

Jointly produced by
CEDEFOP and
EUROFOUND

Labour market change
European Company Survey 2019:
Sampling and weighting report

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



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: WPEF20014

Related reports: European Company Survey 2019: Technical report; European Company Survey 2019: Translation report; European Company Survey 2019: Quality control 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
Statistical population and sample requirements	1
Sampling frames and reference statistics	2
Sampling frames	2
Reference statistics	6
Assessment of coverage and sampling frame quality	8
Pilot sampling strategy	11
Fieldwork sampling strategy	14
Stratification	14
Size of the gross sample	15
Adjustments to the standard approach	16
Adjustments for stratum jumpers and general sample management	19
Screening strategy for company-level frame countries	21
Weighting	30
Design and non-response weights	30
Calibration weighting (Step 5)	35
Summary and evaluation of weights	36
Screener data weighting	43
Annex	44

Introduction

This document outlines the sampling and weighting approach for the European Company Survey (ECS) 2019. The first part of the report focuses on sampling. This part summarises the sampling approach documented in Ipsos' technical proposal and notes any changes that were agreed subsequently during project set-up. It also describes the pilot sampling approach and summarises changes that were made to the main stage sampling based on the pilot. The report includes information on the development work undertaken for the four IPA countries included in this phase, although subsequently these countries were not included in the main survey. The second part describes the weighting approach implemented in ECS 2019 and summarizes the results from the weighting process.

Statistical population and sample requirements

The universe represented by the survey was the population of establishments employing 10 people or more in each of the countries covered by the survey. Establishments in the NACE rev. 2 categories B to N, R and S were included in the universe, while NACE rev.2 categories A, O, P, Q, T and U were excluded. The unit of enquiry was the establishment. Within each establishment the manager responsible for human resources (MM respondent) was interviewed. In addition, if a formal employee representative existed that person was also interviewed (ER respondent).

Table 1: Reference and final planned sample sizes

	Reference N	Final N		Reference N	Final N
Austria	1,000	1,000	Luxembourg	250	250
Belgium	1,000	1,000	Malta	250	250
Bulgaria	1,000	1,000	Netherlands	1,000	1,000
Croatia	500	500	Poland	1,500	1,150
Cyprus	250	250	Portugal	1,000	1,000
Czechia	1,000	1,000	Romania	1,000	1,000
Denmark	1,000	1,000	Slovakia	500	350
Estonia	500	500	Slovenia	500	500
Finland*	1,000	1,000	Spain	1,500	1,500
France	1,500	1,500	Sweden	1,000	1,000
Germany	1,500	1,000	United Kingdom*	1,500	700
Greece	500	500	Total EU28	24,750	21,900
Hungary	1,000	1,000	North Macedonia	500	-
Ireland	500	250	Montenegro	500	-
Italy	1,500	1,500	Serbia	500	-
Latvia	500	500	Turkey	1,500	-
Lithuania	500	500	Total IPA	3,000	-

* In Finland and United Kingdom smaller targets were initially agreed following the pilot, of 350 and 550 interviews respectively. It was possible to increase the final achieved samples in these countries during the main stage when assumptions proved better than expected. The sampling targets were revised proportionally to accommodate the change.

Table 1 provides the target net sample size in each country, which is the total number of establishments in which an interview with the manager was to be secured. The first column for each country provides the tender reference sample size and the second column the agreed main survey sample sizes. Reductions to the planned sample sizes were agreed in some of the countries on the basis of a feasibility assessment following the pilot. North Macedonia, Montenegro, Serbia and Turkey were also not taken forward in the main survey following the pilot assessment.

Sampling frames and reference statistics

Sampling frames

The following table (Table 2) lists the sampling frames that were proposed and agreed during the inception stages of the project. The first column shows the sampling frame source suggested in the tender specifications, and second column confirms whether this frame was agreed for the survey, and if it was not, provides the name of the alternative that was used. The third column shows the level of the sampling frame (establishment or company).

At the tender stage Ipsos suggested a different frame in five countries (Belgium, Bulgaria, Denmark, Italy and Turkey) to the one initially suggested in the technical specifications, all of which were approved by Eurofound and Cedefop. During the subsequent sampling frame information gathering stages further changes were proposed and agreed in Bulgaria (for a second time), Croatia, France, Hungary, Italy, Latvia, Malta, Poland and Spain. In Poland and Spain this has meant changing to a company-level frame from an establishment-level one. The frame used in Italy is available at the company level for sampling, but records can also be provided of the establishments of each selected company (this is discussed in more detail later in this report) Additionally, in Belgium, it was decided that the sampling frame should be used at the company rather than establishment level. Table 3 provides details of the reasons for these changes.

The choices of sampling frames reflect Ipsos' experience, the advice provided in the feasibility study report on sampling¹ and the subsequent work undertaken to evaluate sampling frame quality in greater detail (reported below). In relation to the feasibility study report, one of the findings was that some of the frames used in the previous survey at establishment level appeared to be at company level in reality, due to the low number of subsidiaries reported on the sampling frame. This applied to four out of the five countries where Bisnode/Dun & Bradstreet was used as a sampling frame at the establishment level. This is reflected in the frame choices in all four countries, where, a different sampling frame to the previous survey was selected in Austria, Germany and Sweden; and the same frame was selected but at the company level in Slovenia.

All of the sampling frames listed in Table 2 were tested during the pilot and all were retained for the main survey in the 28 EU Member States.

Table 2: Sampling frames

	Sampling frame – tender specifications	Sampling frame – confirmed	Frame level
EU28 Member States			
<i>Austria</i>	HEROLD Marketing CD / MDOoffline professional	As specification	Establishment

¹ Eurofound (2017), *Feasibility study regarding methodology, design and mode of the European Company Survey. Task 2: sampling modes and frames*, unpublished.

	Sampling frame – tender specifications	Sampling frame – confirmed	Frame level
<i>Belgium</i>	Infobel	Graydon	Company (change from Ipsos initial proposal)
<i>Bulgaria</i>	Bureau van Dijk (ORBIS databank)	Dun & Bradstreet (centrally-sourced)*	Company
<i>Croatia</i>	Bisnode	Annual Financial Statements Registry (RGFI), Financial Agency (FINA)	Company
<i>Cyprus</i>	Business Register (ΜΗΤΡΩΟ ΕΠΙΧΕΙΡΗΣΕΩΝ) ²	As specification	Company
<i>Czechia</i>	Bisnode Albertina	As specification	Company
<i>Denmark</i>	Experience (KOB)	Bisnode	Establishment
<i>Estonia</i>	e-Business Register	As specification	Company
<i>Finland</i>	Bisnode	As specification	Establishment
<i>France</i>	Cegedim CD	Direct Fichier	Establishment
<i>Germany</i>	Heins & Partner	As specification	Establishment
<i>Greece</i>	ICAP directory	As specification	Company
<i>Hungary</i>	KSH (Central Statistical Office)	Bisnode	Company
<i>Ireland</i>	Bill Moss	As specification	Establishment
<i>Italy</i>	Dun & Bradstreet	Cribis	Company/ Establishment
<i>Latvia</i>	Statistikas uzņēmumu reģistrs/Statistical Enterprise Register	Lursoft	Company
<i>Lithuania</i>	Creditinfo	As specification	Company
<i>Luxembourg</i>	Editus	As specification	Establishment
<i>Malta</i>	Internal database of survey agency	Dun & Bradstreet (centrally-sourced)*	Company
<i>Netherlands</i>	Handelsregister	As specification	Establishment
<i>Poland</i>	Bisnode	Dun & Bradstreet (centrally-sourced)*	Company (change from Ipsos initial proposal)
<i>Portugal</i>	Informa D&B	As specification	Company
<i>Romania</i>	Lista Firmelor din Romania	As specification	Company

² The register provider in Cyprus had stopped providing the size of companies and their phone numbers. Therefore, the sampling frame was based on matching the latest database (which included company name, address and sector for in-scope size/sector businesses) to the 2011 version (the last to include size and phone number). Additional work was then undertaken to look up details for unmatched companies.

	Sampling frame – tender specifications	Sampling frame – confirmed	Frame level
<i>Slovakia</i>	Bisnode Albertina	As specification	Company
<i>Slovenia</i>	Bisnode Slovenija	As specification	Company
<i>Spain</i>	DataCentric	Informa D&B	Company (change from Ipsos initial proposal)
<i>Sweden</i>	Bisnode Sverige (PARAD)	As specification	Establishment
<i>United Kingdom</i>	Experian	As specification	Establishment
IPA countries			
<i>Montenegro</i>	Central Registry of Business Entities	As specification	Company
<i>North Macedonia</i>	Central Registry	As specification	Company
<i>Serbia</i>	Business Entities Register	As specification	Company
<i>Turkey</i>	Dun & Bradstreet	TURKSTAT Business Database	Company

* 'Dun & Bradstreet (centrally-sourced)' refers to data sourced from Dun & Bradstreet by the Ipsos coordination team rather than local teams. Database counts were sourced from this supplier for a number of countries with initially unsatisfactory coverage. It was also sometimes the case that it had different numbers to a locally-sourced Dun & Bradstreet solution (see Malta, Table 3).

Table 3: Reasons for changes in frame source or level

	Reasons for change
<i>Belgium</i>	<p>For Belgium, several frames had overall good coverage; Graydon was not included in the feasibility assessment, but was very similar to the other frames in terms of data source (i.e. how the sampling frame is compiled), frequency of updates, coverage and quality of firmographic information. Ipsos proposed Graydon given experience of the quality of service when working with this provider.</p> <p>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. For this reason, the coverage of Infobel was also checked (the frame used in the previous ECS), but it was found to have worse coverage than Graydon. The feasibility report had also noted issues with using this frame at the establishment level.</p>
<i>Bulgaria</i>	<p>At the proposal stage Ipsos suggested Apis as a sampling frame, for various reasons, however, although this frame appeared a good sampling frame to use, upon making the detailed assessment against population counts it was shown to have poor population coverage. Dun & Bradstreet was proposed as an alternative, with good coverage demonstrated. It is worth noting that this frame had almost exactly the same number of units as Bureau van Dijk, which was suggested in the feasibility study and tender specifications.</p>

	Reasons for change
Croatia	During the initial stages of the project it became clear that there would be practical advantages in the implementation of the sampling to using FINA instead of Bisnode, while the quality of the frames was the same, given they were based on the same source (Bisnode was a commercial frame drawn from FINA data). Specifically, the local agency was not able to purchase the full Bisnode database, unlike with FINA, complicating the sampling and survey process.
Denmark	In Denmark more than one high-quality sampling frame was available: Experian, Soliditet and Bisnode (former NN Markedsdata and other frames). Experian and Bisnode used the same sources and received daily information from a variety of providers – public as well as private (company information: e.g. Erhvervsstyrelsen/Danish Business Authority and Danmarks Statistik/Statistics Denmark). Bisnode was selected given previous experience of the quality of service when using this supplier.
France	Cegedim CD (the tender specification supplier) and Direct Fichier (the chosen supplier) were both commercial providers that drew data from the INSEE Sirene file, the source of business sample in France. Ipsos France had an ongoing relationship with Direct Fichier for the provision of sample, and so this provider was chosen for quality of service. The source of data and quality was the same with either provider given they accessed the same data.
Hungary	During the initial stages of the project it was discovered that CSO (Central Statistical Office) would not be able to provide any phone numbers with the sample. Therefore, Bisnode was proposed, where this was not an issue. The frame was able to deliver good population coverage.
Italy	Dun & Bradstreet had stopped operating directly in Italy, and as such, was less suitable for the current survey. Two suppliers were considered during the preparation of the sampling, Consodata and Cribis, and the latter was found to have better quality in terms of coverage and was also the only provider of the two able to provide an establishment-level sample.
Latvia	Lursoft was proposed by the local agency during the initial stages of the project. The services of Lursoft were known to be cheaper, faster, more client orientated and flexible, and that they can provide better quality phone numbers and additional information not otherwise available (including the number of establishments at the company, useful for screening). Additionally, Lursoft and the National Statistical Bureau of Latvia received their information from the same source – the Company Register, State Revenue Service, so coverage was similarly high.
Malta	At the start of the project the local agency advised that it would be possible (and preferable) to use a commercial database, provided by Dun & Bradstreet, rather than their own internal database, and that this would be available at establishment level. Their internal database had been preferred due to inaccessibility of the official register previously used for the survey. This frame was however assessed and had severe issues, specifically, no size information, and far too many units listed (45,000 against a population of 2,200). Instead, Dun & Bradstreet (provided by Ipsos' central contact) was used given this supplier was able to provide counts (at the company level) with more accurate coverage, and size information for part of the sample.
Poland	Bisnode was initially suggested as a sampling frame, the feasibility report noted that it had better coverage than the frame previously used, PCM (Polskie Centrum Marketingowa), with a total of 150,000 units (including the public sector). However, in practice it proved to have around 84,000 units (excluding the public sector, which will explain some of the difference), and fairly substantial under-coverage particularly of smaller establishments. As this comparison was also with company-level population

	Reasons for change
	counts, it called into question that the sampling frame was at establishment level (a greater number of smaller establishments would be expected than the company-level population). Dun & Bradstreet (provided by the Ipsos central team supplier) was able to provide counts (at the company level) with better coverage, and so this frame was selected.
<i>Spain</i>	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).
<i>Turkey</i>	The Business Database was the most up-to-date frame available at the time of sampling. Based on our assessment, the frame was better than D&B, but the quality below that of the frames in many other countries. For example, only about half of the entries on the frame had an up-to-date phone number. Therefore, the frame was changed for the pilot and subsequently main stage fieldwork did not go ahead in Turkey.

Reference statistics

Table 4 below presents the reference statistics that were used to assess sampling frame coverage and to design and weight the survey, along with the unit of the reference statistics. Having reference statistics at the establishment level was important for the weighting to ensure the survey represents the population accurately. In 17 of the 28 EU Member States, and all four of the IPA countries, the reference statistics were available at the company level only, and so the establishment level needed to be estimated. In addition to this issue, in some countries some of the sectors were not available in the population statistics and so had to be estimated. Specifically, in Croatia population estimates for NACE K, R and S were not available and so were based on the sampling frame numbers; in Ireland some small cells were suppressed in the population data to prevent disclosure and so were also based on frame numbers.

