



Industrial relations and social dialogue
**Preliminary framework for a
Eurofound Collective Agreements
Database on collectively bargained
minimum pay rates**

[Regulating minimum wages and other forms
of pay for the self-employed](#)

Authors: Janna Besamusca, Martin Guzi, Kea Tijdens, Ludmila Caminha Barros, Gabriele Medas, Paulien Osse and Gabor Szüdi (Central European Labour Studies Institute)

Research managers: Christine Aumayr-Pintar, Barbara Gerstenberger and Jakub Kostolny

Eurofound reference number: WPEF22043

Acknowledgements: The authors of the report would like to thank the Eurofound expert group that commented and advised on this report; the members of the group, including European Commission staff, are Alfonso Arpaia, Anna Vernet, Tatiana Lopez Garrido, Joe Rieff, Alkistis Zavakou, Duco Claringbould, Jean-François Guillardreau, Nils Trampe, Patrick Karlsson, Torsten Mueller, Paul de Beer, Peter Nachev, Mariusz Zielonka, Enrique Fernandez-Macias, Andrea Garnero, Catarina Braga, Paul Schwab, Anna Fransson, Thorsten Schulten, Claudio Lucifora and Veronika Pálová. Furthermore, the authors thank the national collective bargaining agreement registries in the EU Member States that kindly shared their expertise and data: ÖGB Verlag, Austria; Federal Public Service Employment, Labour and Social Dialogue, Belgium; National Institute for Conciliation and Arbitration (NICA), Bulgaria; Ministry of Labour and Pension System, Family and Social Policy, Croatia; Ministry of Labour, Welfare and Social Insurance, Department of Labour Relations, Cyprus; Ministry of Labour and Social Affairs, Czechia; Confederation of Danish Employers (DA); Ministry of Labour, Estonia; Ministry of Justice, Finland; Ministry of Labour, France; Federal Ministry of Labour and Social Affairs, Germany; Ministry of Labour, Greece; Ministry for Innovation and Technology, Hungary; Labour Court, Ireland; National Council for Economics and Labour, Italy; Official Gazette, Latvia; Ministry of Social Security and Labour, Lithuania; Inspectorate of Labour and Mines, Luxembourg; Department for Industrial and Employment Relations of the Government, Malta; Ministry of Social Affairs and Employment, Netherlands; Ministry of Family, Labour and Social Policy, Poland; Directorate-General of Labour and Employment, Portugal; Ministry of Labour, Family and Social Protection, Romania; Ministry of Labour, Social Affairs and Family of the Slovak Republic; Ministry of Labour, Family, Social Affairs and Equal Opportunities, Slovenia; Ministry of Employment and Social Economy, Spain; and the National Mediation Office, Medlingsinstitutet, Sweden.

© European Foundation for the Improvement of Living and Working Conditions (Eurofound), 2022
Reproduction is authorised provided the source is acknowledged.

For any use or reproduction of photos or other material that is not under the Eurofound copyright, permission must be sought directly from the copyright holders.

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

The European Foundation for the Improvement of Living and Working Conditions (Eurofound) is a tripartite European Union Agency established in 1975. Its role is to provide knowledge in the area of social, employment and work-related policies according to Regulation (EU) 2019/127.

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

Abstract	1
Non-technical summary	2
List of tables	5
Glossary of terms related to pay scales and bargaining	6
1. An EU-wide database on collectively bargained minimum pay rates	8
2. Population and sampling frames	10
2.1 The population	10
2.2 Sampling JOBs.....	12
2.3 Estimates of the number of jobs (not) covered by collective agreements.....	13
3. Sampling collective agreements in different bargaining contexts	16
3.1 The universe of collective agreements.....	16
3.2 Linking collective agreements to selected JOBs and sectors.....	23
3.3 Selecting collective agreements in different bargaining contexts.....	28
3.4 Strategies for selecting collective agreements.....	31
4. Coding parameters for the database	34
4.1 JOBs versus sectors: expected data quality and potential indicators	34
4.2 Measuring metadata	35
4.3 Measuring occupations	38
4.4 Measuring pay rates	39
4.5 Measuring working hours.....	45
5. Discussion and conclusions	47
5.1 Two approaches for the Eurofound Collective Agreements Database	47
5.2 Indicators derived from the two approaches.....	49
References	51
Appendix I: Full list of JOBs	52
Appendix II: Sources for collective agreements universe count	64

Abstract

In this preliminary conceptual and measurement framework, a proposal for the sampling and coding of a new Eurofound Collective Agreements Database is presented. This will result in an EU-wide database on collectively agreed minimum pay rates applicable to a selection of low-paid jobs or sectors, which will be representative, comparable and capable of tracking pay rates in collective agreements ('CAs') over time. Section 2 of this paper discusses two methods for the selection of low-paid jobs: one based on microdata of the industrial sectors in which low-paid workers are found, and one using the Eurofound JOBs methodology ('JObS'). Section 3 lays out the proposal for the selection of CAs linked to JOBs and sectors in the 27 EU Member States. In the selection of CAs, differences in the modalities of collective bargaining on pay across EU Member States are taken into account, as well as the size of and access to the population of CAs. Section 4 details the parameters that need to be coded in the Eurofound Collective Agreements Database in order to be able to estimate collectively negotiated pay rates in a comparable manner. Section 5 summarises two sampling and coding strategies and outlines a proposal for piloting these methods in five EU Member States, as well as outlining the implications of both strategies on the resulting indicators and analyses that can be derived.

Note: This preliminary concept is part of Module 2 of Eurofound's pilot project on minimum wages.¹ It was mainly drafted during 2021 and refined based on discussions with the expert group supporting the implementation of this project module. Subsequently Eurofound developed the prototype of the database and rolled out a 'mini-pilot' covering eight Member States, in which both the JOBs approach and the sector approach were tested, as envisaged and outlined in this preliminary concept. This mini-pilot phase took place between June and September 2022. Based on the results of the mini-pilot, it was decided to go ahead with the implementation of the sector approach for the extension of the data collection to more sectors and all EU countries. In addition, several modifications to the preliminary concept were made. An integrated final report, which will contain a summary of the final concept and measurement framework, as well as an overview of the data obtained, will be compiled as part of the project during 2023.

¹ *Role of the minimum wage in establishing the Universal Labour Guarantee*; see more details at <https://www.eurofound.europa.eu/data/minimum-wages-pilot-project>

Non-technical summary

In 2020, the European Commission at the request of the European Parliament asked Eurofound to carry out a pilot project on minimum wages, which shall feed into the monitoring of the European Commission's initiative on adequate minimum wages. Within module 2 of the pilot project, Eurofound has been asked to 'build a database on minimum wages in collective agreements' and in particular 'A database on minimum wage rates applicable to low paid jobs'.

For the development of this working paper the project team – WageIndicator Foundation, the Central European Institute of Labour (CELSI) and Eurofound – was supported by a group of experts on collective bargaining agreements.² Work on this module started in 2021 with the development of the conceptual and measurement framework presented in this paper.

