipsos ensuring that comprehensive monitoring could take place in all countries thereby minimising the impact on the accuracy and accessibility of the fieldwork monitoring.

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Punctuality	10.7	Number of times that the weekly monitoring data for the preceding week is not delivered on Tuesday by the end of business, without prior agreement.	Monitoring needs to be timely to be effective.	0	2	Emails to client delivering weekly reports

Ipsos submitted 21 weekly monitoring reports (for 21 weeks of CATI screener fieldwork) having agreed with Eurofound and Cedefop to stop reporting after that week. Two fieldwork reports were delivered late (week 1 and week 3) so the target for indicator 10.7 was missed on these two occasions. However, given that these were relatively early in the fieldwork period the impact that this had on monitoring progress was minimal. One further report was delivered on a Wednesday, but this was agreed beforehand (week 14).

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Punctuality	10.8	Number of times that the quantitative indicators in the weekly monitoring data and the progress and projections (of end date) are not checked by the following Thursday by the end of business, without prior agreement.	Feedback on monitoring needs to be provided in a timely manner to be effective.	0	0	Emails responding to weekly reports / Minutes of weekly calls

Eurofound and Cedefop intended to check and provide timely written feedback to Ipsos on the monitoring data as specified in indicator 10.8. In practice, however, any issues detected were raised in the weekly client/contractor calls on the Thursday following delivery on the Tuesday (or Wednesday) the same week. Eurofound and Cedefop raised at least some issues or observations on every call – indicating that the weekly monitoring data had been checked. This is evidenced by the minutes of the weekly calls.

5.3 Data validation

Requirements

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accuracy	10.10	An explicit, comprehensive and discriminative data validation strategy is outlined.	A strategy needs to be in place to assess the quality of the data, and ultimately the acceptability of cases as valid.	YES	YES	Data validation strategy
Punctuality	10.11	The data validation strategy is sent to the Client for approval prior to fieldwork.	Client reviews are intended to ensure a comprehensive document is produced and any problems rectified before implementation.	YES	YES	Email to client delivering data validation strategy

Ipsos developed a fieldwork monitoring strategy that included a data validation strategy, which was agreed prior to mainstage fieldwork – meaning that both indicators 10.10 and 10.11 were met. A strategy for data validation needed to be in place to assess the quality of the data, and ultimately the acceptability of cases as valid.

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accessibility	10.14	Results from the data validation approach are integrated in the weekly fieldwork reports.	To assess to what extent the target sample sizes are being achieved, it is necessary to know what proportion of cases would be excluded due to quality concerns.	YES	NO	Weekly reporting

Indicator 10.14 stipulates that the results from the data validation approach are integrated in the weekly fieldwork monitoring reports enabling the proportion of cases to be excluded because of quality concerns to be assessed and to determine the extent to which target sample sizes are being achieved. Ipsos and Eurofound/Cedefop were unable to finalise the data validation approach in time to include it in the weekly fieldwork reports. Therefore, the projections on the extent to which cases would need to be removed due to quality issues were largely based on a strategy that was developed as part of the pilot. This was run on the interim datasets consisting of 10% and 50% of the data respectively and thresholds estimated of the number of cases that would need to be removed in the end.

Ipsos recommends investing time in developing a quality score to identify low quality completes. A set-up where a quality score is defined based on the pilot data seems ideal, so that the number of low-quality records can be monitored more closely during fieldwork. Pilot samples in ECS 2019, however, were too small to define the quality score, because setting cut-off values is data driven. Ipsos therefore recommends finalising the cut-off values for the quality indicators based on a preliminary data set from mainstage.

Targets

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accuracy	10.12	The approach to data validation is applied in real time during data collection.	To assess to what extent the target sample sizes are being achieved, it is necessary to know what proportion of cases would be excluded due to quality concerns.	YES	YES	Weekly reporting; Dataset
Accuracy	10.13	Percentage of issues identified based on data validation for which a solution is provided.	Part of the data validation approach is to identify issues with data quality that could be addressed with interventions. In that case solutions need to be provided as soon as possible.	100%	100%	Weekly reporting; Dataset

Indicators 10.12 and 10.13 both measure accuracy and relate to the data validation approach – whether it is applied in real time during data collection (10.12) and the percentage of issues identified for which a solution is provided (10.13). The targets for these indicators were met by Ipsos.

On reflection, the indicator and target for 10.13 does not seem to be entirely appropriate in the context of CAWI interviewing. There are not many issues that could be addressed, and, even if they could be, it might not be desirable to adjust the script mid-fieldwork. It would be worth considering re-defining this indicator in future waves of the ECS as discussed in Section 9 of this report.

6 Recoding

The quality indicators and targets relating to recoding focus on the strategy that was developed as well as quality control measures for the ECS 2019. Further details of these areas can be found in the Coding report²⁴ or Technical and fieldwork report.²⁵

6.1 Recoding strategy

Requirements

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accuracy	11.1	A systematic approach to coding multi-format items into a single format and for collapsing complex/long questions is in place.	Where coding needs to be applied a strategy for coding needs to be in place in advance of data collection, as items need to be designed with the ultimate coded result in mind.	YES	YES	Code; Technical and fieldwork report

A recoding strategy was developed by Ipsos and submitted as part of the preparation contract in advance of data collection. The strategy provides a systematic approach to coding multi-format items into a single format and for collapsing complex/long questions enhancing the accuracy of the data collected as required by indicator 11.1.