The approach to estimation of establishment-level statistics from the company-level is given below.

1. Population counts were obtained across top-level NACE code and size categories (interlocking cells), using the primary sources of reference statistics given in Table 4.
2. An assessment was made during the preparation phase of the expected difference between the company and establishment level. This was made by comparing countries where both levels were available – Germany and Italy from the providers directly; or via comparisons with establishment-level figures from the frame and Eurostat Structural Business Statistics (SBS) at company level across common sectors³. The conclusion of this exercise was that the differences could be expected to be small, and therefore using the existing level of the reference statistics would be adequate for the purpose of setting sampling targets prior to the survey (see later description).
3. To estimate establishment-level statistics for the weighting, where only company-level statistics were available, the interview screener data was used to extrapolate to the establishment level in each country, after completion of the survey (see weighting section for further detail).

³SBS figures do not cover NACE K, R or S.

Table 4: Reference statistics

	Source of reference statistics	Units
EU28 Member States		
<i>Austria</i>	Official “Arbeitsstättenzählung” from Statistics Austria	Establishments
<i>Belgium</i>	STATBEL	Companies
<i>Bulgaria</i>	Structural Business Statistics (non-financial enterprises/financial enterprises) from the National Statistical Institute	Companies
<i>Croatia</i>	Structural Business Statistics (non-financial enterprises/financial enterprises) of Croatian Bureau of Statistics; and Statistics provided by FINA (Financial Agency)	Companies
<i>Cyprus</i>	Business Register, CYSTAT	Companies
<i>Czechia</i>	Registr ekonomických subjektů (Company register), Czech Statistical Office	Companies
<i>Denmark</i>	Danmarks Statistik (Statistics Denmark)	Establishments
<i>Estonia</i>	Business Register	Companies
<i>Finland</i>	Statistics Finland produces establishment-based statistics (collected from the Tax Administration’s business taxation file)	Establishment
<i>France</i>	La Base Sirene, Insee (National Institute of Statistics and Economic Studies)	Establishments
<i>Germany</i>	Unternehmensregister, DESTATIS	Companies/ Establishments
<i>Greece</i>	Company statistics, Hellenic Statistical Authority	Companies
<i>Hungary</i>	Business Register of the Hungarian Central Statistical Office	Companies
<i>Ireland</i>	Business Register, Central Statistics Office	Establishments
<i>Italy</i>	Structural Business Statistics, Istat	Companies/ Establishments
<i>Latvia</i>	Statistikas uzņēmumu reģistrs/Statistical Enterprise Register of the Central Statistics Bureau	Companies
<i>Lithuania</i>	Register of Legal Entities, Lithuania statistics department	Companies
<i>Luxembourg</i>	Démographie des entreprises STATEC	Companies
<i>Malta</i>	Business Register (BR), National Statistics Office	Companies
<i>Netherlands</i>	Office for National Statistics/Chamber of Commerce	Establishments
<i>Poland</i>	Baza REGON (Główny Urząd Statystyczny/Central Statistical Office)	Companies
<i>Portugal</i>	Sistema de Contas Integradas das Empresas (SCIE - Integrated business accounts system)/Instituto Nacional de Estatística	Companies
<i>Romania</i>	Statistical Yearbook (2016)	Establishments
<i>Slovakia</i>	DATAcube (Statistical office)	Companies
<i>Slovenia</i>	Number of enterprises, Statistical Office of Republic of Slovenia	Companies
<i>Spain</i>	Directorio de Empresas of INE (National Statistical Office)	Companies

	Source of reference statistics	Units
<i>Sweden</i>	Bolagsverket (Office for company registration)/SCB Office for National Statistics (Business Register is a register of all enterprises, government offices, and organisations as well as their workplaces)	Establishments
<i>United Kingdom</i>	Inter-Departmental Business Register (IDBR)/Office for National Statistics (ONS)	Establishments
IPA countries		
<i>Montenegro</i>	Central Registry of Business Entities of the Tax Administration	Companies
<i>North Macedonia</i>	State Statistical Office of Macedonia	Companies
<i>Serbia</i>	Structural Business Statistics (non-financial enterprises), Statistical Office of the Republic of Serbia	Companies
<i>Turkey</i>	Business Registers System, TURKSTAT	Companies

Assessment of coverage and sampling frame quality

During the initial stages of the project the sampling frames underwent a process of evaluation, considering the level of the frame (establishment preferred); the quality of the source and how up to date it was; practical considerations such as accessibility, availability of key variables and arrangements for ordering the sample (lead times, processes for drawing a reserve sample); and most importantly considering the coverage of the frame against the reference statistics (in detail across the sampling cells, both with and without telephone numbers). Table 5 below summarises the key features of the agreed sampling frames. This includes, commentary on the coverage of the sampling frames based on comparisons between the sampling frame counts and those from the reference statistics listed above (column one); the percentage of the relevant frame listings for which a phone number could be provided, based on what was actually delivered for the pilot (column two); whether the frame included the variables necessary for the planned stratification scheme (column three); and a summary of the adjustments that were required to the sampling to deal with the frame issues identified (column four). This final column describes adaptations required in addition to those that may be required to counter phone number coverage and deviations in the stratification scheme, which are covered in separate columns. The following section (sampling strategy) considers these adaptations.

Table 5: Summary of coverage and sampling frames assessment

	Coverage	Phone numbers	Stratification cells	Sampling adjustments
EU28 Member States				
<i>Austria</i>	58% overall, due to undercoverage (UC) at the smaller sizes (<100), believed due to subsidiaries being classed with total company counts.	98%	OK	'No size' stratum required to improve coverage.
<i>Belgium</i>	89% overall, even across sizes, lower coverage (<50%) in NACE M and R	96%	Size class break at 200 instead of 250	-

	Coverage	Phone numbers	Stratification cells	Sampling adjustments
Bulgaria	110% overall, reasonably even across cells. Overcoverage (OC) within acceptable limits.	96%	'City/town' instead of NUTS region	-
Croatia	92%, even across sizes and sectors	92%	OK	Population estimates for NACE K, R, S based on sampling frame
Cyprus	Overall 116% (counts vs total population), coverage within cells to be confirmed as total includes sizeable number where size is not known (but stats authority confirms eligibility)	94% ⁴	LAU instead of NUTS region	'No size' stratum required to improve coverage
Czechia	100%, the sampling frame and register were exactly the same (same source), confirmed via check of separate online sources	96%	OK	-
Denmark	98% overall, excellent across cells	93%	100+ largest size	-
Estonia	108% overall, UC (73%) and sector misclassification in 250+, other cells excellent	99%	OK	-
Finland	98% overall, excellent across cells	68%	OK	-
France	81% overall and similar across sizes, lower sector coverage in NACE D (<50%)	90%	OK	-
Germany	88% overall, some variation by size (79% 100-249 to 109% 500+) and misclassification by sector	100%	OK	-
Greece	45% overall across cases with size, 89% with 'no size' stratum. Worse UC in 10-49 size, some sectors (R, S at 10%, 6%).	100%	100+ largest size	'No size' stratum required to improve coverage
Hungary	105% overall, OC across all sizes (particularly larger companies). Some sector misclassification.	62%	OK	-
Ireland	93% overall, but frame/population at different levels so quantitatively imprecise. Some misclassification across cells.	100%	OK	-

⁴ Of the cases in the 2016 frame that could be matched to the 2011 frame (so that the size could be determined) and were of size 10+.

	Coverage	Phone numbers	Stratification cells	Sampling adjustments
<i>Italy</i>	110% overall, heavy OC in NACE L otherwise consistent (company-level comparison)	87%	OK	-
<i>Latvia</i>	97% overall, excellent across most cells	91%	OK	-
<i>Lithuania</i>	106% overall, some misclassification across cells.	99%	OK	-
<i>Luxembourg</i>	63% (excluding 'no size' cases) or 84% (including 'no size'), based on company-level counts vs company-level reference statistics, worse coverage of smaller companies.	100%	OK	'No size' stratum likely required to improve coverage (based on company-level evaluation)
<i>Malta</i>	108% overall if including 'no size' stratum (otherwise 33%), but with extensive variation between cells, presumably due to differences in classification between sources. Main gaps appear to be in 10-49 size.	100%	'City/town' instead of NUTS region	'No size' stratum required to improve coverage
<i>Netherlands</i>	87% overall (cases with size), worse UC largest establishments (500+ 79%)	95%	OK	-
<i>Poland</i>	123% overall, UC across most cells, more variation across sectors.	80%	'City/town' instead of NUTS region	-
<i>Portugal</i>	102% overall, excellent across cells	90%	OK	-
<i>Romania</i>	93% overall, but frame/statistics at different levels so quantitatively imprecise. Excellent across cells.	94%	OK	-
<i>Slovakia</i>	100%, the sampling frame and register are exactly the same (same source), confirmed via check of separate online sources	91%	OK	-
<i>Slovenia</i>	85% overall, uniform across cells.	94%	OK	-
<i>Spain</i>	97% overall, fairly uniform across cells	90%	Size class break at 200 instead of 250	-
<i>Sweden</i>	100% overall, perfect (same sources for frame and statistics)	98%	Size class break at 200 instead of 250	-
<i>United Kingdom</i>	84% overall, lowest for 10-49 (81%), over 90% for other sizes, some sector misclassification	99%	OK	-

	Coverage	Phone numbers	Stratification cells	Sampling adjustments
IPA countries				
<i>Montenegro</i>	100% overall, perfect (same sources for frame and statistics)	77%	OK	-
<i>North Macedonia</i>	100% overall, perfect (same sources for frame and statistics), however over half of the cases have no size	94%	OK	'No size' stratum required to improve coverage
<i>Serbia</i>	100% overall, perfect (same sources for frame and statistics)	84%	OK	-
<i>Turkey</i>	100% overall, perfect (same sources for frame and statistics)	36%	OK	-

Pilot sampling strategy

The pilot sampling was required to deliver a minimum of 30 MM and 30 ER interviews per country, with the exception of the smaller countries, where lower targets were agreed (15 and 5 in Cyprus, Luxembourg and Montenegro, and 7 and 5 in Malta). The lower target in Malta was due to the very small amount of sample available.

The pilot followed the full main survey random probability sample design, to provide a full dress-rehearsal test of the procedures. In particular, it provided an evaluation of the screening procedures from company-level sampling frames. In addition, it was used to collect key information to inform the main survey design. Including information to assess (i) the yield rate of the sample, including the CAWI conversion rate, and factoring in sample eligibility; (ii) the characteristics of 'no size' cases on the sampling frame in the countries to include such a stratum; (iii) the match-rate and accuracy of looked up phone numbers; and (iv) the accuracy of additional sampling frame information that might be used to supplement the sampling, such as the number of establishments in a company in the company-level frame countries.

To some extent the requirement to achieve an equal number of MM and ER interviews was incompatible with the information requirements listed above (particularly the yield rate), which would be judged best via a sample with the same structure as the main survey. Therefore, as per the initial Ipsos proposal, a sample with the same structure as the planned main survey sample was issued first – i.e. with nine size/sector strata set as the mid-point between establishment and employee-level estimates, and selected using the same procedures including stratification⁵. The size of this sample was based on ratios of cases to achieved MM interviews of 6 or 9 depending on country (180 or 270 cases) in the first instance, and higher factoring in the availability of phone numbers. Targets were not set for achievement of interviews per strata, unlike in the main survey,

⁵ For the pilot sampling it was not possible to obtain employee-level statistics from the LFS in time, given these would need to be obtained locally (the largest size break available centrally from Eurostat for the LFS is 50+), and so the sampling was set using the ESENER-2 targets, given this survey followed the same sampling strategy and included the same strata breaks. The intention in the pilot was to select the sample in similar proportions to the likely main stage sample, to provide the most accurate estimates of the interview yield rate as possible (as the yield rate varies by size and sector the proportions are important).

and rather the sample was worked fully, aiming to achieve as many MM and ER interviews as possible from it. The expectation was that this sample would not quite yield 30 MM interviews, and so reserve sample was also selected.

In the countries that required a 'no size' stratum (Austria, Cyprus, Greece, Luxembourg, Macedonia, Malta and Spain – at the pilot stage) an additional 100 cases were selected, to provide reasonably reliable eligibility and yield estimates of this stratum. In order to compensate agencies for the additional work involved in surveying these cases (given the boost in numbers, and low expected eligibility and yield) the MM target was reduced to 25 in the larger countries – Austria, Greece, Macedonia and Spain.

The pilot sample targets were prepared on the basis of the assumptions stated above, and then this was checked against the available sampling frame, and where necessary pilot numbers were capped at a maximum of 10% of the sampling frame counts (in order to prioritise most of the sample for the main stage of the survey). The sample counts were then doubled (to provide reserve sample) and the orders placed with providers on this basis. In the following countries/cells 20% of the sample was selected for the pilot (including reserve):

- Malta: services 10-49
- Slovakia: production 250+, construction 250+
- Montenegro: all cells
- Macedonia: all 250+ cells
- Serbia: all 250+ cells

Several other countries had cells with more than 10% of the sample selected (including the reserve), primarily in the 250+ and 50-249 cells, namely Bulgaria, Cyprus, Estonia, Finland, Croatia, Hungary, Ireland, Lithuania, Latvia, Portugal and Slovenia; and further cells in Malta, Slovakia, Macedonia and Serbia.

The Ipsos proposal also suggested an even number of the target number of pilot interviews for each language spoken in multi-language countries. In subsequent discussions it was agreed to adjust this to favour the main language in each country, in order to more closely reflect the population, as follows:

- Belgium: French 15, Dutch 15
- Estonia: Estonian 20, Russian 10
- Macedonia: Macedonian 20, Albanian 10
- Latvia: Latvian 20, Russian 10
- Spain: Spanish (Castilian) 20, Catalan 10

The language was determined based on the location of the company where this was feasible.