Principles and challenges for the data collection

As a matter of principle, it was set out that any EU-level database should strive (1) to be representative in terms of the collective agreements it will include, (2) to be comparable across Member States, as far as the different realities around collective bargaining allow for this, and (3) to be able to follow the development of collectively agreed pay over time in a consistent manner. This task is daunting for many reasons, among them the following.

- It is estimated that there are more than 180,000 collective (wage) agreements in the EU, but less than 1,000 in 20 of the Member States (Tijdens, 2021), so any EU-level database will have to rely on a sample of these agreements.
- Collective agreements are negotiated at different levels in the Member States (for instance, at company, sector or cross-sectoral level), and within a country several bargaining levels can coexist.
- A mapping shows that only some EU Member States have full registers of collective agreements available, which would be also accessible for the purpose of the data collection, and in some country contexts full texts of agreements are not in the public domain.
- The pay rates set in collective agreements can relate to different groups of workers, with different skills, occupations, seniority or responsibilities, and there are different ways in which basic pay rates or extra payments could be included in agreements.

Sectors versus JOBs: two possible approaches for the database

One central question discussed with the expert group was whether the focus of the database should be based on a sector approach or on a JOBs approach.

In the **sector approach**, a number of low- to medium-paid sectors would be selected and any type of collective agreement related to these sectors would be identified and a sample included in the

² This group consists of academic experts in the field of collective bargaining/collective agreements; experts nominated by Eurofound's board members from trade unions, employer organisations and governments; practitioners managing national data registers of collective agreements; and experts from the International Labour Organization, the Organisation for Economic Co-operation and Development, the European Commission's Joint Research Centre and the European Commission.

database. The data collection could then include the lowest pay rates within the relevant agreements, and seek to map their development over time. An advantage of this approach is that it is relatively straightforward to apply and it would yield relevant information for the policymaker in the context of the EU minimum wage initiative on the lowest rates of pay that can be found in a range of collective agreements. It would also make it possible to assess the gap between statutory minimum wages and collectively agreed minima in different sectors, in countries where the former exist. A downside of this approach is that the comparability of the coded pay rates across agreements and countries is limited. For example, the minimum pay rate in one agreement related to the care sector could relate to the cleaning personnel working in the care sector, while the minimum pay rate in another care sector agreement could relate to nurses, and a third agreement might relate exclusively to managerial personnel in the care sector.

The **JOBs approach**, on the other hand, goes a step further in aiming to address non-comparability. A job is defined as a specific occupation in a specific sector, such as a cleaner working in the IT industry as opposed to a cleaner working in domestic households or a technician working in the IT industry. When following this approach, the data collection would select a sample of such low- to medium-paid jobs of interest, identify the collective agreements related to these jobs and include the minimum, median and maximum collectively agreed pay rates related to these JOBS in the data collection. An advantage of this approach is the higher degree of comparability across countries and agreements: the collectively agreed pay rates of cooks in the hospitality industry would be compared across Member States, but they would not be compared with the pay rates of kitchen assistants in the hospitality industry. The policymaker would obtain information on the collectively agreed pay rates related to different JOBS across the EU over time, which again could be compared with prevailing statutory minimum wages. A downside of this approach is that it requires more steps in the sampling, so a bias could be introduced in each of the steps. In addition, the approach assumes that collective agreements contain detailed pay scales, which is not the case in all countries or sectors and might potentially cause difficulties for the construction of indicators.

While most experts spoke in favour of applying the sector approach, there was agreement that both approaches should be tested in the first stage of the project in 2022 when the data collection is piloted in a small number of countries.

How to select a representative sample of relevant sectors or JOBS?

Another important question the working paper deals with is how to select the sectors or jobs of interest in the first place and the collective agreements related to these sectors or jobs, given that numbers are high and resources limited.

For the JOBS approach, the concept envisages drawing a sample based on the only available pan-European source that includes in principle all types of workers and companies: the European Union Labour Force Survey (EU-LFS). (Sector-based sampling would be similar, but would ignore the occupational component of the job.) For each country, a national ranking of lowest- to highest-paid JOBS is established. Then an EU ranking of jobs – from lowest to highest paid – is constructed by averaging the mean incomes in JOBS weighted by the number of employees in a specific job and country. The resulting list contains 200 low- to medium-paid JOBS, which represent 30% of employment in the EU. In order to focus the data collection further on those JOBS with a sufficient amount of collective bargaining coverage, it is proposed to add as an additional criterion for selection an estimate of the degree to which workers in the respective JOBS are indeed covered by

collective bargaining agreements. Such estimates are derived from the Structure of Earnings Survey. It is then proposed to select a number of jobs that represent the largest share of the labour force and have a collective bargaining coverage of at least 33% to be included into the data collection. The final number of jobs, sectors and agreements to be included in the data collection will depend on available resources. All thresholds mentioned are indicative and scalable.

How to select a representative sample of collective agreements?

Once the JOBs (or sectors) of interest are chosen, the concept provides clear guidance on how to identify and select relevant collective agreements for different bargaining regimes and within the different countries, based on actual availability of and access to national archives. The project team is currently also establishing working relations with providers of national databases of collective agreements, with a view to obtaining the relevant metadata for the sampling and access to the full texts of the agreements. In cases where no or incomplete data registers exist, a 'snowball sampling' is proposed, in which local experts seek to obtain information on further agreements through signatory parties to agreements they have already identified. As each sector or job may be related to many agreements, it is proposed that all agreements be coded when there are fewer than 10. Where more than 10 agreements exist for a selected job or sector, a rule-based selection is proposed based on the country's bargaining context.

Which parameters will be coded in the database?

The starting point for the data collection will be the metadata on the collective agreements, for example their names, the country, the bargaining level at which they were concluded, the signatory parties, start and end dates, the (estimated) numbers of employees covered, and information on what sectors, occupations or JOBs they relate to.

The main focus of the data collection will be on pay rates of interest within the collective agreements (namely the minimum, median and maximum) and, in order to harmonise the data collection and compare like with like as far as possible, it is proposed to focus on pay rates for adults and obtain further information on contractual working hours related to the pay rates, as well as on agreed extra payments (e.g. bonuses or holiday allowances).

List of tables

Table 1: Selection of 10 low-paid JOBS based on the largest share of the labour force	14
Table 2: Estimated number of CAs in each EU Member State	17
Table 3: National registries of CAs in the EU Member States.....	19
Table 4: Inclusion of CAs from different bargaining levels in national registries, based on official mandate.....	21
Table 5: Online publication of CA metadata by national registries	26
Table 6: Number of CAs in force at 'enterprise or industry/branch level' and sector level by economic activity in Bulgaria in 2020.....	28
Table 7: Overview of proposed country strategies to select collective agreements.....	32
Table 8: Questions to ascertain bargaining level and names of signatories.....	36
Table 9: Questions to ascertain excluded groups from the collective agreements	37
Table 10: Questions to ascertain which industry is addressed in the collective agreements	38
Table 11: Questions to ascertain which occupation(s) is/are addressed in the collective agreements	38
Table 12: Questions to ascertain whether job descriptions are addressed in the collective agreements	39
Table 13: Questions to indicate whether and how wages are included in the collective agreements	40
Table 14: Questions to ascertain whether and how pay scales are differentiated	42
Table 15: Questions to ascertain the minimum, median and maximum amounts in the pay range for the relevant job or sector	43
Table 16: Questions to ascertain whether extra payments are agreed	44
Table 17: Questions to ascertain standard working hours	45

Glossary of terms related to pay scales and bargaining

Articulation: articulation of bargaining is the degree of control of firm-level bargaining by an outside union. It refers to the mutual dependence of union negotiations at different bargaining levels such that the actions of the centre are frequently predicated on securing the consent of lower levels, and the autonomous action of lower levels is bounded by rules of delegation and scope for discretion ultimately controlled by successively higher levels

Bargaining level: collective bargaining can take place at different levels of centralisation. Commonly recognised bargaining levels are the firm level, the sector level, the cross-industry level and the central level.