6.2 Recoding quality control

Requirements

Quality dimension	#	Indicator	Rationale	Targ	Ach.	Evidence
Coherence and comparability	11.2	Variables with country-specific answering categories are captured in clearly coded cross-country variables.	Good survey practice requires that all variables are consistently coded and labelled to enable comparability.	YES	YES	Code; Coding report; Dataset

²⁴ Eurofound and Cedefop (2020), *European Company Survey 2019: Coding report*, European Company Survey 2019 series, Eurofound working paper, Dublin (<https://www.eurofound.europa.eu/publications/report/2020/european-company-survey-2019-workplace-practices-unlocking-employee-potential#wp-101983>).

²⁵ Eurofound and Cedefop (2020), *European Company Survey 2019: Technical and fieldwork report*, European Company Survey 2019 series, Eurofound working paper, Dublin (<https://www.eurofound.europa.eu/publications/report/2020/european-company-survey-2019-workplace-practices-unlocking-employee-potential#wp-101978>).

A data map was developed by Ipsos incorporating clearly coded cross-country variables for variables with country-specific answering categories to enable comparability across countries (as required by indicator 11.2). The data map was produced prior to the pilot and updated prior to mainstage fieldwork.

Quality dimension	#	Indicator	Rationale	Targ	Ach.	Evidence
Accessibility	11.3	Comprehensive coding report provided, including code (e.g. SPSS syntax) for any automated transformations.	Transparency relies on systematic documentation of all phases of the survey preparation and implementation and avoidance of processing errors.	YES	YES	Coding report

A comprehensive coding report was produced by Ipsos including statistical code/syntax for the automated transformations as required by indicator 11.3. This report was provided to Eurofound and Cedefop for transparency and systematic documentation to avoid processing errors.

7 Data and reporting

The quality indicators and targets relating to data and reporting focus explicitly on micro data and the documentation Ipsos is required to provide focusing on each stage of the ECS 2019. Further details can be found in the Technical and fieldwork report²⁶ as well as the Data editing report.²⁷

7.1 Micro dataset

Requirements

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Punctuality	12.1	A template for the final dataset is agreed prior to fieldwork.	A clear data template is required to ensure appropriate scripting and coding and to ensure no variables are missed or are unusable.	YES	YES	Email exchanged with client; Dataset template

²⁶ Eurofound and Cedefop (2020), *European Company Survey 2019: Technical and fieldwork report*, European Company Survey 2019 series, Eurofound working paper, Dublin (<https://www.eurofound.europa.eu/publications/report/2020/european-company-survey-2019-workplace-practices-unlocking-employee-potential#wp-101978>).

²⁷ Eurofound and Cedefop (2020), *European Company Survey 2019: Data editing report*, European Company Survey 2019 series, Eurofound working paper, Dublin (<https://www.eurofound.europa.eu/publications/report/2020/european-company-survey-2019-workplace-practices-unlocking-employee-potential#wp-101979>).

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accuracy	12.2	Percentage of variables in the dataset that are named and labelled in accordance with the agreed template.	The delivered dataset needs to match the agreed data map.	100%	100%	Dataset
Accuracy	12.3	Percentage of variables for which the metadata (e.g. missing values, measurement level) are properly defined in the dataset.		100%	100%	Dataset
Accuracy	12.4	Percentage of substantive variables included in the dataset.	Datasets need to be comprehensive.	100%	100%	Dataset
Accuracy	12.5	Percentage of auxiliary variables (contact data, paradata, frame data etc.) included in the dataset.		100%	100%	Dataset
Accessibility	12.6	Dataset delivered in specified format.	Dataset needs to be in the appropriate format, to be useful for the client.	YES	YES	Dataset
Accessibility	12.8	Dataset made available for external users.	Transparency and relevance of the study rest on the availability of the resulting datasets.	YES	YES	Dataset available through UK Data Service

Indicator 12.1-12.8 measure requirements relating to the micro dataset. The targets for indicators 12.1 and 12.2 were met. For 12.1 a draft template for the final dataset was agreed prior to mainstage fieldwork ensuring that scripting and coding could be prepared to match it and to ensure no variables were missed or specified in a way to be unusable. This early development allowed for further tweaking throughout fieldwork and was finalised in May 2019, prior to final data delivery. For 12.2, 100% of variables in the dataset are named and labelled in accordance with the agreed template. The metadata (missing values, measurement level) are properly defined in the management and employee representative datasets for 100% of variables – meeting indicator 12.3. Similarly, all substantive variables (100%) are included in these datasets - meeting indicator 12.4 and all (100%) auxiliary variables (contact data, paradata, frame data etc.) are also included – meeting indicator 12.5. For 12.6, Ipsos did deliver the dataset in the specified format to Eurofound and Cedefop enabling analyses to be undertaken - meeting this indicator. Indicator 12.8 is also met. Eurofound and Cedefop have made the management dataset and the employee representative datasets available through the UK Data Service in October 2020. The screener dataset and, contact dataset will be made available on request.