Pilot sample evaluation

As noted in the pilot report the pilot samples achieved fell short of the targets in most countries, in spite of following the approach set out above and working the sample fully.⁶ The reserve sample was used in many of the countries in an attempt to improve numbers, except for those where it was agreed that this sample should be held back to maximise that available for the main survey.

Table 6 below gives achieved the pilot sample sizes. The target MM interviews were achieved or exceeded in Austria, Croatia, Latvia, Lithuania, Malta, Netherlands, Portugal, Romania, Slovenia, Spain and FYR Macedonia. In relation to the ER interviews, even after using reserve and going over the MM target in many of the countries, it was not possible to hit the planned target number of

⁶ Pilot report is made available on request.

interviews in any countries. This was due both to the low yield rates, which affected both interview types, and a lower ER identification rate than in previous surveys.

Table 6: Pilot achieved samples

	<i>MM online interviews</i>			<i>ER online interviews</i>		
	Main	Reserve	Total	Main	Reserve	Total
<i>Austria</i>	22	12	34	4	2	6
<i>Belgium</i>	15	10	25	4	1	5
<i>Bulgaria</i>	10	17	27	3	2	5
<i>Croatia</i>	25	16	41	2	0	2
<i>Cyprus</i>	4	0	4	0	0	0
<i>Czechia</i>	11		11	0		0
<i>Denmark</i>	17	1	18	5	0	5
<i>Estonia</i>	14	10	24	1	0	1
<i>Finland</i>	8	6	14	4	5	9
<i>France</i>	14	15	29	9	8	17
<i>Germany</i>	6	1	7	3	0	3
<i>Greece</i>	24		24	0		0
<i>Hungary</i>	12	15	27	0	0	0
<i>Ireland</i>	6	6	12	0	0	0
<i>Italy</i>	27		27	5		6
<i>Latvia</i>	22	8	30	0	0	0
<i>Lithuania</i>	31		31	7		7
<i>Luxembourg</i>	11	2	13	0	0	0
<i>Malta</i>	23		23	1		1
<i>Netherlands</i>	31	14	45	10	11	21
<i>Poland</i>	3	0	3	0	0	0
<i>Portugal</i>	24	12	36	1	0	1
<i>Romania</i>	18	13	31	2	3	5
<i>Slovakia</i>	3		3	0		0
<i>Slovenia</i>	40		40	3		3
<i>Spain</i>	6	23	29	2	3	5
<i>Sweden</i>	13	5	18	1	6	7
<i>United Kingdom</i>	5	9	14	0	0	0
<i>North Macedonia</i>	28		28	5		5

	<i>MM online interviews</i>			<i>ER online interviews</i>		
	Main	Reserve	Total	Main	Reserve	Total
<i>Serbia</i>	14		15	3		3
<i>Montenegro</i>	4		4	0		0
<i>Turkey</i>	1		1	0		0

Fieldwork sampling strategy

Stratification

Three explicit sector strata were included in the sampling for ECS 2019: NACE B–E (production), NACE F (construction) and NACE G–S (services). In the previous survey NACE F was grouped with NACE B–E; this change was implemented to control the size of the NACE F sample, given cases in NACE F had a lower response rate than NACE B–E cases in the previous survey. In addition, the three previously-used size class strata (10–49, 50–249 and 250 or more) were used, giving a total of nine explicit strata. In Denmark and Greece, the largest size class available in the reference statistics was 100 or more, meaning the sampling was set across the size strata 10–49, 50–99 and 100 or more. In Belgium, Spain and Sweden the nearest category break to 250 employees was 200, giving size sampling strata of 10–49, 50–199 and 200 or more.

Sampling targets were set across the nine strata as the mid-point between establishment-level and employee-level population estimates. The establishment-level proportions were based on the reference statistics (see earlier). As noted in the applicable section company-level reference statistics were used to set the sampling targets in countries where this was the only level available. The employee-level proportions were based on the Labour Force Survey (LFS), which is available for the countries of the EU⁷. As such, the sampling is disproportional to both the establishment-level and the employee-level population. The targets were then reviewed considering the number of cases available on the sampling frame and predicted yield rates, and adjustments were made by capping cells where the targets could not be delivered (see further detail in the next section).

In addition, a layer of implicit stratification (i.e. systematic sampling to ensure the selected sample is in proportion to the population) was included in the sampling by top level NACE code and then region, to ensure that the sample was selected in proportion to the sampling frame on these factors. The region variable was based on what was available on the sampling frame; for the largest countries ideally being based on NUTS1 or NUTS2, and the smaller countries NUTS3 or smaller (given the smallest countries contain fewer, or just one, of the higher-level NUTS regions). Table 5 highlights some variation in what is available on the sampling frames on region. Importantly, this apparent lack of consistency between countries did not compromise the sample design, given region was placed last in the stratification scheme. A total of 45 strata preceded region, as it followed (i) sector category by size category (9 strata) and (ii) top-level NACE code (15 strata in total distributed across sector groupings). This had two implications. First, this placement meant that region had a lower priority than the other variables, and less of an effect on the stratification, the order of the

⁷ The LFS data available centrally from Eurostat covers only one of the planned size strata: 10–49, with a single category available for 50 or more. The national agencies were asked to seek a more detailed breakdown locally, and for many this confirmed that the LFS does not collect any additional detail. In 11 of the countries, including the largest, Germany, it was possible to obtain a split at the 250 level, either based on establishments (seven countries) or companies (four). For the other countries the Eurostat LFS 50+ figures were apportioned based on the average across those for which the split was available.

strata determining the priority. Second, the placement meant that the sampling could accommodate a variety of different numbers of region categories without affecting the other stratification factors⁸.

Size of the gross sample

At the start of the survey it was necessary to estimate the size of the gross sample that would need to be requested from the sampling frame providers to deliver the planned sample sizes. Typically, one would be able to base the assumptions for these calculations on the outcomes of the previous survey, given important features that influence the response rate are held constant between survey iterations (such as the survey topic, sponsor, informant, questionnaire length, etc.). For ECS 2019 this was less straightforward given the change in methodology. To consider the required size of the gross sample the **screening** and **CAWI conversions** were considered separately, with the former informed by the outcomes from the previous survey, and the current survey pilot, and the latter informed by the current survey pilot only.

At the pilot stage it was considered that the screening conversion rate should be slightly better at the current survey than the overall MM yield rate at the previous survey, given the shorter interview at the screening stage. Therefore, for the screening a ratio of 2:1 was tested in the pilot (issued sample per achieved screening interview) in countries where the previous survey overall sample ratio (issued sample to achieved MM interviews) was better than 3:1 (Croatia, Estonia, Latvia, Luxembourg and Romania), and 3:1 in the other countries. The set of worse-performing countries had ratios of up to 9:1 in the previous survey (the worst being in France). Less was known about what might be achieved for the CAWI conversion ratio (screening interviews to MM interviews). A test of this approach was undertaken in Hungary as part of the previous survey development, which achieved a ratio of 3:1. While it was not expected that all countries would match the outcomes of the test in Hungary in the absence of better information this ratio was trialled in the pilot across all countries. This gave an overall ratio of the pilot start sample of 9:1 or 6:1

Pilot implications and adjustments

The pilot gave steer as to the CATI screener and CAWI yield rates that could be expected in the main survey, which was used to formulate the expected size of the gross sample required in the main survey.

In general, the pilot outcomes were in line with expectations for the CATI screener. CATI yield ratios ranged from 1:1.7 to 1:13.6. The Czechia, Finland, Poland and Slovakia all had yields worse than 1:8, above the worst yields seen in the previous survey. Yields were worse than expected in most countries for the CAWI conversation from successful screener. Of the 28 EU Member States 18 countries had CAWI ratios above 3: Bulgaria, Cyprus, Denmark, Estonia, France, Greece, Germany, Hungary, Ireland, Italy, Lithuania, Luxembourg, Poland, Portugal, Slovenia, Slovakia, Spain and United Kingdom.

The pilot findings made it clear that in the majority of countries it would not be possible to achieve the population mid-point-based sampling targets in all cells, particularly in the large size class strata, given limitations on the amount of sample available. Based on the pilot outcomes, Germany, Greece, the Netherlands and the UK were the only countries expected to be unaffected by this issue. In Cyprus, Malta and Slovakia the pilot suggested that a full census of the sampling frame would be required to deliver the target, meaning the sampling would be in proportion to the sampling frame. The remaining countries all required some reductions in the targets in the largest size strata (and sometimes also the medium size strata) to set achievable targets.

⁸ In order to be fully effective a stratified sample should include sampled cases in all strata. This was assured in the majority of countries at the second level, with 135 strata. At the level of region there were more likely to be some empty strata, depending on the number of region categories (for example with 10 region categories there would be 1,350 strata), reducing the benefit of the stratification at this level.

Adjustments to the standard approach

As noted earlier, adjustments to the sampling strategy were required in a number of countries to maintain quality standards (see Table 5 for summary).

Improving coverage via inclusion of 'no size' stratum

In five EU Member States a 'no size' stratum was required to improve sampling frame coverage (Austria, Cyprus, Greece, Luxembourg and Malta, see Table 5) plus in North Macedonia had it gone ahead. This strategy involved including cases in the survey where there was no size information (number of employees) listed on the sampling frame, given that if these cases included establishments with over 10 employees then coverage would be increased. In all of these countries the coverage evaluation suggested frame coverage of under 80%, or, in the case of North Macedonia, both the sample and reference statistics included a sizeable proportion (over half) of cases without the size (so inclusion of the 'no size' stratum reflected the reference population). It is important to note that including a 'no size' stratum in the survey adds to survey costs (see discussion which follows), and as such the strategy was used selectively (i.e. only where necessary to improve quality and shown to be effective at doing so). The countries with coverage of between 80% and 90% are discussed first, as, ideally, coverage in these countries would also have been boosted.

- In Croatia, France and Slovenia the sampling frame did not contain 'no size' cases. These countries had reported coverage of 80%, 81% and 85% respectively. A further strategy that could be considered would be to augment the sampling frame using other sources, however, this is a complex and potentially expensive procedure. Given that the level was above 80% it was agreed that the survey could proceed in these countries without further adjustment.
- Belgium was very close to 90% (at 89%), and there was the suggestion that coverage may be higher (via population statistics provided by the sampling frame provider, where overall coverage was 94%).
- Germany, similar to Belgium, was close to 90% (at 88%). It did include 'no size' cases on the sampling frame however these were suspiciously small in number, at 10% of the number of 10-49 establishments. In most countries the pattern was of a much larger 'no size' stratum, compared with the other size categories, given the stratum will mostly consist of the smaller categories (below 10 employees) where there are many more establishments. The suspicion was that these cases may have been ineligible for other reasons, such as being historic data for establishments no longer in operation. Although the eligibility of these cases could have been checked in the pilot, given the level of coverage, and the expected low yield rates in Germany, this was not considered necessary or efficient.
- Netherlands too was close to 90% (at 87%), it did include 'no size' cases (8% of the total), however the size class of lowest coverage was the 500+ category (79%), which would be much less likely to be filled from a 'no size' stratum (typically the number of employees is available on the sampling frame for larger companies).
- Similarly, the UK coverage was considered to be sufficient to proceed with the survey, at 84% overall, and given coverage was 95% for establishments of over 50 employees (vs. 81% for the 10-49 stratum). As such, the effective coverage of the survey sample, once over-sampling of the larger establishments was taken into account, was over 90%. The UK also experienced low yield rates in the pilot meaning the adverse impact on efficiency of the inclusion of a 'no size' stratum would not have been affordable.

To implement the 'no size' stratum strategy in the main survey the following steps were considered.

1. The size and structure of the eligible population in the 'no size' stratum was estimated based on pilot outcomes. This was required to estimate the overall coverage of the sampling frame that could be achieved, and then to set an appropriate target based on the actual (eligible cases) size of the stratum. Typically, 'no size' cases will have a much lower eligibility rate (i.e.

the proportion of the contacted sample which consists of an establishment with 10 or more employees operating in NACE B-N, R, S) than the rest of the sample. This is because micro (0-9 employees) and small (10-49) establishments are more difficult to classify accurately, and tend to be those in this stratum, both due to fluctuations in size at the margin of survey eligibility (e.g. an establishment may move above/below the 10 threshold on a seasonal basis) and that the information is less likely to be picked up by the sampling frame provider (e.g. information not publicly available, more likely to be unlisted companies, etc.).

2. The overall target number of interviews that would *ideally* be obtained from this stratum was determined, based on the estimated size of the stratum (from Step 1). For example, if the number of eligible cases in this stratum was estimated to be 10% of the total number of eligible cases (cases with size plus those of no size estimated to be eligible) then in principle around 10% of the interviews should come from this stratum (in practice the unweighted number of interviews would be different to 10% as the overall sample is skewed towards the larger sizes). However, if the eligibility rate in this stratum was shown to be lower than the rest of the sample the cost per interview would be higher, meaning it may not be feasible to include the ideal number of cases. Further, a very low eligibility rate would call into question the feasibility and value of including such a stratum at all.
3. The required number of cases to achieve the agreed target would be sampled, factoring in the expected yield rate. The numbers to select across the sectors (given these cases *are* classified on sector) was based, broadly, on the ratios of the overall sample, but also where feasible taking into account the expected sector/size structure based on the pilot outcomes (at a very broad level given the small pilot sample sizes).
4. The cases would then be treated during fieldwork like cases from any other strata, i.e. they would be used to fill the sampling targets across the nine explicit strata, in addition to monitoring the overall number achieved from this stratum (as per the discussion at step 2).

The situation in **Cyprus** was different to the other countries, given that there were a sizeable number of 'no size' cases but all were confirmed to be eligible (10 or more employees) by the national statistical authority. Instead, the issue was that the statistical authority was not permitted to provide the size to us, and so the size information that was available was based on matching to an old 2011 version of the database. It was also the case that a census was required in Cyprus to reach the planned sample size, and as such all cases were used and the country is not discussed further in this section.

Pilot implications and adjustments

Based on the pilot results it was possible to estimate the size of the population contained in the 'no size' stratum and estimate an adjusted level of coverage for the sampling frame (as outlined in the steps above). The analysis is described in the pilot report and shown in Table 7 below, for the EU Member States.