Career family: this covers one or multiple pay grades linked to a part of the business operation (e.g. human resource management or distribution). The grade's pay range is divided into pay rates or broader zones, which correspond to career advancement and are common to different career families in the business.

Collective agreement (CA): a CA is a negotiated agreement between employers (or employer organisations) and trade unions about pay and/or other conditions of employment.

Job family: one or multiple pay grades may be linked to a family of jobs performed in the operation (e.g. researchers). The grade's pay range is divided into pay rates or broader zones, which correspond to career advancement, but these differ across different job families in the business.

Multi-employer agreements: multi-employer agreements are CAs that are negotiated between multiple employers and one or more employee bodies, elected and/or mandated. The most common subtype is sectoral or sector-level agreements, which are CAs signed by trade unions and employer organisations that represent workers and employers in a specific sector (e.g. metal sector, chemical sector).

Nested agreements: CAs are nested when workers are simultaneously covered by multiple CAs, for instance when workers are covered by both industry (or cross-industry) and enterprise agreements, or by national and regional agreements.

Pay grade: a pay grade, sometimes called pay band, is a set of pay rates that refer to a single job or type of employee/job/work.

Pay range: the pay range is the range between the minimum and maximum rate to be paid in a pay grade, for example €12–€16 per hour for workers employed as cleaners. The range starts at the lowest defined pay rate and ends at the highest defined pay rate for this specific type of employee. The rates within the pay range are sometimes referred to as 'steps'.

Pay rate: a pay rate is a single amount reflecting the wage to be paid to/for a specific type of employee/job/work. A 'type of employee/job/work', in this instance, refers to a group of jobs as it is defined by the CA, and could be a specific occupation or a broad group of occupations. Usually, a pay rate is linked to a job, grade or band, for example €15 per hour for a cleaner.

Pay span: the difference between the lowest and highest rate (as a percentage of the lowest rate) within a range is referred to as the span and used to quantify the potential for career progression within the job.

Pay spine: this is a pay structure containing a single pay grade, usually including a high number of pay rates.

Pay structure: pay structures include all negotiated pay rates. This could include one pay scale table containing several pay rates, bands, grades or scales. Alternatively, a pay structure may contain multiple pay scale tables, which set pay for different career families or job families.

Pay system: pay systems include the pay structure as well as premium pay (e.g. for performance or for night work).

Pay (scale) table: this is a single table displaying negotiated pay rates, possibly within pay grades. CAs may contain multiple pay tables, for instance referring to different job/career families, different time periods, units of measurement, and so forth.

Sampling frame: a sampling frame is a frame that could be used as a basis for sampling and allows the determination of the probability of selection. It is normally any list, material or device that delimits, identifies and allows access to the elements of the surveyed population.

Sampling unit: a sampling unit is one of the units selected for the purpose of sampling. Each unit is regarded as individual and indivisible when the selection is made. A sampling unit can be selected with known probability from a sampling frame.

Single-employer agreement: single-employer CAs are CAs between an employer and a trade union or between an employer and an employee body, elected and/or mandated by the company's staff. Single-employer agreements are also commonly referred to as 'firm-level', 'enterprise-level' or 'company-level' agreements, although these technically exclude public sector single-employer agreements.

1. An EU-wide database on collectively bargained minimum pay rates

At the request of the European Parliament and European Commission, the European Foundation for the Improvement of Living and Working Conditions (Eurofound) is developing an EU-wide database on collectively agreed minimum pay rates applicable to a selection of low- (and medium-) paid jobs. This will result in a representative, comparable and EU-wide database that will be capable of tracking pay in collective agreements over time. Eurofound has contracted the WageIndicator Foundation and the Central European Labour Studies Institute (CELSI) to support the development of the database, which includes, among other things, the development of this conceptual and measurement framework. This work is further supported by an expert group, consisting of academic experts in the field of collective bargaining/collective agreements; experts nominated by Eurofound's Board Members from trade unions, employer organisations and governments; practitioners managing national data registers of collective agreements; and experts from the International Labour Organization, the Organisation for Economic Co-operation and Development (OECD), the European Commission's Joint Research Centre and the European Commission.

In this working paper we propose such a conceptual and measurement framework that can be implemented to sample low-paid JOBS, select matching collective bargaining agreements ('CA'), and link pay rates to JOBS. The working paper aims (1) to arrive at an agreed definition for the measurement framework to monitor minimum pay rates agreed in collective agreements in the EU27 applicable to a selection of low- (and potentially medium-) paid JOBS over time, and to estimate how many employees are covered by these minimum pay rates, and (2) to develop a methodology for a harmonised data collection of collectively agreed pay rates.

In this conceptual framework, we set the following guiding principles that the Eurofound Collective Agreements Database should meet. First, the database should be based on a representative sample of CAs related to low-paid JOBS, or as near to representativeness as possible within the bounds of the existing data infrastructure, as will be detailed in sections two and three. This is a precondition for the publication of indicators on collectively negotiated pay in low-wage employment. For this reason, the sampling of low-paid JOBS and selection of CAs are front and centre in this working paper.

Second, the database should present comparable data in order that country comparisons may be made. This is reflected in the harmonised coding scheme, as well as the proposal to select the same JOBS in all countries. Finally, the database should be sustainable and scalable. This means that the database should be designed in such a way that it can become a part of Eurofound's data infrastructure in the future and that it should be possible to implement new waves of data collections and include new countries or more JOBS and CAs. The proposed sampling procedures and the coding schemas take account of this necessity.

Keeping these principles in mind, we propose a conceptual framework for the Eurofound Collective Agreements Database that codes CAs linked to a representative sample of low-paid JOBS. One central question, raised by the expert group, is whether the focus of the database should be based on a sector approach or on a JOBS approach. Using a sector approach, low-paid (sub)sectors are sampled based on EU microdata, CAs covering those same (sub)sectors are selected using national

CA archives, and the lowest pay rates found in those CAs are coded. Using a JOBs approach, occupations nested in sectors (for example, a cleaner in the cleaning industry or a shelf stacker in the retail sector) are sampled using microdata, CAs covering the sector component of the job are selected using national CA archives, and applicable pay rates/ranges in the agreement are coded based on the occupational component of the job. The choice between the two approaches, both of which could feasibly result in a representative, comparable, scalable and sustainable database, represents one of the central questions addressed in this paper.