7.2 Reporting

Requirements

Quality dimension	#	Indicator	Rationale	Targ.	Ach.	Evidence
Accessibility	13.1	Comprehensive reports following an agreed format are provided for all stages of the survey process (see Terms of reference for requirements).	Transparency relies on systematic documentation of all phases of the survey preparation and implementation.	YES	YES	See list of deliverables

The target for indicator 13.1 was met. Comprehensive reports following an agreed format were provided by Ipsos for all stages of the survey process as stipulated in the Terms of reference. This ensured transparency of the 2019 ECS through documentation of the project from preparation through to implementation.

8 Punctuality

All the indicators in this section are intended to measure the quality dimension 'punctuality' in terms of completing a task by an agreed date. The rationale behind these indicators is that good project management relies on good time keeping. Timely delivery makes the indicators more useful and reduces the impact in case of any problems.

Requirements

Theme	#	Indicator	Target	Achieved	Evidence
Gross sample	1.25	Gross sample for the Pilot provided to national agencies at agreed date.	YES	YES	Table detailing dates of gross sample sent to/ receipt by national teams
Gross sample	1.26	Gross sample for the mainstage provided to national agencies at agreed date.	YES	YES	Table detailing dates of gross sample sent to/ receipt by national teams
Sampling strategy	1.2	Final sampling strategy delivered to client at the agreed date.	YES	YES	Email to client submitting final sampling strategy
Sampling plan	1.22	Final Pilot sampling plans delivered to client at agreed date.	YES	YES	Email(s) to client submitting final country-level sampling plans
Sampling plan	1.23	Final mainstage sampling plans delivered to client at agreed date.	YES	YES	Email(s) to client submitting final country-level sampling plans
Sampling plan	1.24	Sampling plans approved by client at agreed date.	YES	YES	Email(s) from client approving final country-level sampling plans
Weighting strategy	2.2	Final mainstage weighting strategy delivered to client at agreed date.	YES	YES	Email to client delivering weighting strategy

Theme	#	Indicator	Target	Achieved	Evidence
Design weight	2.7	Design weights delivered to client at agreed date.	YES	NO	Email to client delivering final dataset
Post-stratification weight	2.12	Post-stratification weights delivered to client at agreed date.	YES	NO	Email to client delivering final dataset
Supra-national weights	2.16	Supra-national weights delivered to client at agreed date.	YES	NO	Email to client delivering final dataset

The punctuality targets and indicators relevant to the sampling plan and weighting strategy were met. However, those relevant to the timely delivery of the weights (indicators 2.7, 2.12 and 2.16) were not met. This was because it took longer than anticipated to agree the weighting approach and to subsequently finalise the weights. The knock-on impact was that the final weighted dataset (indicator 12.7) was not delivered at the agreed date.

For indicator 1.23, the Ipsos sampling team determined the need to review the sampling plans in four countries prior to the mainstage. Ipsos suggested providing these by 8 February 2019 and duly delivered revised plans for BE, CY, ES and IT on that date. Related to this Eurofound and Cedefop agreed to the changes in the plans prior to the sampling plans being updated by Ipsos (indicator 1.24). For 2.16, the result was met through the timely delivery of the calibration weights.

Theme	#	Indicator	Target	Achieved	Evidence
Questionnaire development	3.3	Timeline for questionnaire development is defined and kept.	YES	YES	Email providing final draft questionnaire to Ipsos (15/12/2017)
Advance translation	3.8	Advance translation delivered to Ipsos at agreed date.	YES	NO	Email from client sending results of advance translation to Ipsos

A timeline for questionnaire development was defined and kept by Eurofound and Cedefop (indicator 3.3) with the final draft version of the questionnaire being delivered to Ipsos on 15 December 2017. Whilst the advance translation was not delivered to Ipsos at the agreed date (indicator 3.8) the findings from the advance translation were discussed in a meeting held on 29 March 2018 and were used to update the questionnaire and add translator instructions. The updated questionnaires were also shared with Ipsos on 9 April 2018, however the report detailing the results of the advance translation was not sent to Ipsos until June 2018. Although indicator 3.8 was not met, the results and relevant information from the advance translation was shared with Ipsos in advance of delivery meaning that the quality of the questionnaires was not compromised by the delay.

Theme	#	Indicator	Target	Achieved	Evidence
Cognitive test	3.11	Final strategy for cognitive testing is delivered to client at the agreed date.	YES	YES	Email to client delivering cognitive test strategy

Theme	#	Indicator	Target	Achieved	Evidence
Cognitive test	3.14	Cognitive test results delivered at the agreed date.	YES	YES	Email to client delivering cognitive interview report
Selecting of questions eligible for translation	4.1	Existing translations of relevant questions delivered at agreed date.	YES	N/A	Email providing existing translations to Ipsos
Initial translation	4.8	Initial translations delivered to client at agreed date.	YES	N/A	Emails to client delivering initial translations

The final strategy for cognitive testing (indicator 3.11) was delivered a day late but this was due to problems with email sending/receipt from the Eurofound email address. Eurofound, Cedefop and Ipsos agreed to classify this target as 'met' during a meeting in Paris (January 2018).

Eurofound reviewed indicator 4.1 in June 2018, determining that it was no longer relevant to the ECS 2019 since the existing translations were available on the Eurofound website, which is where Ipsos collected them from.²⁸

For Indicator 4.8 - the delivery of the initial translations from the countries (for Translator 1 and Translator 2 only) was not stated in the agreed project timetable so these were not delivered to Eurofound and Cedefop for the ECS 2019.