Table 7: Pilot outcomes and estimated coverage for the no size stratum

	N pilot	Pilot eligibility rate (a)	Frame no size count (b)	Frame no size estimated eligible (a*b)	Frame total estimated eligible	Coverage (without no size)	Coverage (with no size)
Austria	190	19%	169,909	33,038	51,691	58%	93%
Greece	100	27%	12,196	3,344	14,842	45%	53%
Luxembourg	200	37%	3,370	1,237	4,614	63%	106%
Malta	102	68%	1,619	1,099	1,681	33%	77%

In Austria, Luxembourg and Malta the estimates suggest greatly improved survey coverage. In Greece improvements are more modest however as coverage was already low inclusion of the 'no size' stratum was considered worthwhile.

To determine the optimal number of interviews from each stratum one would ideally consider the proportion the 'no size' strata cases comprise of the total population. In other words, the proportion of column four ('frame no size estimated eligible') out of column 5 ('frame total estimated eligible'). This would have suggested that around two-thirds of the interviews should have come from 'no size' cases in Austria and Malta, and a quarter in Greece and Luxembourg. However, the interview yields from 'no size' cases were lower in the pilot than other strata, as expected, and so more costly for fieldwork. This was the case particularly in Austria, with yields over 5 times lower for 'no size' cases compared with other cases. The agreed strategy was therefore to aim for 10% of the interviews from the no-size stratum in Austria, 20% in Greece and Luxembourg, and 50% in Malta⁹. Improving coverage with phone number look-ups

The coverage figures reported were based on a comparison of all of the cases on the sampling frame, including those for which the provider did not hold a phone number. Of course, these cases could not be contacted unless a phone number could be identified. Excluding cases without a phone number from the survey would have reduced survey coverage further, in a way that could be biasing, given it would be reasonable to expect that cases without phone numbers might be different in ways that were related to survey measures. In addition, it is good practice to look up phone numbers during a survey where the initial phone number turns out to be invalid.

Ipsos proposed to use Dun & Bradstreet as an external source for looking up missing/wrong phone numbers, a successful strategy according to our experience from other surveys. The exception to this was countries where Dun & Bradstreet was the source of the sampling frame itself¹⁰. This approach was trialled in the pilot.

The pilot demonstrated that in most countries the proportion of delivered cases with a phone number was high. The rate of phone number inclusion in the pilot was over 90% in all EU Member States except Finland, Hungary, Italy and Poland (see Table 5). It also showed that a high level of case usability¹¹ could be attained following look-ups of missing numbers using Dun & Bradstreet. Here, most countries were above 95%, except for Croatia (92%), Finland (65%), Italy (84%), Poland (81%) and Slovakia (94%) out of the EU Member States. In France, Portugal and Spain the look-up procedures were not tested in the pilot as the sample provided was mistakenly of cases with a phone number only, but similarly high rates were expected. The Dun & Bradstreet provided phone numbers also resulted in an acceptable bad number rate (of 17%, compared with the average of 9%). Dun & Bradstreet was not able to process sample from Cyprus, due to the alphabet used, and the phone numbers had to be looked up manually by the local team.

Based on the pilot this strategy was maintained, supplemented by manual look-ups by the local agencies in countries with lower usability rates.

Working with company-level sampling frames

In addition, an adjustment could be considered for countries using a company-level sampling frame. In these countries some of the (typically larger) multi-site companies would deliver additional (smaller) establishments, which could affect the total number and balance of the establishment interviews. On reflection following the pilot no adjustment was made to the sampling targets in the company-level frame countries. Overall, in the pilot, only 13% of the screened establishments in

⁹Malta subsequently required a census approach to get closer to the target interviews.

¹⁰ This was also considered an issue where Bisnode was used as the sample source, applicable to Denmark and Hungary, given Bisnode and Dun & Bradstreet were understood to be based on the same source.

¹¹Usability rate refers to the proportion of cases that could be issued for fieldwork as they had a phone number that fit the expected format.

these countries identified as multi-site and consisted of two or more establishments with 10 or more employees. Only these companies would therefore provide additional establishments for the survey. Factoring in non-response the additional establishments were not expected to make much difference to the structure of the unweighted sample, and hence adjustments were not necessary.

A further potential variation was considered for company-level sampling frame countries, of adjusting the sampling based on the expected number of establishments in each company. If this information was available on the sampling frame, then it should be possible to deliver a more efficient sample of establishments. For example, if a company consists of five establishments, then a higher probability of selection could be given to this company in the sampling, knowing that this will reduce the size of the weights required to represent the establishments accurately (see further discussion under screening strategy below). A risk with this approach was that if the information on the sampling frame was inaccurate then this could result in a less efficient sample.

To consider this, at the pilot stage information on the type of establishment and numbers of establishments in multi-site companies was requested from all countries, so that sampling frame accuracy could be reviewed post-pilot. Only two countries, Italy and Spain, had information on number of branches that looked to be of any use for the sampling¹². Of these, the establishments could be sampled in advance in Italy (see below), leaving Spain with this potential solution. The correspondence between the number of branches on the sampling frame, and the information collected in the screener, was fairly low, with a correlation of 30% (n=137 cases in Spain completing the screener). Using the distribution of pilot outcomes in Spain different sample designs were simulated allowing the selection probabilities to vary according to the number of branches on the sampling frame, to check the effect on sample precision. The benefit this could achieve is heavily reduced due to the cases where the sampling frame was incorrect and so weight variation is increased instead of decreased. Based on the pilot outcomes, a modest improvement to sample precision of about 5% could be realised, however it would introduce significant complexity at the sampling stage and may not be something the frame provider could actually deliver. Therefore, this approach was not followed in Spain.

In Italy the sampling was done differently in the pilot. Here, the sampling frame provider was able to sample cases at the company level only, but provided an additional file containing all the subsidiaries associated with the selected companies. Prior information was not available on the number of establishments per company, however, the provision of the additional establishments meant that the subsidiaries could be selected in advance, instead of via the screening questions. This process circumvented the need for establishment screening and collection of contact details, a source of non-response in the pilot. To provide a test of procedures, the same sampling rules were applied in Italy as in the other company-level frame countries, of selecting up to 3 establishments per company (in advance of fieldwork). As outlined in the pilot report, the pilot sample in Italy had the highest proportion of subsidiaries of all the company-level sampling frame countries, demonstrating that the approach was effective at bringing in this type of establishment. This approach was therefore used in the main survey in Italy.

Adjustments for stratum jumpers and general sample management

The sampling was designed to achieve target numbers of MM interviews in the sector and size stratification cells. As noted earlier, the sampling targets were set across the nine strata as the mid-point between establishment-level and employee-level population estimates. The issued sample was selected using the sampling frame information for NACE sector and size of company with the aim of achieving the target number of interviews, within a certain tolerance level, for each cell. Ideally this

¹² Belgium, Croatia and Ireland also had the number of branches indicated for a small part of their samples, however this was indicated for between 1% and 7% of the companies, much less than the proportion found in the survey, and so was not considered reliable.

would be achieved by controlling the size of the issued sample in each cell, and our strategy for this is discussed later in this section. However, it is inevitable that some of the targets will be met while there remains issued sample that is still 'live'.

To manage this issue, many telephone surveys of businesses include a 'stopping rule', whereby cells in which the target has been reached are allowed to be closed, and any remaining sample that belongs to that cell according to the sampling frame information is no longer dialled. This was the case for 14% of the gross sample on the previous survey (the same proportion as ESENER-2), and there are reasons to believe this practice could bias the sample, both because of the presence of 'stratum jumpers' (establishments that end up in a different analysis cell to the one in which they were sampled), and because cases that are abandoned after being dialled are more likely to be harder to reach. Stratum jumpers are a problem because they result in some of the cells filling up more quickly than others, i.e. those which more of the stratum jumpers move to, meaning the stopping rule is applied disproportionately. This adjusts the profile of the sample, and there is evidence that establishments which 'jump' could be significantly different to establishments that had the correct classification on the sampling frame, potentially biasing the sample.

Two strategies were used in the main survey to deal with these issues. The first was to alter the profile of the issued sample, taking into account information from previous survey iterations that showed the relationship between the sampling frame and reported information, and also the response rates in each cell. This adjustment was only considered for countries that used the same sampling frame as the one used in the 2019 survey – where there was a choice between the previous ECS and ESENER-2, ECS was used¹³. The response rates were also adjusted within each country to the overall pilot levels across the strata, for both the screener (CATI) and MM interview (CAWI) stages, but basing the variation between cells on the previous survey data.

In practice this process worked as follows. First, MM interview targets were set, as already described. As noted earlier the targets for large establishments were reduced for most countries, away from the mid-point population distribution, so that they were more realistic factoring in the limited numbers of these establishments in most countries. Second, the MM targets were extrapolated to CATI screener 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. Third, 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 previous survey (ECS3), and additionally factoring in stratum jumpers in the 17 countries that used the same sampling frame as the previous survey.

The second strategy was to manage the sample responsively during fieldwork, using observations from the early stages of fieldwork to inform decisions on sample release in the later stages. This was achieved, first, by firstly randomly allocating the gross sample into batches, and loading these iteratively during fieldwork. Second, the fieldwork outcomes were reviewed on a fortnightly basis during fieldwork, and strata that were predicted to reach target based on the sample loaded were removed from batches of sample not yet loaded. And third, periodic adjustments were made to the CATI screener targets on the basis of main stage outcome information.

In relation to the final point above, a challenge for the management of the sample in this survey was the delay between achieving a CATI screener interview and the subsequent CAWI MM interview. This meant that it was not feasible for the fieldwork agencies to monitor achievement of the MM targets effectively, and instead it was the CATI targets that were monitored by the CATI centres. However, the CATI targets were themselves not useful in any sampling sense, rather they were used

¹³ Across EU Member States in 15 countries the adjustment could be made based on the previous ECS and two on ESENER-2. In 11 no adjustment could be made.

as a tool to achieve the MM targets. It was therefore necessary during the survey to adjust the CATI screener targets based on the overall and within-stratum yield rates between this stage and the CAWI stage, which at the start of fieldwork had been based on a more limited sample size from the pilot.

Evaluation of main stage sample management

In combination these strategies were intended to ensure that the sampling targets were met and that no more than 10% of the gross sample remained “live” at the end of fieldwork (i.e. the situation where the within-cells and overall interview targets have been met but there remains sample that has not been fully dialled). Overall, 15 of the 28 countries were within +/- 5 percentage points of the sampling targets on all nine sampling strata, while remaining within the 10% not closed sample level. For the other countries it was usually no more than two cells that were outside of this range, and at an overall level across all countries only 20 cells missed the +/- 5 percentage points out of a total of 252 cells. In terms of sample closed, across all countries, a total of 5% of the gross sample was not closed – i.e. this sample was contacted at least once, but discarded upon realising the net sample. This compares favourably to the 14% of the previous survey, particularly given the challenges with monitoring a survey with a time lag from CATI screener to CAWI MM interview completion.

Error in the sample in Slovenia

In Slovenia, there was an error at the sample selection stage for the main stage which affected the quality of the achieved sample. The sample provider in Slovenia, Bisnode, omitted large numbers of cases in individual NACE service sectors (most cases in NACE K-S were omitted) when it delivered the sample to Ipsos. Upon receipt of the sample, Ipsos checked that the total number of sampled services at the stratum level was correct and that there were observations in all sector cells but not the distribution across the NACE sectors (the implicit stratification levels). This omission was not picked up during fieldwork monitoring either as reporting focused on the stratum level not the categories within stratum cells. The problem was detected during the weighting process, by which time it was too late to rectify the problem. The error meant that 19% of the intended target population in Slovenia was not included in the sample, 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 point¹⁴.

Screening strategy for company-level frame countries

This section outlines the screening strategy for company-level frame countries. A separate document explains the contact strategy.

Overall, the sampling in 17 of the countries of the EU28 was planned based on a company-level sampling frame (see Table 2, this count excludes Italy). In these countries it was necessary to have an additional sampling stage within the screener interview to select establishments, so that the survey correctly reflected the survey population. Had this not been included the sample in these countries would be one predominantly of company headquarters, leading to a biased sample if these types of establishments differed to subsidiary establishments. This section outlines in detail the various options for approaching the establishment screening and the evidence that underpins the choice of final approach.

At the proposal stage, Ipsos suggested an approach to screening for establishments from company-level frames that involved the following steps in the event of contacting a multi-establishment company:

1. The initially contacted establishment would be asked to participate in the survey (if eligible).

¹⁴ Based on a worst-case scenario of estimating the size of the sectors K-S in the EU.

2. Of the remaining establishments in the organisation, one or more of those that were eligible would be selected at random, varying the number to select depending on the number eligible.

This suggested procedure was based on the ESENER-2 approach, an approach advocated in the ECS feasibility study report, except that in ESENER-2 only one additional establishment was selected at step 2. It differed to the previous survey and ESENER-1 where the procedure was to randomly select a single establishment from all those eligible (including the contacted establishment) and attempt an interview with this establishment only, skipping Step 1.

The rationale for these proposals was two-fold:

- The proposal to include the first contacted establishment was to ensure **accurate representation of multi-site establishments in the sample**. Although this procedure over-represents company headquarters compared to their subsidiaries¹⁵, relative to the single-site establishments, establishments that are part of multi-site companies are actually in the correct proportion, everything else such as non-response being equal. Ipsos endorsed this approach given the outcomes of ESENER and recommendations of the previous ECS feasibility study, which suggested that accurate representation of multi-site establishments was not obtained in the previous survey (or ESENER-1)¹⁶. Our own analysis, using ESENER-2 data, also demonstrated that there were significant differences on the key survey estimates between different types of establishments: both between additional (usually subsidiary) and first-contacted (usually headquarter) establishments, and also between multi-site and single-site establishments/companies¹⁷. This part of the proposal aimed to deal with the second of these issues, while accepting the cost of the second.
- The proposal to sample more than one additional establishment was designed to help control the size of the selection weights which are required to **represent subsidiaries in their correct proportions**. The required selection weights in the bigger multi-site companies become very large, both due to their size¹⁸, and also the ESENER-2 outcomes¹⁹ suggest a very low participation rate can be achieved amongst the additional establishments (a second establishment interviewed for only 1 in 8.4 main establishments part of companies with eligible second establishments). This meant that the selection weights would need to be trimmed heavily, to limit loss of precision, resulting in weighted samples where subsidiaries remain under-represented. Given the significant differences between main and subsidiary establishments observed on the ESENER-2 key survey estimates, and assuming this would also apply to ECS variables, a recommendation was made to boost the size of the sample of additional establishments.