This paper is structured to follow the process of sampling and coding of the future database, discussing the benefits and challenges of both approaches. Section 2 proposes a method for the selection of low- (and potentially medium-) paid JOBs using the JOBs and sector approaches. Section 3 lays out the proposal for the selection of CAs linked to JOBs and sectors in the 27 EU Member States. In the selection of CAs we take account of differences in the modalities of collective bargaining on pay across EU Member States, as well as the size of and access to the population of CAs. Section 4 details the parameters that need to be coded in the rep in order to be able to estimate collectively negotiated pay rates in a comparable manner. In coding, we consider all parameters that are necessary for the sampling of CA as well as for the calculation of hourly pay rates. Section 5 summarises the two approaches for sampling and coding the Eurofound Collective Agreements Database.

2. Population and sampling frames

2.1 The population

The first methodological question of the Eurofound Collective Agreements Database is to determine and implement a strategy to sample CAs in a manner that allows conclusions to be drawn about low- (and potentially medium-) paid work in the EU Member States. This section contains the general approach to measuring CA minimum pay rates in a representative way. Four broad sampling strategies are discussed by Tijdens and colleagues (Tijdens, 2021; Tijdens et al., undated): sampling from (1) bargaining actors, (2) the stock of CAs, (3) enterprises³, and (4) the dependent labour force using linked (4a) or non-linked (4b) data.

A sampling strategy can start by using bargaining actors as the primary sampling unit and the CAs they have concluded as the secondary sampling unit. This requires a list of actors to be collected through their employer organisations, trade unions or professional organisations. However, such lists do not currently exist, nor is it evident that all actors will maintain archives of negotiated CAs. This approach is vulnerable to the risk of double (or multiple) counting of CAs with multiple signatories, which would be collected from multiple sources. In addition, there is no clear method of identifying and choosing between conflicting versions if the multiple signatories archived different versions of the CA, for example one version as agreed by negotiators and one as extended to the sector as a whole. This approach was therefore not deemed feasible for the Eurofound Collective Agreements Database.

A representative random or stratified sample from the stock of CAs uses collective agreements themselves as the sampling unit and assumes the existence of lists of all CAs as a sampling frame (Nicolitsas, 2020; Yerkes and Tijdens, 2010). In the absence of such exhaustive lists of CAs, a sampling strategy based on the stock of CAs would have to rely on non-probability snowball sampling methods. Another issue is that collective agreements do not usually stipulate the number of workers covered (directly or by extension) or how many workers are covered by a certain minimum pay rate. However, data on workers covered are crucial to probability sampling. Without data on workers covered, sampling based on the stock of CAs gives equal weight to collective agreements, regardless of whether they cover just a handful of workers or tens of thousands. Additionally, a random or stratified sampling of the CA universe requires well-developed views about the sampling of multi-employer bargaining (MEB) versus single-employer bargaining (SEB) CAs, nested agreements, separate agreements for manual and non-manual workers, and annexes to agreements. Finally, this strategy needs to take into account the volatility of the stock of CAs. As will be discussed in section 3, this option is available for a limited number of EU countries only. For any CA database that aims to include all EU Member States, this is therefore not a feasible option.

A third sampling strategy proposes using employers as the sampling unit by sampling companies through enterprise surveys (Charni et al., 2020; Garnero et al., 2020). The sampling methods employed by such surveys are typically stratified by company size. The employer could be questioned about collective bargaining coverage, as is done in Eurofound's European Company Survey. In a linked employer–CA strategy, covered employers could be asked to submit a copy of the

³ Where we use 'enterprises', we also refer to public sector organisations.

CA or answer a limited set of questions about its content. In the case of an MEB CA, the respondent must identify the domain of the CA following the Nomenclature of Economic Activities (NACE) classification. This method has been applied in the WageIndicator Support for Bargaining – 3 company survey, in which respondents found responding relatively easy (van Klaveren and Gregory, 2019). This strategy, however, requires the sampling of CAs to occur during the data collection for the enterprise survey or that researchers have access to company IDs and matching lists of applicable CAs.

A final sampling strategy proposes using employees as the primary sampling unit by sampling from the dependent labour force, as in the (EU-LFS). To match employees directly to CAs, surveys need to include the questions ‘Are you covered by a collective agreement?’ and if so ‘What is the name of the agreement?’ There are several disadvantages to this approach. First, it relies on employees’ knowledge of collective wage bargaining in their workplace, which might be limited, especially with regard to types of agreements and the existence of multiple agreements covering one workplace. Second, to merge this survey data with data from a CA database, the name of the employer should be known for countries with predominantly SEB, whereas the industrial sector should be known for countries with predominantly sectoral bargaining. The former might be hampered by the privacy concerns of the survey holders; the latter by the fact that sector demarcations in the NACE industry classification and those in CAs often do not overlap, requiring a domain identification connected with the NACE coding.

The dependent labour force can also be taken as the primary sampling unit without direct matching to CAs using a JOBs or sector-based approach. In this approach, a sample of the labour force is drawn based on the distribution of employees across occupations or sectors, and CAs are matched to this sample in the second instance. As a result of the post hoc matching (i.e. separating the process of selecting CAs from the field work for the survey), CAs are not linked one on one to respondents in the survey instrument, and respondent characteristics (e.g. working hours, attitudes, parental status, gender) are not linked to the CAs, implying a loss of individual data but also circumventing a range of privacy concerns. On the aggregate level, average characteristics and demographics of workers in a job or sector can still be associated with CAs linked to that same job or sector. European survey instruments exist that can provide the frame from which to draw such samples, including data on wages to allow the sample to be limited to low- (and medium-) paid workers. This sampling strategy is feasible in the current data infrastructure and is able to handle differences in access to the CA stock across EU Member States (as will be described in section 3). Therefore, this is the sampling strategy we recommend.

Samples can theoretically be drawn based on the distribution of low- (and medium-) paid workers across different sectors, occupations or JOBs, as defined by Eurofound in its JOBs Monitor.⁴ Occupations were rejected as a basis for sampling, because bargaining based on occupational groups is rare and metadata on the inclusion of occupational groups in CAs nearly non-existent (Tijdens, 2021). Sector-based sampling, where low-paid sectors are identified using microdata, is more in tune with bargaining practices and is relatively straightforward. This approach presents minor challenges. First, bargaining often takes place at detailed sector level, which does not necessarily match sector

⁴ For how the European JOBs Monitor understands ‘JOBs’, see its methodology, available at <https://www.eurofound.europa.eu/nb/observatories/emcc/european-JOBs-monitor/methodology>

classifications such as the NACE. The precision of the sampling strategy is therefore dependent on the level of detail in the sector codes provided in the survey instrument. Sampling based on sector would therefore require access to higher-digit NACE codes, which are currently unavailable in the publicly released (EU-LFS) and Structure of Earnings Survey (SES), and large enough survey samples to break down to higher digits, which are unavailable in the European Company Survey. Second, sector-based sampling frames provide little guidance in the selection of pay rates from the CAs by coders, which will probably result in more diversity and less comparability in terms of the types of occupations associated with the selected pay rates across countries and sectors: for example, the lowest pay rates in the information and communications technology (ICT) sector may or may not be applicable to ICT occupations.