Theme	#	Indicator	Target	Achieved	Evidence
Within country adjudication	4.10	Within country adjudication (overall) delivered at agreed date.	YES	NO	Emails to client delivering within county adjudication outcomes
Cross country adjudication	4.13	Final cross-country review (overall) delivered to client at agreed date.	YES	NO	Emails to client delivering outcomes of cross-country reviews
Edit final translated questionnaires	4.15	Final translated questionnaires (language version) delivered to client at agreed date.	YES	NO	Emails to client delivering final translations

The within country adjudications for Luxembourg German, Montenegrin and Serbian were not delivered to Eurofound and Cedefop by 22 August 2018 (the agreed date) (indicator 4.10).

For Luxembourg German, it was not possible to hold the review meeting on adaptation before the second half of August due to the limited availability of the linguist. For Montenegrin and Serbian, Eurofound and Cedefop requested that an additional cross-national review meeting should be organised with Croatia to ensure harmonization across all languages and enhance the quality of the

²⁸ <https://www.eurofound.europa.eu/surveys/european-company-surveys/european-company-survey-2013/ecs-2013-questionnaire/ecs-2013-questionnaire-translation>

translations. This session was not foreseen in the original timetable or costing. The session took place at the end of August with the final translation made available at the beginning of September for final proofreading. (The first version of the Croatian adjudication was delivered by 22 August 2018 to Eurofound and Cedefop as planned.)

The outcomes from the French, Dutch and German cross-country reviews were delivered to Eurofound and Cedefop on the agreed date (these were included in the country translation file delivered by 22 August 2018, referenced in indicator 4.10). However, the cross-country review meeting for Croatia, Serbia and Montenegro was only organised in the last week of August with the country translation being finalised in the first week of September for final proofreading by the project manager. This meant it was not possible to deliver the final review by the original agreed date (indicator 4.13). The final translated questionnaires (language version) were not delivered to Eurofound and Cedefop at the agreed date (indicator 4.15). The agreed delivery date was 7 January 2019 for delivering the post-pilot updated questionnaire translations, however it was not possible to deliver all versions on this date. 82% were delivered on 7 January; 88% by 9 January and 100% by 22 January. The delays can be attributed to the limited availability of linguists over the Christmas period.

The timely delivery of the within country adjudication, cross-country review (overall) and the final translated questionnaires for these countries may have allowed Eurofound and Cedefop more time to check and (where applicable) sign off the materials. However, all the translations were finalised in time for fieldwork so there was no detrimental impact on the quality of the translations, or the timing of subsequent activities created by these small delays for some language versions.

Theme	#	Indicator	Target	Achieved	Evidence
Scripting	6.3	Scripting finalized and sent to client at agreed date.	YES	YES - Pilot and mainstage	Email to client delivering final script
Meeting(s) of local project leaders and country coordinators	7.2	Meeting(s) of local project leaders and country coordinators held at the agreed date.	YES	YES - Pilot and mainstage	Meeting agenda(s)
Interviewer training materials	7.5	Final translated interviewer training materials delivered to client at agreed date.	YES	NO	Emails to client delivering interviewer training materials
Fieldwork finalisation	9.6	Percentage of countries where fieldwork is finalised on the agreed date.	100%	71%	Technical and fieldwork report; Weekly reporting; Contact data
Fieldwork completion	10.9	Number of days that fieldwork continues after the agreed date.	0	6 calendar days (4 working days)	Technical and fieldwork report; Weekly reporting

The final scripts for both the pilot and mainstage fieldwork were finalised and sent to Eurofound and Cedefop at the agreed date (indicator 6.3). Similarly, the meetings(s) of local project leaders and country coordinators were held at the agreed dates for both the pilot and the mainstage (indicator 7.2).

The final translated interviewer training materials were not delivered to Eurofound and Cedefop at the agreed date (indicator 7.5). For the Pilot – Ipsos conveyed feedback on the materials to the countries and these were revised accordingly before the training was conducted. The revised versions were not delivered to Ipsos before fieldwork started thus it was not possible to deliver all documents to Eurofound and Cedefop by the agreed date. For the mainstage, 25 out of 30 manuals were sent to Eurofound and Cedefop before the meeting with local project leaders (agreed date) with the remaining five sent in the week after this meeting. Near final versions were generated by the local partners before the meeting and all final versions were sent to Eurofound and Cedefop in due time before mainstage fieldwork started so these delays had no detrimental impact on the quality of the briefings or the resulting fieldwork.

Ipsos closed the CATI screener on 14 June 2019. However, eight countries (BG, CZ, FR, GR, PL, SK, ES and RO) continued fieldwork for a further four working days after that date (indicator 10.9) meaning that only 71% of countries finalised fieldwork on the agreed date (indicator 9.6). In the additional three days, these countries did manage to complete extra interviews so whilst this delay was in breach of the punctuality quality target, it did help Ipsos come closer to the target number of interviews (indicators 10.15 and 10.16).

Theme	#	Indicator	Target	Achieved	Evidence
Dataset	12.7	Substantive dataset delivered at agreed date.	YES	NO	Email to client delivering final substantive dataset
Reporting	13.2	All reports are delivered at the agreed dates.	YES	NO	Emails to client delivering agreed deliverables

The final, weighted substantive dataset (indicator 12.7) was not delivered at the agreed date as the weighting took longer than anticipated to finalise. To try to mitigate the delays to subsequent analysis, Ipsos delivered unweighted datasets followed by the weights.