A different approach was trialled during the pilot, following further review of the previous surveys, which suggested that ECS 2013 had not under-represented multi-establishment companies. This finding called into question the value of the ESENER-2 approach, given it came at the cost of over-

¹⁵ Because the first contacted establishment – the one listed on the sampling frame – is always included in the survey when following this method, and it is usually the headquarters.

¹⁶ This would happen under the ECS 2013 and ESENER-1 approaches if the screening was subject to a high level of refusal, meaning that overall there was a lower response rate amongst establishments from multi-site companies compared with single-site companies.

¹⁷ Technical assessment of the expansion of the Second European Survey of Enterprises on New and Emerging Risks (ESENER-2). Ipsos and CWERC.

¹⁸ To represent multi-site establishments correctly requires inverse probability weights to be applied. For example, if one establishment is selected to complete the interview, from a company with three establishments, its chance of selection is 1 in 3, and to represent these establishments correctly in the weighted sample it would require a weight of 3 (relative to single-site establishments).

¹⁹ Second European Survey of Enterprises on New and Emerging Risks (ESENER-2), Technical Report. TNS Infratest.

representing headquarters. Additionally, the very low response rates amongst second establishments seen on ESENER-2 were a cause for concern²⁰. The pilot approach therefore involved dropping the automatic selection of the contacted establishment (i.e. no step 1) but allowing for selection of multiple establishments (step 2). In the sections that follow the findings in relation to the outcomes from the previous surveys were described first, followed by an outline of the pilot approach and modifications for the main survey.

Do the previous survey procedures under-represent multi-site establishments?

As mentioned, a key reason for switching to the ESENER-2 approach, of automatic inclusion of the contacted establishment, was to avoid under-representation of multi-establishment companies. The previous survey procedures could under-represent these establishments if lower response rates were achieved with this group due to the difficulties associated with establishment screening.

As the ECS feasibility study report points out, a comparison of the outcomes of the screening (company-level frames) and non-screening (establishment-level frames) countries might initially suggest that multi-site establishments were under-represented in the previous survey. Excluding the public sector, 21% of the establishments in the screening countries sample were from multi-site companies, compared with 39% in the non-screening countries (unweighted). The proportions of subsidiaries in the overall samples were also different, at 7% and 13% of these overall samples respectively. As acknowledged, however, that there are a number of reasons why these samples cannot be compared directly, the most important of which are the types of sampling frame (only one establishment listed per multi-site company on the sampling frames of the screening countries, vs. all such establishments listed) and the nature of the countries in each group (generally, the larger economies used establishment-level frames, where multi-site companies are probably more prevalent).

A better way to check this would be to compare the equivalent outcomes from ESENER-2. Given that in this survey the contacted establishment was always targeted for interview, it is able to provide an estimate of the proportions of multi-site establishments on the sampling frames of the screening countries that is not affected by screener response issues. This comparison shows that the figures for the screening countries are very similar to those of the previous ECS. Based on the same population definition as the figures above²¹, 20% of the main (first contacted) establishments in the ESENER-2 screening countries sample were from multi-site companies (in line with the 21% for ECS 2013). The proportions of subsidiary establishments in the screening countries were also similar, at 6% of the overall sample (in line with the 7% for ECS 2013).

Table 8: Screening countries (marked X) in ECS 2013 and ESENER-2

	ECS 2013	ESENER-2		ECS 2013	ESENER-2
<i>Albania</i>	n/a	X	<i>Latvia</i>	X	X
<i>Austria</i>		X	<i>Lithuania</i>	X	X
<i>Belgium</i>		X	<i>Luxembourg</i>		
<i>Bulgaria</i>	X	X	<i>Malta</i>	X	X
<i>Croatia</i>	X	X	<i>Montenegro</i>	X	X

²⁰ It is also worth noting that at ESENER-2 the sampling of additional establishments was greatly simplified. Instead of asking the initial informant to list other establishments in their organisation, they were asked to nominate the site located furthest away from their own establishment, avoiding the need for listing. The poor subsidiary response rate, in spite of this simplification, lead us to believe that the previous European Company Survey could have been similarly adversely affected when attempting to select a different establishment.

²¹ All figures in this section cover the samples for NACE B-N, R and S and size 10+ for both surveys.

	ECS 2013	ESENER-2		ECS 2013	ESENER-2
<i>Cyprus</i>	X	X	<i>Netherlands</i>		
<i>Czechia</i>	X	X	<i>Norway</i>	n/a	
<i>Denmark</i>			<i>Poland</i>		
<i>Estonia</i>	X	X	<i>Portugal</i>	X	X
<i>Finland</i>			<i>Romania</i>	X	X
<i>France</i>			<i>Serbia</i>	n/a	X
<i>Macedonia</i>	X	X	<i>Slovakia</i>	X	X
<i>Germany</i>			<i>Slovenia</i>		X
<i>Greece</i>	X	X	<i>Spain</i>		
<i>Hungary</i>	X	X	<i>Sweden</i>		
<i>Iceland</i>	X	X	<i>Switzerland</i>	n/a	
<i>Ireland</i>			<i>Turkey</i>	X	X
<i>Italy</i>	X		<i>United Kingdom</i>		

These outcomes suggest (i) everything else such as non-response being equal, no evidence was found of an issue with under-representation of multi-site establishments with the previous survey procedures, and (ii) each of the survey procedures brought in a similar proportion of subsidiaries. In other words, the reasons for allowing the less stringent ESENER-2 procedures in the first place (over-representing headquarters as a result) do not appear to be realised.

Do the previous survey procedures over-represent headquarters?

As mentioned, the ESENER-2 procedures are expected to over-represent headquarters (compared with subsidiaries). These types of establishment could also be over-represented on the previous survey if interviewers and field managers ‘shortcut’ the process and opt to interview a willing participant at the initially contacted establishment, in situations where they should seek an interview with a different establishment (as is suggested may have been done in the previous ECS feasibility report on sampling). It can be considered to what extent this happened based on the data from the surveys.

On ESENER-2, 16% of the screening country cases were a headquarter site, out of all of the interviewed cases that were listed on the sampling frame and so were contacted first (which was 77% of all the multi-site establishments). A high proportion of headquarters is to be expected, given the procedures.

In comparison, 14% of the equivalent previous survey sample was a headquarter site (which was 68% of the multi-site establishments). This figure seems high, given the contacted site was not automatically selected. If an assumption is made that each company had a single headquarter site, and all of headquarters contained at least 10 employees, then it can be simulated what the proportion should be, based on the distribution of company size (number of establishments) across the previous survey sample. For example, if an establishment was part of a multi-site company with three eligible establishments then the chance of selecting the headquarters was 1 in 3. This analysis

suggests that 58% of the multi-site establishments should have been a headquarters site²², compared with the 68% from the survey, suggesting there may have been some ‘shortcutting’ of the rules by interviewers²³.

The main implication of this analysis is to cast some doubt on how scrupulously the screening procedures were applied on the previous survey. If there was interviewer shortcutting, then this is likely to have supported higher response rates among multi-site establishments than would otherwise have been obtained.

What can be learnt from the response rates to the screening approaches?

As already highlighted, the response rates for the additional establishments at ESENER-2 were low, of 12% on average. Refusing to complete the screening section of the interview was the main reason for non-response at ESENER-2 (63% of all asked to provide a second contact did not agree to). These rates cannot be compared to those from ECS 2013 for referrals to an alternative establishment as the information was not recorded. However, it is notable that ECS 2013 achieved proportions of multi-establishment companies and subsidiaries that are in line with ESENER-2 when selecting a single establishment vs. ESENER-2’s two.

The ESENER-2 rates may have been affected by the placement of the screening questions. On ECS 2013 the screening happened up-front and interviewers could not achieve an interview unless the screening was completed. On ESENER-2 the screening section was placed after the main respondent had agreed to do the interview, and the interview could still go ahead if the screener was refused. It is possible that this led to the ESENER-2 interviewers not pushing for the participant to agree to the screening.

Sampling multiple establishments

As noted earlier it makes sense to sample additional establishments from the larger companies, to ensure better representation of subsidiaries. This is because with larger companies the probabilities of selection of each establishment are very small, requiring large weights to put them in their correct population proportions. However large weights can be very damaging to survey precision, and therefore would need to be trimmed (see discussion which follows). Therefore, increasing the number of establishments to sample in line with the number of establishments reported at the main establishment is a sensible strategy to limit the size of the weights. For practical reasons it is advisable to cap the number, e.g. at three or five, to limit the burden on the respondent during screening (and so limit non-response to the request).

Considering the distribution of establishments in the previous survey (private sector, screening countries only) a number of options can be compared, see Table 8. This is based on the distribution across the full sample, factoring in estimates from S3_C3 (whether the establishment is single or multi-site) and S5A_C5A (number of establishments with 10+ employees in the company in which the multi-site establishment is based)²⁴. The first column shows the unweighted estimates, with a mean size (number of establishments) across the full sample of 1.7, and proportion of the sample that was multi-site of 14%. The unweighted estimates are however biased as multi-site

²² Note: these figures seem unintuitive (too high) and the reason for this is that the question that establishes whether a company is multi-site does not mention the required size of the establishments to be eligible, whereas when respondents are asked to give the number of establishments the question asks about those with 10 or more employees only. As such, a number of companies turn out to have just one establishment with 10 or more employees (36% of those that are initially multi-site). If these companies are redefined as single-site (i.e. remove them from the base) then the ‘correct’ proportion of headquarters in the sample would be 34%, and the realised interview proportion approximately 47%.

²³ Although one can also expect that some companies would have more than one headquarters, which would mean the 58% is under-stated.

²⁴ A single outlier (3000 establishments site) has been removed as it makes the results a bit easier to interpret

establishments have been under-represented in the sample, given only one is interviewed per company. The second column shows what unbiased (accurate) estimates would look like from this sample, with full inverse probability selection weights without any trimming²⁵. This shows that the actual mean establishment size (number of establishments in the company), when the sample is fully weighted to represent subsidiaries in their correct proportions, is 26.5, and the proportion of multi-site establishments is 49%. In reality, these weights would be trimmed heavily, given the design effect from just this stage of the weighting is 15.7²⁶ (final row of the table), and the effective sample size is 783 from a sample of 12,268.

Table 9: Options for capping the number of establishments to select

	Unweighted	Fully weighted (no trimming)	As ECS 2013 (select 1)	Select up to 2	Select up to 3	Select up to 5
Mean size	1.7	26.5	3.4	4.8	5.9	7.7
% multi-site	14%	49%	34%	39%	41%	43%
Effective N ²⁷	n/a	783	8,306	9,605	10,398	11,167
N	12,268	12,268	12,268	12,268	12,268	12,268
Design effect ²⁸	n/a	15.66	1.48	1.28	1.18	1.10

Four different options were then considered on how the screening could be implemented. In all of these options it was assumed that the weights would be trimmed at a maximum of 5, which is the sort of level of trimming that is likely, although a different level could be used (as the previous survey did not include this step of weighting, it cannot be based on this). Column three shows the ECS 2013 screening approach, with one establishment selected, and then the next three options select up to 2, 3 or 5 establishments (where numbers allow). As the weights are trimmed, different estimates for each of the options are reached, and can also see what the effect is on the precision of the sample. The options which involve selecting additional establishments are more efficient (smaller design effect) as fewer of the cases require any weighting and where they are needed the weights are smaller for more of the sample. For these options the same trimming rule affects less of the sample, meaning they are also more accurate (i.e. the estimates are closer to the 'fully weighted' figures). However, this assumes that response rates are equal, in reality one expects to get a lower level of response where more establishments are asked to participate (due to the longer interview and increased burden of the request), which could reduce the numbers of establishments actually added in the sample. Given this it was agreed to cap the number of establishments to select at three, as beyond that level there was less benefit observed and a greater risk of a reduction in response. In terms of the numbers of companies where this would apply, based on the previous

²⁵ The weight for each establishment is the reciprocal of its probability of selection, which is (n selected, i.e. 1) / (n establishments in the company). In effect each of the single establishments that were interviewed from multi-establishment companies are weighted to represent all the establishments in these companies.

²⁶ The weighted sample is very inefficient due to the presence of large weights – weights of up to 330 in this example (equal to the number of establishments in the largest company)

²⁷ The effective sample size is equal to the sample size divided by the design effect. It is a measure of the actual precision of the sample, i.e. the width of the confidence intervals obtained around survey estimates are based on a sample of this size, rather than on the full sample.

²⁸ The design effect from weighting is an approximation based on Kish's formula: (N * sum of squared weights) / (sum of weights squared).

survey distribution 8% of the companies in the sample had three or more establishments and 6% of the sample two establishments.

Recommendation for pilot test on overall screening approach

In summary, the previous survey screening approach – of a straight random selection of an establishment from multi-establishment sites – appeared the favourable one for a number of reasons. First, there was no apparent benefit of the ESENER-2 approach – of including the contacted establishment automatically – in terms of the sample composition of multi vs single-site establishments. Second, the ESENER-2 approach came at the cost of over-representing headquarters. And third, the ESENER-2 approach appeared to be associated with a much lower response rate amongst additional establishments (i.e. those where a referral is required by the contacted establishment), most likely because the interviewer was able to prioritise the contacted establishment, while the available evidence suggested that the previous survey did not suffer from this issue. Therefore, our recommendations at the pilot stage were as follows:

1. to base the screening on the previous survey approach, in terms of sampling out of all available establishments and not assuming an interview with the one first contacted;
2. to follow a similar screener questionnaire structure to the previous survey and re-use questions where possible; and
3. that up to three establishments were selected from companies of this size or larger, to reduce the size of the weights required to represent subsidiaries accurately.