Alternatively, sampling can be performed based on JOBS, which are measured as a combination of occupations (two-digit International Standard Classification of Occupations (ISCO) code) and industrial sectors (one-digit NACE code). At this level of aggregation, there are 42 occupations and 21 economic sectors, which generates 882 JOBS. The JOBS methodology is used by Eurofound across various statistical products, including its JOBS monitor, tested and documented extensively, and reflects the view that working conditions differ by both occupation and sector (see, for example, Eurofound and European Commission Joint Research Centre, 2019). Sampling based on JOBS will give coders guidance in the selection of pay rates for low- (and medium-) paid workers, as will be described in section 4, and has the advantage of presenting intuitive indicators due to the clear units of observation (e.g. negotiated minimum pay rates of cleaners in the cleaning industry). Because the JOBS-based sampling strategy is less straightforward than a sector-based strategy, we describe the sampling strategy for JOBS in more detail in the remainder of this section. A sector-based sampling strategy would be similar to the approach detailed below, except that it ignores the occupational component of the job in all steps described in sections 2.2 and 2.3.

2.2 Sampling JOBS

In the first step towards drawing a sample of JOBS for the Eurofound Collective Agreements Database, we follow the Eurofound methodology, which was kindly shared with the WageIndicator and CELSI teams. This step aims to identify low- (and medium-) paid JOBS. The JOBS are defined by crossing two-digit ISCO codes (42 categories) and one-digit NACE codes (21 categories). For reasons of cross-country comparability, we aim to select the same JOBS across EU Member States. While it is technically possible to use the method described below to select different JOBS for each Member State, the interpretation of country differences will be hampered by having to compare the pay of, for example, Dutch plumbers with that of Romanian kindergarten staff. We therefore assume in the remainder of this section that the list of selected JOBS will be the same across Member States.

The identification of low- to medium-paid sectors and occupations (referred to as JOBS or JOB cells) in the EU27 labour market relies on EU-LFS individual data. Following the existing Eurofound JOBS methodology, the analytical sample includes employees working full time (at least 30 hours per week), with known classification of occupation (two digit ISCO code) and economic sector (one-digit NACE code). The income of a worker is characterised by income decile, since EU-LFS data do not include precise wage information. Data from three years (2017, 2018 and 2019) are pooled together to increase the reliability and robustness of identification. Two countries are dropped from this step of the analysis due to missing information on income decile (Sweden) and insufficient occupation classification (Malta). Due to the aim of selecting the same set of low-paid JOBS across all EU

Member States, the unavailability of data for two countries does not impair the inclusion of these countries in the Eurofound Collective Agreements Database; alternatively, national data from Malta and Sweden may be used instead.

The following steps describe the identification of low- to medium-paid JOBs. In the first step, the national ranking of JOBs is constructed by average income decile in each country. The ranking of JOBs are obtained starting with the lowest average income decile. JOBs with fewer than 20 individual observations are dropped from the sample in order to increase the reliability of the final ranking. In the second step, the EU ranking of JOBs is constructed by averaging the national rankings of JOBs weighted by the employment size. The final list then includes 200 low- to medium-paid JOBs that represent 30% of employment in the EU. Alternative thresholds were tested, including the JOBs with the lowest 10% or 20% of rankings, but yielded significantly lower shares of the labour force (under 5% of European workers were in the 10% lowest-ranked JOBs). This threshold was chosen for the conceptual note with the aim of including both low- and medium-to-low-paid JOBs according to the Eurofound JOBs methodology. The strategy is scalable and the method does not change if applied to include the medium- or even medium-to high-paid JOBs.

2.3 Estimates of the number of jobs (not) covered by collective agreements

These first two steps thus result in a list of 200 low- and medium-to low-paid JOBs. Coding the CAs associated with all 200 JOBs, however, is beyond the resources of the Eurofound Collective Agreements Database. Therefore, selection criteria need to be applied to sample a reduced number of JOBs from this list. First, the final EU ranking of JOBs is created by averaging the national ranking of JOBs by the employment size, and therefore the share of the labour force is considered as the selection criterion. Second, collective bargaining coverage is often lower in low-wage employment than in the labour force as a whole. Selecting JOBs from the abovementioned list of 200 thus runs the risk of selecting JOBs with very low or no bargaining coverage. Therefore, we propose to match the JOB cells constructed using the EU-LFS data with data on bargaining coverage using the SES. In this way, data on bargaining coverage are added to each JOB cell that was selected in the previous step. We choose this approach over a strategy that samples JOBs exclusively from the covered labour force for two reasons: first, sampling from the full dependent labour force (with the limitations described in the previous section) results in data that are more comparable across Member States than sampling based on the labour force covered by collective wage bargaining. Second, sampling from the full labour force allows us to perform the construction of the JOBs using the EU-LFS, which has a more representative sample than the SES. The SES excludes firms with fewer than 10 employees from its sampling frame.

In this step, we thus estimate the coverage and non-coverage of JOBs by collective wage bargaining agreements. The information about the share of workers covered by the collective agreement is taken from the SES, since these questions are not covered in the EU-LFS. For the purpose of this working paper and due to the availability of data to the researchers, the 2014 SES was used for the examples presented in this methodological note, while the more recent 2018 SES data will be used going forward. First JOBs are constructed in the SES data by crossing occupations (two-digit ISCO

code) and sectors (one-digit NACE code) in the same way as in the EU-LFS, and then the coverage and non-coverage of JOBs is determined for the jobs selected in the previous step⁵.

Since the JOBs ranking is produced at EU level, the lower participation of EU Member States in the SES than in the EU-LFS does not cause problems in the selection of JOBs. The coverage of JOBs at EU level is calculated applying EU population weights.

Six forms of collective agreements can be distinguished in the SES⁶. For the purpose of selecting JOBs, we disregard the division across different bargaining levels and measure the total share of workers in a job that is covered by any kind of CA. Note, however, that the level of bargaining will be taken into account in the selection of CAs described in section 3.3.

In the final step the EU ranking of low- to medium-paid JOBs obtained from EU-LFS data is matched with information on the collective bargaining coverage obtained from the SES. The final ranking of JOBs then includes information about the level of collective bargaining coverage in the low- to medium-paid JOBs at EU level. More than 95% of JOBs have collective bargaining coverage rates (of any kind of CA) higher than 33%. We include only these JOBs in the database, to avoid selecting JOBs with sparse bargaining and therefore with low prevalence of CAs in the EU Member States. From this ranked list, we propose to select 10 JOBs that represent the largest share of the labour force and have collective bargaining coverage rates of at least 33%. The number of selected JOBs can be adjusted based on Eurofound resources. Alternatively, JOBs can be selected from the 200 JOBs based on their distribution across different sectors, political interest or any other criteria preferred by Eurofound. Table 1 presents such a list of JOBs based on criteria described above. A full list of all 200 JOBs with NACE and ISCO codes is included in Appendix I; the table in the appendix includes additional information on the size of each job as a share of total employment, the average collective bargaining coverage rate in the job and the shares of workers in the job whose earnings fall in the lowest three and lowest five deciles. The strategy is scalable and the method does not change if applied to include more JOBs. When applied to sectors, the approach is implemented on the distribution of workers across industrial sectors only, without reference to occupations, but remains otherwise unchanged.