Not all of the draft reports (indicator 13.2) were delivered to Eurofound and Cedefop at the agreed dates. The delivery of the draft Sampling and weighting report and final Technical and fieldwork report were delayed due to earlier delays in finalising the project and agreeing the weighting approach.

The delays have created difficulties for Eurofound and Cedefop internally by putting pressure on the team to review and approve deliverables outside of pre-allocated time to do so leading to extra time pressures and the potential risk of not being able to ensure that the reports are of the highest quality. In addition, the delays in delivery had a knock-on effect on the preparation of the data and documentation for external users, therefore Eurofound and Cedefop were obliged to amend the timetable for the IZA conference, which had a negative impact on the time available to researchers

for submission of abstracts and their review and preparation of the conference. Furthermore, the delays in data delivery and delivery of the fieldwork reports have reduced the amount of time available to Eurofound and Cedefop for data analysis and report writing to meet pre-agreed deadlines for dissemination. This means that there is an increased risk of error and consequently in quality loss in terms of the accuracy of the results of analysis as any postponement of the publication date for Eurofound/Cedefop reports would compromise their impact.

9 Conclusions

9.1 Reflections on quality control results in the 2019 ECS

This section contains Ipsos's reflections on the results from the quality control activities undertaken in the 2019 ECS. It covers the results overall, those for the targets that were set as requirements as well as those that were defined as real-world targets.

Overall results

In total, 132 quality indicators were specified for the 2019 ECS but during implementation, three indicators were dropped. For two further indicators, the results were either 'Not applicable' or 'TBC' by the end of the project. These five indicators have been excluded from the overall results. Of the remaining 127, 95 were 'requirements' and a further 32 were 'real world targets'.

Table 1.7: Overall quality control targets and achievements

Category	Description	Targets	Achieved	%
Requirements	Targets that have to be achieved	95	72	76%
Real-world targets	Targets that are expected to be achieved. If they are not met, an explanation for not reaching them has to be provided	32	17	53%
Total		127	89	70%

In total, over three quarters of the requirements were met (76%); over half of the real-world targets were met (53%) and 70% of all targets were met (Table 1.7).²⁹

This section discusses the targets that were missed, how close each party came to meeting them and the extent to which missing the targets had serious implications for the quality of the survey.

Requirements met and missed

Quality assurance targets were set for different activities within each stage of the survey life cycle. For some aspects of the survey, only requirements were set (fieldwork training, pilot, data processing micro data and reporting) and for others (fieldwork infrastructure) only real-world targets

²⁹ Eurofound and Cedefop missed 23% (3 out of 13 targets) and Ipsos missed 31% (35 out of 114 targets) respectively.

were defined. Table 1.8 summarises the requirements for each stage and the extent to which the targets were achieved.

Sampling requirements: Almost all the sampling requirement targets were met. The only one that was not met was 1.29 – where the achieved sample sizes were smaller than those planned in nine countries. As discussed earlier in the report, the main impact of not achieving or exceeding the planned sample size is the loss of precision of estimates based on the achieved sample.

Table 1.8: Required targets (by survey life cycle activity) and achievements

	Requirements		
	Targets	Achieved	%
Sampling	14	13	93%
Weighting	14	11	79%
Questionnaire	12	10	83%
Translation³⁰	13	7	64%
Scripting	4	4	100%
Fieldwork training	7	5	71%
Pilot	5	2	40%
(Mainstage) Fieldwork	4	3	75%
(Mainstage) Fieldwork monitoring and data validation	11	6	55%
Data Processing	3	3	100%
Micro data	8	7	88%
Reporting	2	1	50%

Weighting requirements: Out of the 14 requirements, three were not met. These were all punctuality targets relevant to the timely delivery of the weights (indicators 2.7, 2.12 and 2.16), which were not met. It took longer than anticipated to agree the weighting approach and to finalise the weights themselves. As a result, the final, weighted, substantive dataset (indicator 12.7) was not delivered at the agreed date, which had further consequences for the timing of the dissemination activities planned by Eurofound and Cedefop.

Questionnaire requirements: Just over 80% of the requirements were met. One of the two requirements that were not met, was indicator 3.8. Although this was not met, the results and relevant information from the advance translation was shared with Ipsos in advance of delivery meaning that the quality of the questionnaires was not compromised by the delay. The other requirement that was not met was 3.13 relating to the proportion of countries in which the selection

³⁰ For two of the translation requirements, the results were 'Not applicable' so a base of 11 has been used to calculate the percentage of targets achieved.

of respondents and composition of the sample corresponded with those agreed. If this indicator was re-defined in future to refer to the proportion of quotas rather than the proportion of countries that was met, then it might be possible for this to be met in future. In ECS 2019, despite the deviations from the agreed quotas, the composition of the achieved sample meant that a range of respondents were recruited across almost all the cells and that the questions were tested on those with characteristics closely resembling the target population, thus Ipsos were confident that missing this target did not impact the quality of the testing process itself.