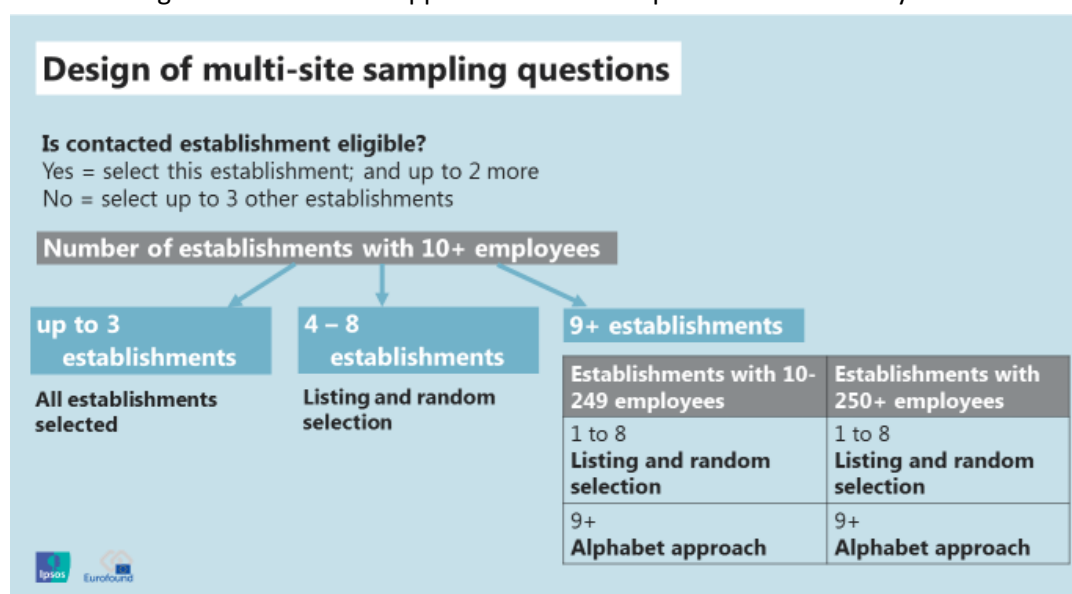
The main areas of concern in relation to this recommendation were (i) whether the move to requesting contacts from multiple establishments would affect overall response and (ii) to what extent (if at all) previous survey interviewers shortcut the procedures. Both areas of concern could have meant that a lower level of response was attained from multi-establishment sites in the current survey than was the case in the previous survey. On the second point, the risk was that shortcuts were taken to keep companies on board (e.g. progressing with an interview with the contacted site even though a subsidiary was or should have been selected), and that if tighter monitoring procedures were able to reduce or eliminate this practice, response rates would suffer.

Implementation of the screener questionnaire

The establishment screening needed to produce comparable data across countries and facilitate an over-sample of the larger establishments. To do this it was necessary to capture the size of the subsidiary establishments in some way. In the previous survey this was done by first asking respondents to provide counts of establishments in each of five size classes, then randomly selecting one size class (if more than one contained establishments), at which point over-sampling of the larger size classes could be accommodated. Thereafter, a random selection of one establishment was taken from within the selected size class. For the current survey this approach was problematic as it would sometimes be necessary to select establishments from more than one size class (given the move to sampling three). Furthermore, asking respondents to count establishments by size class could be very demanding, especially when the company has many establishments, and so it was preferable to simplify the process and sample out of all establishments.

As in the previous survey, having participants first list the establishments in their company, and then sampling from them, is generally considered the most reliable method. However, above a certain number of establishments the listing approach becomes very demanding and it is preferable to use a different method to shorten the length of the screener. In the previous survey this involved the program nominating a random letter and then asking respondents to select the site at a location starting with that letter, or the next nearest letter, as applicable. This approach was repeated in the current survey.

The following chart outlines the approach used in the pilot and main survey.



Pilot implications and adjustments

The strategy described above was tested in the field pilot, with the evidence suggesting that the establishment sampling was improved compared with what had been delivered by the previous surveys. Out of all 'additional' establishments identified in the screener, meaning establishments in addition to the contacted establishment, the contact details were provided for 45%, and the majority of these establishments were subsequently successfully contacted during the pilot fieldwork (67%). The rate of contact detail provision was slightly higher for the first additional establishment (49%), which compares favourably to ESENER-2 (37%). In addition, the rate of agreement for the second and third establishments where applicable, out of those agreeing to provide the details for the first establishment, was very high, at 91%²⁹. This suggested that the

²⁹ The overall rate is a bit lower at 45% as all the cases which said 'no' to the first establishment are automatically 'no' to the 2nd and 3rd where applicable.

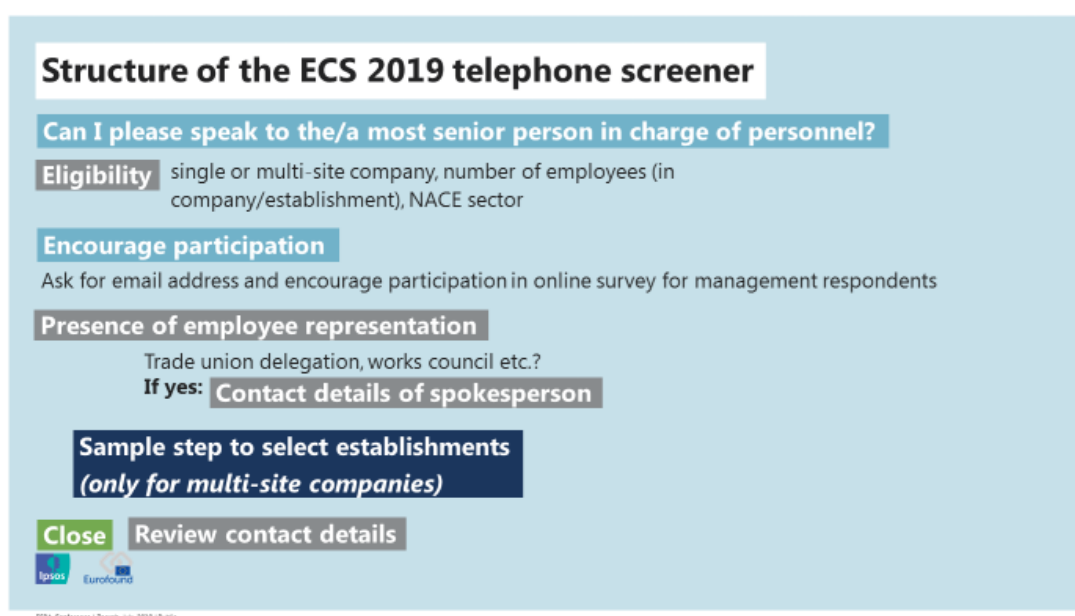
strategy of selecting additional establishments from the larger multi-site companies, to reduce the weights, is effective.

However, based on interviewer feedback from the pilot, the approach was difficult to manage if the contacted establishment was not selected in the screening sampling step. This would happen relatively rarely, as it could only apply to companies with four or more sites, however it was raised as an issue, particularly given that the contact strategy to be adopted for the main survey focused on completing the screener with the MM respondent in order to get their buy-in for the survey. In the pilot approach it could happen that the MM participant was willing to do the interview but then their establishment was not selected. Although rare, it was felt that response to the screening (provision of additional contact details) could be improved if MM participation could be confirmed first, as could response to the ER identification questions.

The following changes to the establishment screening were therefore agreed, affecting the 17 establishment screening countries:

- (i) that the contacted establishment would be selected automatically, in addition to up to another two establishments, and
- (ii) the order of the screener would be revised to improve its flow, to first recruit the MM and collect the ER contact details if applicable, followed by the sampling of further establishments (in the pilot the sampling step was completed first in this sequence).

The following chart shows the placement of the multi-site sampling questions in the main survey. In the pilot survey these questions – the navy-blue block – were placed directly after the eligibility questions.



A potential risk identified this approach was that interviewers and respondents may not place as much emphasis on the sampling step, having completed recruitment for the MM interview. Unfortunately, this risk was realised and in the main survey the response rate for the screening step was 24%, compared with 49% in the pilot. This contributed to the overrepresentation of headquarters in the sample.

Weighting

This section describes the weighting approach implemented in ECS 2019 and summarizes the results from the weighting process. Weighting is required to correct for any disproportionalities in representation due to sampling design and non-response. In addition to these requirements, there is value in weighting to ensure equivalent outcomes to the previous ECS and continuity in trend estimates. As such, any proposed deviations from the previous survey's weighting strategy were considered when formulating the weighting approach.

Four types of weights were estimated: design weights, non-response weights, calibration weights and cross-national weights. An additional weighting stage was required if the sampling frame was at the company level.

Design and non-response weights

Design weights (Step 1)

The first step of weighting was undertaken to account for the disproportionate sampling probabilities across the sampling cells (design weighting)³⁰. The weights were calculated as the reciprocal of the probability of selection of each sampled case within each of the sampling strata (i.e. three-sector-by-three-size-class strata, giving a total of nine cells). The probability of selection of a given sampled case at this stage of the selection is based on the sampling frame numbers and as such is agnostic to the level of the sampling frame (establishments or companies). Further, the probability is determined by the information on the sampling frame rather than that given in the interview. In other words, sampled cases that turned out (during the interview) to belong to a different stratum to the one they were sampled from (termed 'stratum jumpers') were given the same weight as other cases in their original sampling stratum.

Because cases can jump across the strata it is usual to end up with an extreme range of weights within the analysis cells, which was the case for this survey and would impact on the efficiency for analyses of sub-groups if left unchecked³¹. In order to reduce the impact of this, it is standard practice to trim the weights. The WERS approach (Workplace Employment Relations Survey 2011³²) is to trim any sample design weights that are three times larger/smaller than the expected weight within each interview-recorded weighting cell (i.e. across the nine strata) and this is the approach that was adopted. Trimming the weights does slightly alter the weighted profile within each analysis cell but is a price worth paying for increased efficiency. See Table 15 for the comparison of untrimmed and trimmed weights for the weighting steps where trimming was applied.

Non-response adjustments

The previous ECS calculated the design weights based on the probability of being in the net sample. This in effect adjusted, in a single step, the probability of selection into the gross sample and non-response between the gross and net samples on the nine sampling strata. This is an approach that was considered for ECS 2019 however it was agreed to take a more elaborate approach, to provide more

³⁰For an example of this from another high-quality survey see Workplace Employment Relations Survey (WERS):

http://doc.ukdataservice.ac.uk/doc/7226/mrdoc/pdf/7226_the_design_and_administration_of_the_2011_wers_5_august_2013.pdf, page 16-17.

³¹ If, for example, an analysis was to look at small establishments with 10-49 employees, where the analysis would be based on the set of cases that gave this size in the interview, sample precision will be reduced if the weights were not trimmed because of the range of weights in the small firms sub-group; given the cases in this cell will also have been sampled in other cells on the sampling frame (termed 'stratum jumpers').

³² http://doc.ukdataservice.ac.uk/doc/7226/mrdoc/pdf/7226_the_design_and_administration_of_the_2011_wers_5_august_2013.pdf, page 16-17.

control of non-response bias at each of the survey stages. In particular, it was felt that it would be valuable to include adjustments to account for non-response between the CATI screener and CAWI interview stages, which was a new feature of ECS 2019.

The approach that was used was based on 2011 WERS. In this survey, non-response was modelled using a logistic regression model where the predictors were a set of characteristics from the sampling frame. A number of variables were significantly related to non-response, including variables available on ECS 2019 sampling frames (sector, size and region) and others that were not (legal status and the number of sites operated by the company). The approach involves using the model to generate predicted probabilities of response, the inverse of which can be used to generate non-response adjustments to the design weights. If the sampling strata were the only model predictor, then this approach would be equivalent to basing the design weights on the net (instead of gross) sample, i.e. the ECS 2013 approach.

Non-response adjustment – screener stage (Step 2)

Following the sequence of the survey the first modelling stage was to account for response to the screener stage. Following a stage of testing different approaches, a first logit model was used to model the outcome of CATI screener completion, on each country separately, with the following characteristics:

- Outcome variable: productive screener interviews in which MM or ER contact details were obtained.
- Base: all cases in the dialled gross sample, excluding cases that were out of sample during dialling. Cases found to be ineligible in the screener interview were included in the base, given that this information is only available for cases that are successfully screened.
- Predictors (from the sampling frames): top level NACE category, sampling stratum size category (3 categories) and region³³ (see tables below). Any cells with <30 cases in the base at a country level were first combined with a neighbouring category (if one exists) or with the largest category.

Table 10: NUTS level used in the non-response models

	CATI model		CAWI model	
	NUTS category	# of regions in final (collapsed) variable	NUTS category	# of regions in final (collapsed) variable
<i>Austria</i>	NUTS3	35	NUTS2	9
<i>Belgium</i>	NUTS3 (collapsed categories)	40	NUTS2	11
<i>Bulgaria</i>	NUTS3	28	NUTS2	6
<i>Croatia</i>	NUTS3 (collapsed categories)	20	NUTS1	2
<i>Cyprus</i>	LAU1	5	LAU1 (collapsed categories)	3
<i>Czechia</i>	NUTS3	14	NUTS3 (collapsed categories)	13

³³ Additional information was available but only for small numbers of countries, e.g. the founding year of the company. It was agreed to take a consistent approach to the weighting and include variables available for all countries only.