Table 1: Selection of 10 low-paid JOBs based on the largest share of the labour force

One-digit NACE code and label	Two-digit ISCO code and label			Bargaining coverage (%)
G Wholesale and retail trade; repair of motor vehicles and motorcycles	52 Sales workers			56

⁵ We have to drop three economic sectors that are not covered in SES (Agriculture, forestry and fishing; Activities of households as employers; Activities of extraterritorial organisations and bodies). In total, 2.5% of EU labor force is employed in these economic sectors.

⁶ These are: (1) an agreement at national level, or an interconfederal agreement, covering employees of more than one industry, and usually signed by one or more trade-union confederations and by one or more national employers' organisations; (2) an industry agreement setting the terms and conditions of employment for all or most workers and employees in an individual industry or economic sector; (3) an agreement for individual industries in individual regions; (4) an enterprise or single-employer agreement covering only those employees with the same employer, regardless of size. The agreement may cover only certain local units or groups of employees within the enterprise; (5) an agreement applying only to the employees in one local unit; (6) any other type of agreement not covered above.

Preliminary framework for a Eurofound Collective Agreements Database
on collectively bargained minimum pay rates

F Construction	71 Building and related trades workers, excluding electricians			69
I Accommodation and food service activities	51 Personal service workers			56
Q Human health and social work activities	53 Personal care workers			79
C Manufacturing	75 Food processing, wood working, garment and other craft and related trades workers			55
C Manufacturing	93 Labourers in mining, construction, manufacturing and transport			55
C Manufacturing	82 Assemblers			61
G Wholesale and retail trade; repair of motor vehicles and motorcycles	72 Metal, machinery and related trades workers			52
G Wholesale and retail trade; repair of motor vehicles and motorcycles	93 Labourers in mining, construction, manufacturing and transport			53
N Administrative and support service activities	54 Protective services workers			36

Source: Authors' own estimations based on EU-LFS 2017-2019 and SES 2014

3. Sampling collective agreements in different bargaining contexts

Following the identification of a European sampling strategy for JOBs or sectors, the Eurofound Collective Agreements Database must develop a strategy for selecting CAs associated with these JOBs or sectors. This stage aims to enrich each selected job or sector in each EU Member State with CA information. Due to uneven access to CA metadata and vast differences in the number of CAs that are negotiated in Member States, the Eurofound Collective Agreements Database will need to implement different sampling techniques across Member States. Part of this differentiation stems from differences in bargaining systems, which will remain an issue of interest in the future, while some of it is due to differences in the quality of data, which long-term cooperation between Eurofound and the national registries in charge of archiving CAs might be able to address in years to come. We propose a sampling strategy that aims to obtain access to CAs within each country, depending on country resources and the universe of CAs.

First, the proposed sampling strategy defines the universe of CAs, including both the estimated number of CAs at different bargaining levels and access to structured metadata and full texts of these CAs (sections 3.1 and 3.2). Section 3.3 sets out the proposed country-specific strategy for selecting CAs from available archives according to information on bargaining levels and other relevant characteristics of collective bargaining. Finally, section 3.4 summarises the proposed strategy for selecting CAs in different bargaining contexts. This paper outlines the overall strategy for doing so. In the implementation phase of the Eurofound Collective Agreements Database, however, we foresee that CA registries and/or national industrial relation experts from the Network of Eurofound correspondents will be consulted on the country-specific sampling strategies to ensure local knowledge is utilised.

3.1 The universe of collective agreements

The first step in sampling CAs for the Eurofound Collective Agreements Database is to identify the population of CAs in EU Member States. According to conservative estimates, around 110,000 CAs are currently in force in the European Union⁷. The vast majority of the current number of CAs (approximately 90%) are CA in Germany and France (Tijdens, 2021). This estimate is based on desk research in all EU27 countries, including information in academic publications, information released by national CA archives and information available on the Eurofound and ETUI websites. At the time of writing of this report, the WageIndicator and CELSI teams, supported by Eurofound, are in the

⁷ This only contains CAs of which we confirmed the existence and is therefore a large underestimation since data on enterprise level CAs are missing for a number of countries, including Italy. Estimates in the previous version of this paper put the number of CAs substantially higher. Differences between the 2018/29/20 and the 2021 counts stem primarily from: For Germany, the counted CAs dropped with more than 4,000 to 73,000; For Czechia and Romania no recent information could be found for the initial estimates of more than 3,500 resp. 8,000 enterprise CAs.; For France the counting method was changed. In the 2018 count all non-expired documents in the archive were taken into account, but in the recent count only the documents classified as 'Vigour' and 'Force' were counted.

process of contacting all national CA registries to verify these estimates and establish long-term cooperation⁸.

The number of CAs varies considerably across countries (Table 2). Tens of thousands of agreements are in force in France and Germany (although the number of sector-level agreements is much lower in the former), compared with fewer than 100 each in Luxembourg, Malta, Slovakia and Slovenia. The coding of the full population of CAs appears feasible in some countries, especially when the CAs to be coded are reduced to those covering specific JOBS or sectors. Strategies for the selection of CAs need to be developed for the remaining countries.

Table 2: Estimated number of collective agreements in each EU Member State

Country	CAs counted	Estimated number	Of which MEB	Of which SEB	Year
Austria	In force	560			2021
Belgium	In force	240			2021
Bulgaria	Concluded	754	14	740	2020
Croatia	In force	570	16	554	2014
Cyprus	In force	14			2021
Czechia	Sectoral CAs in force	68			2021
Denmark	In force	600			2017
Estonia	In force	677			2021
Finland	In force	158			2021
France	In force	27,924			2021
Germany	In force	73,000			2021
Greece	In force	271			2021
Hungary	In force	2,869	87	2,782	2019
Ireland	In force	n/a			2019
Italy	In force	1,359	1,359		2020
Latvia	In force	1,152	8	1,144	2016
Lithuania	In force	456			2021
Luxembourg	In force	33			2021

⁸ Establishing contact with the national registries is part of work package 2 of this project. This includes the establishment of cooperation, the organization of two meetings with national registries, and gaining access to meta data and full text CAs from their archives.