Translation requirements: Seven out of 11 of the requirements were met. One of those missed relating to making translation materials publicly available (4.4), could be met in future if the indicator is re-defined or re-phrased to better match current practice or if current practice is improved by Eurofound and Cedefop. The remaining three requirements that were not met (4.10, 4.13 and 4.15) all relate to punctuality targets. Whilst the relevant work referred to in these indicators was not completed by the agreed date(s), it was completed following the agreed translation procedures (intended to ensure high quality translations) so Ipsos did not think that these delays had an impact on the quality of the translation process or the related outputs.

Scripting requirements: All the required targets were met.

Fieldwork training requirements: Five out of the seven requirements were met. Ipsos was unable to meet indicator 7.5 as it was not possible to deliver all translated interviewer manuals on time to Eurofound and Cedefop. Twenty five out of 30 manuals were delivered on time, with the remaining five following on within a week. And even though these were late, all packs were delivered prior to fieldwork starting. The other requirement that was not met was 7.1 – as only 78.6% of local project leaders/project coordinators attended the fieldwork manager instruction meeting. However, all those attending training were responsible for managing and monitoring fieldwork on a day-to-day basis and had a key role in the delivery of the ECS 2019. Ipsos did not think that the quality of the fieldwork was adversely affected by these other team members attending the training compared to the project leader/country coordinator since those that were trained were responsible for the fieldwork.

Pilot requirements: Two of the five requirements set for the Pilot were met. One of those not met (8.3) could have been regarded as met since Ipsos did complete interviews in all the local languages covered by the Pilot. However, once the target numbers specified in the clarification notes for this indicator were considered, it became clear that the target was not met. As discussed elsewhere in this report this had an impact on the quality of the pilot findings and Eurofound, Cedefop and Ipsos's ability to make decisions for the mainstage based on the pilot results. Furthermore, indicators 8.1 and 8.2 were also missed with the number of achieved interviews falling short of target for both the employee representative respondents and the management respondents. By missing the target number of interviewers the ability to assess the functioning of the questionnaire in each language and to test the gross sample requirements were limited. However, the ability to test the contact strategies were not compromised. (Mainstage) Fieldwork requirements: Only one of the requirements relating to this aspect of the survey was not met. Indicator 9.6 related to the timely completion of fieldwork. Whilst this was not met in all countries, in the additional three days, extra

interviews were completed. So, whilst this delay was in breach of the punctuality quality target, it did help Ipsos come closer to the target number of interviews overall.

(Mainstage) Fieldwork monitoring and data validation requirements: Seven of the 11 targets were met. Of the five missed, three were arguably close to being met. For 10.6 – only the omission of data for one country from the combined monitoring data meant that the target was missed; for 10.7, the monitoring data was delivered late on two occasions without prior agreement and for 10.9, fieldwork continued for three working days beyond the agreed date. For each of these the quality of the fieldwork monitoring data was not compromised and in the case of 10.9 may have been enhanced by the delay since additional interviews were completed in the time period. Indicators 10.14 and 10.15 were also missed. For indicator 10.14, Ipsos and Eurofound/Cedefop were unable to finalise the data validation approach in time to include it in the weekly fieldwork reports. Therefore, the projections on the extent to which cases would need to be removed due to quality issues were largely based on a strategy that was developed as part of the pilot. Ipsos recommends investing time in developing a quality score to identify low quality completes – ideally based on pilot data but if these are too small then based on a preliminary data set from mainstage (e.g. the 50% MM and ER datasets). For indicator 10.15, five countries did not achieve the 90% target. In DE, PL and RO, samples of sufficient size were collected, and the distributions were of acceptable quality. In CY and MT, too few interviews were completed to facilitate proper analyses.

Data Processing requirements: All three targets were met.

Micro data requirements: Only one of the targets was not met. This related to a punctuality indicator ‘Substantive dataset delivered at agreed date’ as this dataset was not delivered at the agreed date because the weighting took longer than anticipated to finalise. To try to mitigate the delays to subsequent analysis, Ipsos delivered unweighted datasets followed by the weights. However, this delay still had an impact on Eurofound and Cedefop internally and on external users who were unable to access the data in a timely fashion for the IZA conference.

Reporting requirements: One of the two targets was not met. Nearly all the draft reports (indicator 13.2) were not delivered to Eurofound and Cedefop at the agreed dates (the only exceptions were the sampling and quality control reports). The delivery of the draft Sampling and weighting report and final Technical and fieldwork report was delayed due to earlier delays in finalising the project and in agreeing the weighting approach.

Real-world targets met and missed

Table 1.9 summarises the real-world targets for each stage and the extent to which these were achieved.

Sampling real-world targets: Six of the 15 sampling targets were met. The nine that were not met related to the sampling frames (1.3, 1.4, 1.10), the reference statistics (1.12, 1.13, 1.14, 1.18), the sampling plan (1.20) and the net sample size (1.27).

Table 1.9: Real world targets (by survey life cycle activity) and achievements

	Real-world targets		
	Targets	Achieved	%
Sampling	15	6	40%
Weighting	3	2	50%
Questionnaire	2	1	50%
Translation	2	2	100%
Fieldwork infrastructure	1	0	0%
Scripting	2	1	50%
(Mainstage) Fieldwork	2	1	50%
(Mainstage) Fieldwork monitoring and data validation	5	4	80%

Due to access restrictions, fewer countries than anticipated could access an establishment level frame of sufficient quality to meet Eurofound and Cedefop's requirements (1.3) and those that were accessible offered lower population coverage than desired, increasing the potential for coverage error (1.4) and offered different size class categories in four countries (1.10, 1.18, 1.20), which may have led to a reduction in the cross-national comparability of the data for these countries against the remaining 24 countries.