	CATI model		CAWI model	
	NUTS category	# of regions in final (collapsed) variable	NUTS category	# of regions in final (collapsed) variable
<i>Denmark</i>	NUTS3	11	NUTS3 (collapsed categories)	10
<i>Estonia</i>	NUTS3	5	NUTS3	5
<i>Finland</i>	NUTS3	18	NUTS3 (collapsed categories)	15
<i>France</i>	NUTS3 (collapsed categories)	76	NUTS2 (collapsed categories)	20
<i>Germany</i>	NUTS2	38	NUTS3 (collapsed categories)	33
<i>Greece</i>	NUTS2 (collapsed categories)	25	NUTS2 (collapsed categories)	10
<i>Hungary</i>	NUTS3 (collapsed categories)	20	NUTS3 (collapsed categories)	19
<i>Ireland</i>	NUTS3	8	NUTS3	8
<i>Italy</i>	NUTS2	25	NUTS2 (collapsed categories)	15
<i>Latvia</i>	NUTS3	6	NUTS3	6
<i>Lithuania</i>	NUTS3	10	NUTS3 (collapsed categories)	9
<i>Luxembourg</i>	No NUTS regions available	0	No NUTS regions available	0
<i>Malta</i>	No NUTS regions available	0	No NUTS regions available	0
<i>Netherlands</i>	NUTS3 (collapsed categories)	39	NUTS2 (collapsed categories)	11
<i>Poland</i>	NUTS3 (collapsed categories)	73	NUTS2 (collapsed categories)	17
<i>Portugal</i>	NUTS3 (collapsed categories)	23	NUTS2 (collapsed categories)	6
<i>Romania</i>	NUTS3	60	NUTS2 (collapsed categories)	8
<i>Slovakia</i>	NUTS3	8	NUTS3	8
<i>Slovenia</i>	NUTS3	12	NUTS3 (collapsed categories)	10
<i>Spain</i>	NUTS3 (collapsed categories)	51	NUTS2 (collapsed categories)	16

	CATI model		CAWI model	
	NUTS category	# of regions in final (collapsed) variable	NUTS category	# of regions in final (collapsed) variable
Sweden	NUTS3	21	NUTS3 (collapsed categories)	16
United Kingdom	NUTS2	42	NUTS1	12

Note: If not mentioned in the table, NUTS region was not collapsed. The cell with a low count was added to the largest category within the higher NUTS level if collapsing all cells in that category to the higher NUTS level meant collapsing several large cells.

Table 11: Collapsed top-level NACE sectors used in the CATI non-response model

	NACE
Austria	No collapsing needed
Belgium	(B, C)
Bulgaria	No collapsing needed
Croatia	(B, C) (D, E) (K, L)
Cyprus	(B, C, D, E) (K, L)
Czechia	No collapsing needed
Denmark	(B, C) (D, E)
Estonia	(B, C) (D, E)
Finland	(B, C)
France	(B, C)
Germany	No collapsing needed
Greece	(B, C)
Hungary	No collapsing needed
Ireland	(B, C) (D, E)
Italy	(B, C)
Latvia	(B, C) (K, L)
Lithuania	(B, C) (D, E) (K, L)
Luxembourg	(B, C, D, E)
Malta	(B, C, D, E, F) (G, I) (K, L) (M, N, R, S)
Netherlands	(D, E)
Poland	No collapsing needed
Portugal	(D, E)
Romania	No collapsing needed
Slovakia	(B, C) (D, E)
Slovenia	(B, C, D, E) (J, L, M, R, N)

	NACE
Spain	No collapsing needed
Sweden	(B, C)
United Kingdom	No collapsing needed

The inclusion of top-level NACE at this stage was done to provide additional control across the sectors, beyond the three categories of the sampling strata. It was included at this stage given it offered the largest sample sizes, whereas at the interview stage some countries had samples that were too small for so a finely-grained weighting scheme.

The inverse probability of response to the screener stage, from the model, was used to form the weight for this stage.

Non-response adjustment – CAWI interview stage (Step 3)

To account for non-response between the screener and interview stages a second logit model was run of the outcome of the interview stage (on the CATI screener sample). At this stage response was modelled to both the MM and ER interviews, in one model, in order to be able to weight companies with only an ER interview. This model had the following characteristics:

- Outcome variable: response to the MM or ER interview – i.e. any site with response to one of these was counted as productive.
- Base: cases where an invitation email to participate in the MM or ER interview was sent, including, for company-level sampling frame countries, the additional establishments that were identified in the screener interview for the first site contacted.
- Predictors (from the sampling frames): sector group (3 categories), size group (3 categories), region (see Table 10)³⁴.

The inverse probabilities of selection from the second model were also used to create the non-response weighting adjustment for this stage.

The same trimming approach was used in both the CATI and CAWI non-response adjustments, and it was applied only if the range of weights was higher than 10. Three lower and upper percentiles (1st, 2.5th, 5th and 95th, 97.5th, 99th) were evaluated and the pair that would bring the weight range to 10 or lower was used to trim the very small and very large weights. That is, maximum trimming is applied at 5th and 95th percentiles.

Design weight for establishment selection in company-level sampling frame countries (Step 4)

When a multi-site company was identified in the screening countries (i.e. those with sampling frames of companies) up to three of those subsidiary establishments were randomly sampled (see Section 'Working with company-level sampling frames'). Although sampling up to three was an

³⁴ Additional variables that were available from the CATI interview itself were also considered, particularly whether the company had been making a profit (which was a significant predictor when tested at the pilot stage), whether the company/establishment was multi-site or not, whether it was a headquarters or subsidiary, presence of an ER function in the company, and whether frame size or sector was updated in the screener. However, these variables were removed from the CAWI non-response adjustment to simplify the models and improve precision given smaller sample sizes at the CAWI stage. The significance of these variables varied across the countries while most often, with the exception of the profit question, they were not good predictors of response.

increase on the number of the previous surveys (of sampling one), the subsidiaries were still under-represented whenever a company had four or more eligible establishments.

The weights for these establishments were calculated as the reciprocal of their inclusion probability, being the number of establishments selected (i.e. one, two or three) divided by the number of eligible establishments in the company.

Given that this approach resulted in very large weights for companies with a lot of establishments the weights were also trimmed at this stage, with the trimming set at a maximum weight of five (equating to a company of 15 establishments – with three establishments selected the weight for each of these establishments would be five: the reciprocal of 3/15). This was the same level of trimming as in the previous ECS.

Calibration weighting (Step 5)

The final weighting stage was to adjust the design (selection) and non-response weighted interview sample to match the population estimates for selected key measures that were available, namely, activity sector and establishment size; the same variables that were used in the previous ECS.

Given the small sample sizes in many of the countries the same grouped categories were used as those used as sampling strata (three categories for each variable). It was found that a more finely grained calibration approach was not practical as it would introduce large weights.

The weighting targets were based directly on establishment-level statistics, wherever these were available. In 18 countries reference statistics were available only at the company level. To bring these statistics to the establishment level, the calibration targets were adjusted using an inflation factor estimated from the (design weighted) screener survey data. This involved computing the mean number of establishments per company in the applicable countries and using this to adjust to the reference statistics to the establishment level. The inflation was applied based on sector grouping, given that it is reasonable to expect that the establishments would be in the same sector as the parent company, whereas this assumption obviously did not hold for size (number of employees). This approach was used in 17 of the countries that collected information on number of establishments per company. In Luxembourg this was not possible as the sampling frame was at establishment level and for this country information on the number of establishments was not collected in the survey. Instead, the inflation factors were calculated based on the differences in numbers between the company-level and establishment-level sampling frame counts. See the annex (Table 15) for the inflation factors used by country and sector group.

The bounded linear regression method was used to generate the calibration weights. The 'calibrate' command in Stata was used to estimate the weights. The bounded linear regression method did not converge or resulted in negative weights for two countries: Greece and Malta. Therefore, the logistic regression method was used for these two countries.

It is also worth noting that in Cyprus, Ireland, Luxembourg and Malta, construction and production sectors were collapsed into one category and calibrated to the population total for construction sector plus production sector as the cell count was lower than 30 for one of these sector groups (in the MM-only data file). Similarly, medium and large sized companies were merged when calibrating for size in Cyprus, Estonia, Ireland, Luxembourg, Latvia and Malta as cell count for large companies was below 30 in these countries.

The calibration was based on the MM sample in the first instance, creating weights for this sample to represent establishments in Europe. To handle the ER sample, the full MM and ER samples were calibrated separately, to the same reference statistics as were used for the MM sample. This sample hence comprised of one of three types of establishments: (i) those for which there is both a MM and ER interview, (ii) those with only a MM interview and (iii) those with only an ER interview. Groups (i) and (iii) by definition have employee representation, whereas some group (ii) cases do not. This second group was, of course, not required for ER interview analysis, however the rationale for calibrating to this full population is that it is aligned in terms of population definition with the

reference statistics. More specifically, had the MM-only cases been *excluded* in the weighting scheme the weighting would apply to the population of businesses with employee representation in place, which is a set of businesses for which there were no population statistics. Hence, the approach to weighting the ER sample was, in summary, to weight the full sample and then drop MM-only cases, the idea being that the sample remaining should be a representative sample of ER establishments.

Summary and evaluation of weights

Tables 12 and 13 compare unweighted and weighted distributions of MM data to the population distribution by sample strata. It is clear from Table 12 that the weighted distributions effectively approximate the population distributions. The deviations by stratification cell remained below 5 percentage points for all cells.

Table 12: Population, unweighted and weighted distributions by stratification cell (%) - MM online survey

	Sector/size	POPULATION			UNWEIGHTED			WEIGHTED		
		Small	Medium	Large	Small	Medium	Large	Small	Medium	Large
AT	Construction	10.9	1.4	0.1	7.2	2.8	1.3	11.4	1.0	0.1
	Production	11.3	3.1	0.9	10.0	8.9	7.6	12.8	2.3	0.3
	Services	63.3	8.2	0.9	42.0	12.6	7.6	61.3	9.5	1.4
BE	Construction	8.5	1.2	0.2	10.7	3.3	0.7	8.8	0.9	0.3
	Production	11.8	3.5	1.4	11.6	7.0	2.3	11.7	3.6	1.2
	Services	61.0	9.3	3.1	48.8	11.9	3.9	61.1	9.4	3.0
BG	Construction	7.4	1.4	0.1	7.0	3.8	0.6	6.7	2.0	0.1
	Production	19.2	5.9	1.2	15.4	9.8	2.9	19.7	5.5	1.0
	Services	56.0	7.7	1.2	44.0	12.0	4.4	56.2	7.5	1.2
CY	Construction	7.2	0.8	0.0	7.4	3.3	-	6.3	0.8	-
	Production	9.2	1.4	0.2	12.3	3.3	-	10.4	1.3	-
	Services	68.0	11.4	1.8	49.2	23.0	1.6	67.7	12.4	1.1
CZ	Construction	7.0	0.9	0.1	6.1	3.0	0.6	7.0	0.9	0.1
	Production	18.2	6.9	2.1	16.0	14.4	6.4	18.0	6.6	2.5
	Services	52.9	9.7	2.2	36.4	14.0	3.1	53.2	10.0	1.7
DE	Construction	9.9	1.0	0.1	7.3	1.3	0.3	10.0	0.7	0.1
	Production	13.3	5.1	1.3	11.1	9.1	9.7	15.6	3.1	1.1
	Services	56.1	11.6	1.6	34.9	20.0	6.3	53.8	14.0	1.7
DK	Construction	8.8	0.9	0.3	6.1	1.4	1.8	8.7	0.8	0.5
	Production	9.6	1.8	1.6	5.7	5.5	5.3	9.3	2.2	1.5
	Services	66.6	6.3	4.0	49.1	10.8	14.2	67.1	5.9	4.0
EE	Construction	14.0	0.8	0.1	12.8	2.4	-	13.3	1.2	-
	Production	18.7	6.0	0.9	14.8	12.6	1.4	16.7	6.9	0.8
	Services	50.3	7.8	1.4	44.7	9.8	1.6	53.2	7.0	1.0
EL	Construction	3.6	0.3	0.1	5.2	1.0	1.4	3.7	0.2	0.2
	Production	11.7	1.3	1.3	11.6	3.6	3.6	12.5	1.1	0.7
	Services	73.0	5.0	3.7	53.1	11.0	9.6	72.1	5.3	4.3
ES	Construction	9.2	0.8	0.1	4.9	2.1	0.2	8.7	1.3	0.1
	Production	14.9	2.6	0.7	14.8	11.0	3.6	15.6	1.6	1.0
	Services	60.0	8.9	2.8	37.0	20.2	6.1	59.8	9.3	2.6
FI	Construction	12.4	1.6	0.2	9.8	1.6	0.3	12.8	1.5	0.2
	Production	13.5	4.7	1.0	11.5	13.5	3.3	14.0	4.1	0.8
	Services	54.8	10.1	1.7	40.9	15.7	3.4	54.2	10.5	1.9
FR	Construction	10.8	1.1	0.1	5.4	2.2	0.4	10.9	1.0	0.1
	Production	12.7	3.4	0.6	10.1	7.3	6.3	13.4	3.0	0.5
	Services	60.2	9.8	1.5	42.9	15.4	10.1	59.7	9.9	1.4
HR	Construction	10.3	1.5	0.2	7.5	5.4	0.9	10.3	1.6	0.2
	Production	21.3	5.3	1.3	12.0	14.3	5.7	20.6	5.5	1.7
	Services	51.7	6.8	1.6	36.3	15.0	3.0	52.8	6.4	1.1
HU	Construction	9.3	0.8	0.0	10.8	2.1	0.2	9.1	1.0	0.1
	Production	18.5	5.9	1.7	16.7	8.8	1.5	18.7	6.1	1.2
	Services	54.4	8.0	1.4	46.2	11.1	2.6	54.4	7.5	1.9