Country	CAs counted	Estimated number	Of which MEB	Of which SEB	Year
Malta	In force	34	34		2021
Netherlands	In force	658	485	173	2020
Poland	In force	61	61		2021
Portugal	In force	39	15	24	2021
Romania	In force	36	17	19	2019
Slovakia	In force	18	18		2021
Slovenia	In force	50			2021
Spain	In force	1,893	370	1,523	2021
Sweden	In force	700			2020
Total		113,794			

Source: Authors' own elaboration (see more detail in Appendix II)

To be able to sample from this population, access to a sampling frame (i.e. to metadata about the population of CAs) will be crucial. The most promising sampling frames are archives of CAs maintained by national registries in the EU Member States. To be declared legally binding, either on the enterprises covered or on all enterprises in the industry, CAs in most EU countries need to be registered with the government, or an institution acting on its behalf. Registration typically aims to verify whether the bargaining actors are eligible and whether general and specific validity criteria are met. In some countries, such a registry is maintained by the ministries of labour, employment, social affairs, or similar. In other countries, an agency has been assigned this task, such as the National Council for Economics and Labour in Italy. In a third group of countries, social partners maintain a registry of CAs, such as the WSI Tarifarchiv in Germany, on behalf of the trade union movement. More than one registry has been identified in some countries, such as in the Netherlands, where the organisation servicing the employers' association, the largest trade union confederation and the Ministry of Social Affairs and Employment all maintain separate databases.

As Table 3 shows, all Member States except Denmark, Latvia and Sweden require CAs to be registered with a public institution in order to acquire legal status. In Sweden, the trade union Unionen and the National Mediation Office (Medlingsinstitutet) maintain archives of CAs, while no such trade union or employer based archive has been identified in Denmark. Latvia requires MEB CAs to be published in the Official Gazette, but the gazette is not tasked with maintaining an archive. In Malta, we have found no evidence that the Department for Industrial and Employment Relations of the Government maintains an archive of CAs, although scans of a handful of older CAs are uploaded on the website. In Ireland, registration with the Labour Court Register of Employment Agreements is required, which is in turn mandated to maintain and publish an archive, but only three CAs are currently included in the archive. The Eurofound Ireland correspondent confirms that the Labour Court does not currently maintain an archive of CA, although debates in Parliament imply it may soon be mandated to do so. We therefore propose to make use of the Unionen and

Medlingsinstitutet registries in Sweden and the gazette in Latvia, and to use a snowball sampling strategy in Denmark, Ireland and Malta. In Ireland, potential sources include *Industrial Relations News* (IRN), which maintains an incomplete database of CAs since 1979 and, potentially, the Irish Congress of Trade Unions (ICTU) and individual unions. In Denmark, potential sources include individual trade unions and employer organisations.

Table 3: National registries of collective agreements in the EU Member States

Country	Mandatory registration	Registration by	Archive
Austria	Yes	Chamber of Labour Federal Ministry of Social Affairs, Health, Care and Consumer Protection Austrian Trade Union Confederation	Yes
Belgium	Yes	National Labour Council	Yes
Bulgaria	Yes	Ministry of Labour and Social Policy – National Institute for Conciliation and Arbitration	Yes
Croatia	Yes	Ministry of Labour, Pension System, Family and Social Policy	Yes
Cyprus	Yes	Department of Labour Relations of the Government – Ministry of Labour, Welfare and Social Insurance	Yes
Czechia	Sector-level CAs only	Ministry of Labour and Social Affairs	Yes
Denmark	No	N/A	No
Estonia	Yes	Ministry of Labour	Yes
Finland	Yes	Ministry of Justice	Yes
France	Yes	Directorate General of Labour, Ministry of Labour	Yes
Germany	Yes	Federal Ministry of Labour and Social Affairs Highest labour authorities of the individual <i>Bundesländer</i> German Trade Union Confederation	Yes
Greece	National, sectoral and occupational agreements	General Labour Inspectorate of the Ministry of Labour	Yes
Hungary	Yes	Ministry of Innovation and Technology	Yes

Country	Mandatory registration	Registration by	Archive
Ireland	Yes	Labour Court Register of Employment Agreements IRN	No
Italy	Yes	National Council of Economy and Labour, a governmental agency	Yes
Latvia	Publication in gazette	N/A	No
Lithuania	Yes	Ministry of Social Security and Labour	Yes
Luxembourg	Yes	Inspectorate of Labour and Mines	Yes
Malta	Yes	Department for Industrial and Employment Relations of the Government	No
Netherlands	Yes	Ministry of Social Affairs and Employment	Yes
Poland	Yes	Ministry of Family, Labour and Social Policy (for MEB collective agreements) National Labour Inspectorate (for SEB collective agreements)	Yes
Portugal	Yes	General Direction of Employment and Labour Relations at Ministry of Labour	Yes
Romania	Yes	Ministry of Labour, Family and Social Protection Territorial labour inspectorates (for SEB collective agreements)	Yes
Slovakia	Sector-level CAs only	Ministry of Labour, Social Affairs and Family of the Slovak Republic	Yes, the older higher-level CAs are available on the website
Slovenia	Sector-level CAs only	Ministry of Labour, Family, Social Affairs and Equal Opportunities	Yes
Spain	Yes	Ministry of Employment and Social Economy (for national agreements) Regional registers (for regional agreements)	Yes

Country	Mandatory registration	Registration by	Archive
Sweden	No	Unionen, Medlingsinstitutet	Yes

Note: N/A, not applicable.

Source: Authors' own elaboration

Not all archives include the complete universe of CAs. In some countries access to CAs from different geographical levels, such as local or regional agreements, is uneven. We propose to evaluate this with the help of country experts in countries where this issue occurs, such as Germany and Romania. In five Member States (Czechia, Greece, Latvia, Slovakia and Slovenia) only MEB CAs need to be registered. As Table 4 shows, the CAs from SEB agreements are missing in three of these archives, while Latvia does not maintain an archive at all. Greece does register them; further contact with the Greek registry is required to ascertain how representative the sample of registered SEB CAs is. Therefore, in Czechia, Slovakia and Slovenia, we propose to use the national registries for the selection of sector-level CAs and use a snowball sampling strategy to select enterprise-level CAs. In such a snowball sampling strategy, coders attempt to identify as many CAs covering the relevant JOBS or sectors as possible using desk research and by reaching out to trade union and employer organisations with requests for data. The resulting list of CAs will not be a complete list of CAs unless all parties involved in collective bargaining cooperate, but will provide a basis for non-probability sampling. The degree of bias in these samples will depend on the extent to which the identified CAs are representative of all CAs in the sector or job and should be evaluated with the help of local experts.