Fewer countries than anticipated could access reference statistics at the establishment level (1.12), which was not ideal since establishment level statistics are required for weighting however where this was not available it was estimated based on the survey data. Furthermore, in only 87% of countries, the reference statistics fully covered the population (1.13), increasing the risk of generating coverage errors in the sampling design and weighting phases. The statistics were also from 2017 in two countries (1.14) but the impact on quality is expected to be minimal given that population distributions change slowly over time.

Finally, the deviations in the proportional size of each strata between the achieved sample (net) and the targets for some of the strata exceeded 5 percentage points in 13 countries. This meant that the unweighted distributions did not match the sampling targets.

Weighting real-world targets: One of the targets was met. Although 2.13 was not met, this was close to being met with a result of 93% out of 100 recorded. This reflected the fact that for two of the 28 mainstage countries (Greece and Luxembourg), the reference statistics used in the weighting were not based on data that was within two years preceding fieldwork since no recent reference statistics were available.

Questionnaire real-world targets: Half of these targets were met. The one that was missed (3.9) was because four questions from the cognitive test where 'major' issues were detected were retained. Two were replaced by questions that were fielded in the previous wave of the ECS in 2013; for the

remaining two, alternative questions were formulated that were considered less complex but were included in the ECS 2019 untested, which was not ideal but could not be avoided.

Translation real-world targets: All these targets were met.

Fieldwork infrastructure real-world targets: Only one real-world target was set (5.1) and this was missed as Ipsos did not have a fully integrated system for sample management. This meant that for the ECS 2019, there were occasional lags in reporting, which meant that there were gaps in the weekly fieldwork monitoring information, although no major problems occurred.

Scripting real-world targets: Two real-world targets were set for scripting and one was achieved. Indicator 6.1 was missed as one scripting error was detected during the pilot test in the routing for the random assignment of scenarios for additional establishments in the screener questionnaire. The scripting error had an impact on the size of the groups assigned to each scenario – but the impact on was minor because few records were added to the samples via this procedure.

(Mainstage) Fieldwork real-world targets: Two real-world targets were set for scripting and one was achieved. Indicator 9.1 was missed – this related to the size of the gross sample used compared to what was planned. Unfortunately, missing this target does have implications for quality of the sample as the cases sampled later were subject to the same contacting rules as those sampled from the outset, creating delays to the end of the fieldwork in some countries.

(Mainstage) Fieldwork monitoring and data validation real-world targets: Almost all these targets were met – only one was missed - indicator 10.3. This related to the number of issues that had not been resolved by the end of the fieldwork period. By the end of fieldwork, almost 88% of issues had been resolved, but four issues were still outstanding. In an ideal world, all issues raised during weekly monitoring would be resolved. However, for at least two of the issues it would have been impossible to resolve these since they related to fundamental issues with the sample frames in two countries. If these two were excluded, then the result would be 94%. The implication of the issues that were not resolved is that the target of 'at least 90% of interview targets to be achieved' in 100% of countries (indicator 10.15) was not met.

9.2 Reflections on the quality control approach

The quality control approach adopted for the ECS 2019 was a reasonably effective and an efficient process. The volume of indicators in total was not ideal but given that not all indicators applied at all stages of the survey this could be managed successfully to avoid over-burden in assessment and reporting.

The decision to deliver the monthly updates/results to Eurofound and Cedefop for indicators and targets in the quality assurance plan (Excel file) rather than in a word document as well reduced the

effort for Ipsos in documenting QA activities for the remainder of the project³¹. Ipsos recommends adopting this approach in future waves of the ECS and are currently implementing this on EWCS7.

Close monthly monitoring of the indicators and targets worked well. Forthcoming targets were reviewed upfront and on occasion, unclear indicators were further defined, and discussions held with Eurofound and Cedefop on the best way to calculate results. Ipsos recommends this approach is followed in future waves and for similar surveys.

Ipsos maintained evidence of where targets were met or missed throughout the project as well as calculations for results where needed - this meant that the process of reviewing and reporting on results at the end of the project ran smoothly. Ipsos would continue this practice for similar projects in future.

Defining indicators and targets

Most indicators and targets were defined clearly and precisely at the kick-off meeting or prior to reporting for that indicator meaning that it was only necessary to change a small number during implementation. Ipsos recommends repeating this process to ensure that all parties understand what the indicator and accompanying target refer to and how success is measured. To enhance this process, Eurofound and Cedefop could provide the rationale for indicators and targets earlier than was the case for ECS 2019.

A mistake was noticed for the target set for indicator 9.2 during reporting. Rather than 100% the target was changed to 0%. Indicators 8.1 and 8.2 relating to the pilot interview targets in the smaller countries were redefined in the quality assurance plan to reflect changes agreed in the Pilot Sampling strategy. Finally, the definition for indicator 7.1 of what the target group being trained was defined differently for the pilot compared to the mainstage. Further details on the amended indicators and targets are provided earlier in this report. Whilst changing any indicator/target once the process has started is not ideal, it is probably unavoidable.