	Sector/size	POPULATION			UNWEIGHTED			WEIGHTED		
		Small	Medium	Large	Small	Medium	Large	Small	Medium	Large
IE	Construction	7.4	0.7	0.1	4.0	2.3	0.3	4.4	1.6	0.2
	Production	7.9	2.5	0.8	9.7	5.3	2.7	9.5	2.2	1.5
	Services	67.7	11.1	1.9	52.0	21.7	2.0	69.1	10.3	1.3
IT	Construction	7.4	0.4	0.0	4.7	0.6	0.1	7.3	0.6	0.0
	Production	27.3	4.1	0.6	23.1	11.1	5.1	27.4	3.8	0.8
	Services	53.5	5.6	1.1	41.3	9.5	4.5	53.6	5.6	0.9
LT	Construction	11.7	1.9	0.2	8.0	3.5	0.2	11.9	2.0	0.1
	Production	14.8	5.3	1.0	11.2	12.2	5.7	15.2	3.8	1.5
	Services	55.4	8.5	1.3	33.7	23.7	1.8	55.5	9.3	0.7
LU	Construction	15.2	2.7	0.4	10.1	3.8	0.4	15.9	3.1	0.3
	Production	5.4	2.2	0.6	3.8	3.4	1.3	4.0	2.1	1.1
	Services	59.1	11.9	2.6	50.6	21.5	5.1	59.7	11.7	2.1
LV	Construction	10.8	1.5	0.2	11.1	3.9	-	10.6	1.9	-
	Production	16.7	5.2	0.6	13.6	12.5	1.4	15.4	6.2	0.8
	Services	54.9	8.7	1.4	44.7	11.7	1.2	56.3	7.5	1.2
MT	Construction	4.4	0.7	0.2	0.7	1.4	0.7	5.2	0.8	0.1
	Production	8.9	1.9	0.5	1.4	13.8	2.1	5.3	4.2	0.9
	Services	69.1	11.8	2.5	37.9	29.0	13.1	72.1	8.1	3.4
NL	Construction	7.9	1.2	0.3	4.3	3.7	0.4	7.5	1.8	0.2
	Production	8.2	11.1	1.5	5.4	12.5	6.4	10.1	8.5	1.6
	Services	58.4	9.4	2.0	39.2	19.1	8.9	57.7	10.7	1.9
PL	Construction	10.1	1.1	0.1	5.3	2.4	0.5	10.0	1.2	0.1
	Production	21.6	5.6	1.3	18.4	20.4	4.9	22.4	4.9	1.2
	Services	51.5	7.4	1.3	28.7	15.9	3.4	50.8	8.0	1.4
PT	Construction	9.6	1.0	0.1	8.1	2.2	0.3	9.7	0.9	0.1
	Production	22.3	5.1	0.7	17.4	16.2	2.0	21.7	5.9	0.6
	Services	52.6	7.2	1.4	38.0	12.9	2.9	53.2	6.4	1.5
RO	Construction	10.5	1.5	0.1	11.2	4.5	1.1	10.0	1.8	0.2
	Production	17.0	5.3	1.5	15.0	13.4	7.2	17.0	5.3	1.4
	Services	54.7	7.9	1.5	36.4	7.1	4.0	55.1	7.7	1.5
SE	Construction	12.4	1.4	0.1	9.9	2.7	0.4	11.8	1.9	0.2
	Production	10.1	2.7	0.7	5.6	3.1	1.0	11.8	1.5	0.2
	Services	61.5	9.7	1.6	49.7	20.9	6.8	60.5	10.2	1.8
SI	Construction	10.4	1.0	0.1	9.2	2.2	0.4	9.7	1.7	0.1
	Production	20.3	6.6	1.5	21.9	5.6	5.0	20.7	5.5	2.3
	Services	50.5	8.0	1.6	43.5	9.2	3.1	50.8	8.4	0.9
SK	Construction	5.9	0.8	0.1	4.7	3.0	-	4.6	1.6	-
	Production	8.7	3.3	1.0	14.1	8.6	2.2	10.4	2.5	0.8
	Services	65.8	11.8	2.6	55.7	10.5	1.1	65.4	12.8	2.0
UK	Construction	5.2	0.8	0.1	5.3	3.2	0.3	5.3	0.8	0.1
	Production	8.0	2.4	0.4	7.6	7.6	4.0	8.7	1.8	0.3
	Services	69.9	11.3	1.9	46.6	16.9	8.5	69.2	12.0	2.0

Table 13: Weighted vs. population distribution, deviations by stratification cell (percentage points) – MM online survey

	Sector/size	Small	Medium	Large
AT	Construction	-0.5	0.4	0.0
	Production	-1.5	0.9	0.6
	Services	1.9	-1.2	-0.6
BE	Construction	-0.3	0.3	-0.1
	Production	0.1	-0.2	0.2
	Services	-0.1	0.0	0.1
BG	Construction	0.6	-0.6	0.0
	Production	-0.5	0.5	0.1
	Services	-0.2	0.2	-0.1
CY	Construction	0.9	0.0	NA
	Production	-1.2	0.1	NA
	Services	0.3	-1.0	0.7
CZ	Construction	0.0	0.0	0.0
	Production	0.2	0.3	-0.4
	Services	-0.3	-0.3	0.5
DE	Construction	-0.2	0.3	-0.1
	Production	-2.3	2.0	0.2
	Services	2.4	-2.4	0.0
DK	Construction	0.1	0.1	-0.2
	Production	0.3	-0.4	0.1
	Services	-0.5	0.4	0.1
EE	Construction	0.7	-0.4	NA
	Production	2.0	-0.9	0.1
	Services	-2.9	0.8	0.5
EL	Construction	0.0	0.1	0.0
	Production	-0.9	0.2	0.6
	Services	0.9	-0.3	-0.6
ES	Construction	0.5	-0.5	0.0
	Production	-0.6	1.0	-0.2
	Services	0.2	-0.5	0.2
FI	Construction	-0.5	0.1	0.0
	Production	-0.5	0.6	0.2
	Services	0.6	-0.3	-0.2
FR	Construction	-0.1	0.0	0.0
	Production	-0.7	0.3	0.1
	Services	0.4	-0.1	0.0
HR	Construction	0.1	-0.1	0.0
	Production	0.8	-0.2	-0.4
	Services	-1.1	0.4	0.5
HU	Construction	0.2	-0.2	-0.1
	Production	-0.3	-0.3	0.5
	Services	0.0	0.5	-0.5

	Sector/size	Small	Medium	Large
IE	Construction	3.0	-0.9	-0.1
	Production	-1.5	0.3	-0.8
	Services	-1.4	0.9	0.6
IT	Construction	0.1	-0.2	0.0
	Production	-0.1	0.3	-0.2
	Services	-0.1	0.0	0.2
LT	Construction	-0.2	-0.1	0.2
	Production	-0.4	1.4	-0.5
	Services	-0.1	-0.8	0.5
LU	Construction	-0.7	-0.4	0.1
	Production	1.3	0.2	-0.5
	Services	-0.6	0.1	0.5
LV	Construction	0.2	-0.4	NA
	Production	1.2	-1.0	-0.2
	Services	-1.4	1.1	0.3
MT	Construction	-0.8	-0.1	0.1
	Production	3.6	-2.3	-0.3
	Services	-3.0	3.7	-0.9
NL	Construction	0.4	-0.6	0.2
	Production	-1.9	2.5	-0.1
	Services	0.7	-1.3	0.1
PL	Construction	0.1	-0.1	0.0
	Production	-0.8	0.7	0.1
	Services	0.7	-0.6	-0.1
PT	Construction	-0.1	0.0	0.0
	Production	0.6	-0.8	0.2
	Services	-0.6	0.8	-0.1
RO	Construction	0.5	-0.3	-0.1
	Production	0.0	-0.1	0.1
	Services	-0.4	0.3	0.0
SE	Construction	0.6	-0.5	-0.1
	Production	-1.7	1.1	0.4
	Services	1.0	-0.6	-0.3
SI	Construction	0.6	-0.7	0.0
	Production	-0.3	1.1	-0.8
	Services	-0.3	-0.4	0.7
SK	Construction	1.3	-0.9	NA
	Production	-1.7	0.9	0.3
	Services	0.4	-1.0	0.6
UK	Construction	-0.1	0.1	0.0
	Production	-0.7	0.6	0.1
	Services	0.8	-0.7	-0.1

Trimming

Trimming was applied at each weighting step (apart from step 3, where no trimming was required in any of the countries). Trimming reduced the range of weights as intended. In CATI non-response weighting, trimming was applied in 17 out of 28 countries. Out of these 17 countries, only in two countries, Malta and Poland, trimming failed to bring down the range of weights less than 10, the threshold set for the trimming in Step 2. In Malta, the CATI non-response models revealed that the establishments with no size information (in the sample frame) were much less likely to respond to the screener survey which resulted in larger non-response weights for these cases.

No explicit trimming was applied at the calibration step (Step 5), instead a bounded generalized regression approach was used where the minimum value of the ratio exit weight to entry weight (lbound) was set to 0.1 and the maximum value of the ratio exit weight to entry weight was set to 2 (ubound)³⁵. In two countries, Malta and Greece, logistic regression calibration was used as the linear approach did not converge; no trimming was applied in these countries as inserting bounds was not possible in the logistic method.

Overall, trimming reduced variation and increased precision. Trimming did not lead to significant deviations from the population targets.

Design effects

A design effect (DEFF) in its general form measures the relative increase or decrease in the variance of an estimator due to deviations from simple random sampling. The associated design effect for each weighting step is estimated using the following formula:

$$DEFF_p = n * \frac{\sum_{i=1}^n w1_scaled_i^2}{(\sum_{i=1}^n w1_scaled_i)^2}$$

Note that this calculation of DEFF is based solely on the weights themselves whereas DEFF can also be calculated as a variable specific measure. Table 16 shows the design effects across countries by weighting step. The DEFF remained below two for all countries and across weighting steps except for Malta where it was 2.8 at the calibration stage. This was mainly due to substantive differences between the unweighted sample and the population distribution, and small sample size achieved in this country. Malta had a high proportion of stratum jumpers as well which contributed to the imbalance between the achieved sample and population distribution.

³⁵ Please see the documentation on calibration command in Stata for more information.

Table 14: Design effects by weighting step and country – MM online survey

		STEP1	STEP2	STEP3	STEP4	STEP5
		Design weight	Design & CATI nr wgt	Design & CATI nr & CAWI nr wgt	Design & CATI nr & CAWI nr & Establishment wgt	Design & CATI nr & CAWI nr & Establishment & Calibration wgt
	Sample size	<i>s1_wgt</i>	<i>wgt_s12</i>	<i>wgt_s123</i>	<i>wgt_s1,234</i>	<i>s5_wgt_final</i>
Austria	1,010	1.4	1.6	1.7	1.7	1.8
Belgium	1,011	1	1.1	1.2	1.2	1.3
Bulgaria	1,024	1	1.1	1.2	1.2	1.2
Croatia	560	1.2	1.3	1.3	1.3	1.4
Cyprus	122	1	1.1	1.3	1.3	1.4
Czechia	904	1	1.2	1.3	1.3	1.3
Denmark	1,011	1.1	1.2	1.2	1.2	1.3
Estonia	501	1	1	1.1	1.1	1.1
Finland	1,032	1.2	1.2	1.2	1.2	1.2
France	1,360	1.3	1.4	1.4	1.4	1.5
Germany	711	1.2	1.2	1.3	1.3	1.4
Greece	501	1.1	1.2	1.4	1.4	1.5
Hungary	1,087	1	1.2	1.2	1.2	1.2
Ireland	300	1.1	1.2	1.2	1.2	1.3
Italy	1,498	1.2	1.2	1.3	1.3	1.3
Latvia	514	1.1	1.1	1.2	1.2	1.2
Lithuania	510	1.2	1.4	1.4	1.4	1.5
Luxembourg	237	1	1.3	1.6	1.6	1.7
Malta	145	1	1.6	1.9	1.9	2.8
Netherlands	1,030	1.3	1.3	1.4	1.4	1.5
Poland	842	1.1	1.5	1.5	1.5	1.7
Portugal	973	1.2	1.2	1.2	1.2	1.3
Romania	815	1.1	1.2	1.3	1.3	1.3
Slovakia	361	1.2	1.2	1.3	1.3	1.3
Slovenia	556	1.1	1.1	1.1	1.1	1.1
Spain	1,477	1.2	1.4	1.4	1.4	1.6
Sweden	1,080	1.1	1.1	1.2	1.2	1.3
United Kingdom	697	1.2	1.3	1.3	1.3	1.4

Screening data weighting

The calculation of the weights for the CATI screener data file involved four steps: design weights, CATI non-response weighting, establishment selection weights and finally calibration weighting. The same population targets (inflated to the establishment level where necessary) as for the MM and ER online surveys were used for the screener data.

Design weights (Step 1), CATI response weights (Step 2) and establishments weights (Step 3) were calculated in the same way as the MM and ER data weighting. In the MM online survey CATI non-response model, the definition of a successful interviewer was a case for which MM and/or ER contact details were obtained. For the screener data non-response model, this definition was changed to completed screener interviews with eligible establishments (for which size and sector information was recorded).

The calibration targets were revised for two countries due low cell count in the screener data. In Malta, construction and production sectors were collapsed into one category and calibrated to the population total for the construction sector plus production sector. In addition, medium and large sized companies were merged when calibrating for size in Cyprus. In all countries, bounded generalized linear regression was used for calibration.

The maximum design effect for the final screener weights was 2.4 for Malta. All other countries had design effects that remained below two.

Annex

Table 15: Inflation factors used to bring company level reference statistics to establishment level

	Construction (NACE F)	Production (NACE B-E)	Services (NACE G-N, R and S)
<i>Belgium</i>	1.189	1.464	1.688
<i>Bulgaria</i>	1.029	1.165	1.224
<i>Cyprus</i>	1.126	1.096	1.548
<i>Czechia</i>	1.135	1.293	1.559
<i>Estonia</i>	1.053	1.023	1.108
<i>Greece</i>	1.018	1.163	1.282
<i>Spain</i>	1.195	1.199	1.594
<i>Croatia</i>	1.051	1.109	1.274
<i>Hungary</i>	1.051	1.265	1.235
<i>Luxembourg</i>	1.285	1.586	1.537
<i>Lithuania</i>	1.037	1.013	1.102
<i>Latvia</i>	1.047	1.037	1.064
<i>Malta</i>	1.369	1.369	1.446
<i>Poland</i>	1.063	1.186	1.292
<i>Portugal</i>	1.060	1.114	1.312
<i>Romania</i>	1.073	1.130	1.253
<i>Slovenia</i>	1.017	1.089	1.272
<i>Slovakia</i>	1.019	1.102	1.116

WPEF20014

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.