During monthly reporting for the implementation phase of the project it became apparent that some of the targets that had been earmarked for assessment in the mainstage could also be assessed for the pilot. Rather than replicate indicators for the pilot and mainstage (which would have increased the number of indicators to monitor overall) separate results are provided for the pilot and the mainstage (where applicable). Several indicators³² have two results. Ipsos suggests that indicators/targets that are relevant for both the pilot and mainstage are identified at the outset to ensure that these are not overlooked during the reporting of either stage of the project.

³¹ Only the Excel file was provided from November 2018 onwards.

³² Indicators: 1.3, 1.4, 1.6, 1.12, 1.14, 6.3, 6.5, 7.1 and 7.2.

Proposed amendments to specific indicators/targets

As mentioned earlier, two of the requirements related to the translation indicators were regarded as ‘not applicable’ in the final assessment of the targets (4.1 and 4.4). Ipsos recommends considering whether these should both be retained for future waves or only 4.4. Indicator 4.4 could be re-defined or re-phrased to better match current practice. On the other hand, Eurofound and Cedefop could improve on their current practice, by ensuring that all translation materials are available. This would allow users to assess the quality of the data collected in their language, as well as the comparability of the data across countries and languages. In this case, indicator 4.4 could be retained without amendment.

There is one indicator where the result was still ‘TBC’ by the end of the project (indicator 12.8, relevant to the micro data – Dataset made available for external users). For ease of completing the Quality Control report by the end of the project (for all indicators specified) it might be sensible to omit this indicator in future waves. As it is not possible to achieve it by the end of the project it seems incongruous to include it.

Based on the assessments in this report, there are a small number of other amendments that have been suggested to improve specific quality control indicators and targets for future waves of the ECS. All proposed amendments are summarised in Table 1.10.

Table 1.10: Suggested improvements for specific indicators for future waves

Theme and Quality dimension	#	Indicator	Proposed amendment
Weighting - Accuracy	2.3	Percentage of countries where the agreed weighting strategy is used.	Re-consider the wording for the rationale for this indicator, currently worded as ‘Using an equivalent approach for weighting in all countries enhances comparability. The agreed weighting approach will attempt to achieve this equivalence and should therefore be applied in all countries’.
Weighting - Accessibility	2.14	Supra-national weights included in dataset.	Consider re-defining or re-phrasing to focus on gross weights or calibration weights rather than supra-national weights.
Weighting - Accessibility	2.15	Procedure for constructing of, and sources used for, supra-national weights described in weighting report.	Consider re-defining or re-phrasing to focus on gross weights or calibration weights rather than supra-national weights.
Weighting - Punctuality	2.16	Supra-national weights delivered to client at agreed date.	Consider re-defining or re-phrasing to focus on gross weights or calibration weights rather than supra-national weights.
Questionnaire - Accuracy	3.13	Percentage of countries in which the selection of the respondents and composition of the sample corresponds with	Rephrase/re-define indicator to refer to proportion of quota cells met rather than proportion of countries.

Theme and Quality dimension	#	Indicator	Proposed amendment
		the agreed approach and design.	
Mainstage fieldwork & data validation - Accuracy	10.13	Percentage of issues identified based on data validation for which a solution is provided.	Rephrase/re-define indicator and target.
Mainstage fieldwork and data validation - Accuracy	10.15	% of countries where at least 90% of [interview] target was achieved.	Consider moving indicators 10.15 and 10.16 to the sampling section rather than in the fieldwork monitoring section.
	10.16	At least 95% of [interview] target across countries reached.	
Translation - Punctuality	4.1	Existing translations of relevant questions delivered at agreed date.	Delete – result was not applicable in this edition.
Translation - Accessibility	4.4	Translation materials are made publicly available.	Consider re-defining or re-phrasing.
Fieldwork training - Accuracy	7.6	Percentage of interviewers that take part in the training.	Delete – overlaps with 7.7 (Percentage of interviewers that are trained before they participate in fieldwork).

New indicators/targets

During ECS 2019, there were two serious threats to data quality: the translation mistake in Sweden affecting the accuracy of the script overlay process and the sample extraction in Slovenia affecting the representativeness of the sample.

Table 1.11: Suggested new indicators for future waves

Survey area	Theme and quality dimension	#	New indicator wording	Target
Translation	Coherence & comparability	4.16	Number of errors detected in the language overlay process.	0
Sampling	Accuracy	1.30	Percentage of countries where the total number of sampled services at the stratum level AND the distribution across the NACE sectors (the implicit stratification levels) matches the sample specification placed with the sample provider.	100%

Neither of these errors were directly captured by any of the existing indicators in ECS 2019, so Ipsos would like to propose two new indicators to capture checking processes to try to avoid a repetition of such mistakes in future waves of the survey. Given the potentially serious consequences of missing either of these targets, Ipsos suggests that both of these are set as requirements. The two new indicators are shown in Table 1.11.

WPEF20015

The European Centre for the Development of Vocational Training (Cedefop) is the European Union's reference centre for vocational education and training. It provides information on and analyses of vocational education and training systems, policies, research and practice. Cedefop was established in 1975 by Council Regulation (EEC) No 337/75.

The European Foundation for the Improvement of Living and Working Conditions (Eurofound) is a tripartite European Union Agency, whose role is to provide knowledge in the area of social, employment and work-related policies. Eurofound was established in 1975 by Council Regulation (EEC) No. 1365/75, to contribute to the planning and design of better living and working conditions in Europe.