Table 4: Inclusion of collective agreements from different bargaining levels in national registries, based on official mandate

Country	National or inter-industry CAs	Sector-level CAs	Firm-level CAs	Extended/binding CAs	Non-binding CAs	Expired CAs	Start year archive data
Austria	Yes	Yes	Yes	Yes	Yes	Yes	Last 25 years
Belgium	Yes, all	Yes, all	Yes, all	Yes, all	No	Yes	1999
Bulgaria	No	Yes	Yes	Yes	No	Yes	2010
Croatia	Yes	Yes	Yes, but mostly at the local labour directory	Yes	No	Yes	n.a.
Cyprus	Yes	Yes	Yes	N/A	Yes	Yes	2010
Czechia	Yes	Yes	No	No	No	Yes	2005

Preliminary framework for a Eurofound Collective Agreements Database
on collectively bargained minimum pay rates

Country	National or inter-industry CAs	Sector-level CAs	Firm-level CAs	Extended/binding CAs	Non-binding CAs	Expired CAs	Start year archive data
Denmark	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Estonia	Yes	Yes	Yes	Yes	No	Yes	2005
Finland	Yes	Yes	Yes	Yes	No	Yes	N/A
France	Yes	Yes	Yes	Yes	No	Yes	Varies by level
Germany	Yes	Yes	Yes	Yes	Yes	Yes	N/A
Greece	Yes	Yes	Yes	Yes	Yes	Yes	1974
Hungary	Yes	Yes	Yes	Yes	Yes	Yes	1990
Ireland	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Italy	Yes	Yes	Yes	Yes	No	Yes	1946
Latvia	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Lithuania	Yes	Yes	Yes	Yes	No	Unclear (deregistration possible)	2016
Luxembourg	Yes	Yes	Yes	Yes	N/A	Yes	1994
Malta	No	No	Yes	Yes	No	Yes	N/A
Netherlands	N/A	Yes	Yes	Yes	No	Yes	Most since 2011
Poland	Yes	Yes	Yes	Yes	No	Yes	1995
Portugal	Yes	Yes	Yes	Yes	No	Yes	From 2005 onwards
Romania	Yes	Yes	Yes	Yes	No	Yes	1993
Slovakia	Yes	Yes	No	No	No	Yes	2011
Slovenia	Yes	Yes	No	No	No	Yes	1991 (independence)

Country	National or inter-industry CAs	Sector-level CAs	Firm-level CAs	Extended/binding CAs	Non-binding CAs	Expired CAs	Start year archive data
Spain	Yes	Yes	Yes	Yes	No	Yes	Varies by sector
Sweden	No	Yes	Yes	Yes	No	Yes	N/A

Notes: N/A, not applicable; n.a., not available.

Source: Authors' own elaboration

In summary, we propose to implement a snowball sampling strategy for the selection of Danish, Irish and Maltese CAs, as well as for SEB CAs from Czechia, Slovakia and Slovenia. In Latvia, sector-level CAs can be manually identified in the Official Gazette publications, while SEB agreements need to be snowballed from social partners. In all other countries, we propose to rely on the national registries, including the Unionen and Medlingsinstitutet registries in Sweden and the IRN and ICTU archives in Ireland. Currently, efforts are under way to secure collaboration with the registries as a part of work package 2 of this project.

3.2 Linking collective agreements to selected JOBS and sectors

For all remaining Member States, we propose to collect information about all CAs registered in the national registries in the EU27 in order to be able to select CAs linked to the NACE, or occasionally ISCO, codes of the selected JOBS from their archives.⁹ In this step we aim to match the NACE/ISCO code combination of each selected JOBS to the NACE/ISCO code combinations of the CAs.¹⁰ This means that the metadata of national registries are used to identify all CAs covering the selected JOBS and/or sectors. This may result in a one-to-one job–CA mapping, but also in a one-to-many or many-to-one mapping. If a CA has no ISCO mapping, we select based on NACE codes only. Because most CAs will be selected in reference to the sector component of the job using national CA archives, these steps are the same regardless of whether a job- or sector-based approach is used.

Metadata at CA level are necessary to be able to select CAs. Necessary information includes two types of data. First, the start and end dates, bargaining level (sectoral/enterprise CA¹¹), (estimated)

⁹ This data collection is currently underway as a part of WP2 of this project.

¹⁰ Where no NACE/ISCO codes are available, such matching would in the first place be based on the title of the agreement, and then based on further content contained within the agreement, such as the scope of the agreement, if explicitly stated, reference to certain professions, and/or the domain of the signatory parties. In practice this will already require country and sector specific industrial relation knowledge, so Eurofound's correspondents will be asked to carry out this selection step.

¹¹ These are abbreviated as respectively multi-employer bargaining (MEB) CAs and single-employer bargaining (SEB) CAs.

number of employees covered, NACE codes¹² and, when applicable, ISCO¹³ codes of these CAs are required to be able to link the CAs in the registry to the selected JOBs. Second, the name of the CA and, where possible, names of signatories, are required to identify the selected CA and retrieve their full texts.¹⁴

In the ideal case, the national registries maintain archives including the abovementioned metadata in a structured format. In that case, full lists of CAs associated with each job–country cell can be retrieved from the registry metadata. Structured data in any format can generally be converted to the chosen data format of the Eurofound Collective Agreements Database with minimal effort, and such conversion scripts can be reused as long as the registry does not change the format of its metadata. In cases where metadata are missing or unstructured, the team of the Eurofound database will need to undertake manual coding efforts for the first release of the database, which can be reused in future waves when stored in application programming interfaces (APIs).

We assume that for quite a few CAs the NACE and ISCO codes will be missing. In countries where this is the case, we propose to undertake NACE coding based on the name of the CA and the signatories, resulting in an API to identify a two-digit NACE code for each CA in the registries of these Member States, and to assign the missing value to ISCO.¹⁵ If the overarching two-digit NACE codes apply to a CA, all applicable NACE codes will be assigned. Regarding the ISCO coding, in the case of an occupational CA the ISCO code will be assigned, plus the most likely NACE codes for the ISCO code (for which we will use the ISCO-to-NACE prediction, freely available from the survey codings website: [Articles \(surveycodings.org\)](https://surveycodings.org)). See section 4 for further information.

In the context of work package 2 of this project, the WageIndicator and CELSI teams are currently contacting all national registries to seek long-term collaboration and to inventory which registries maintain structured metadata and which do not. This inventory will reveal whether metadata are kept in structured formats. Provisional results based on the examination of legal documents, annual reports and web data suggest that metadata will be available in 18 out of 27 countries (Table 5). There are probably more Member States that maintain archives including metadata but do not release these data on their websites or at CA level. Bulgaria, for example, does not publish metadata at CA level, but the published aggregated data, broken down by NACE code and other characteristics, suggest that the data are maintained in this way internally. Achieving collaboration with the registries will be vital in reducing the workload associated with linking CAs to JOBs. Finally, we believe that the coding of NACE sectors of CAs for national registries that do not keep structured metadata will increase the potential benefits of long-term collaboration with Eurofound for these actors. For Eurofound, improving the metadata for these Member States so that they may be sampled from directly in future data collections has obvious advantages.

¹² Agreements which will be selected to be included into the database will be tagged with the full range of NACE (2-digit) codes they relate to. Agreements which cover more than one sector of interest could therefore be located under different NACE codes. For agreements overlapping several (NACE 1 digit) sectors of interest, a ‘multi-sectoral’ sector of interest category could be included, to avoid issues with double counting.

¹³ Note that most CAs will be NACE-based CAs, covering a sector or an enterprise. However, there will also be ISCO-based CAs, covering one or more occupational groups, e.g. airline pilots, or domestic workers.

¹⁴ Note that the names of the signatories are only needed in case the CA name does not provide sufficient information for the NACE or ISCO classification.

¹⁵ An API allows to provide NACE/ISCO codes for CAs which are to be archived in the coming years.

The identification of CAs with selected JOBs or sectors from the metadata of national registries will result in a list of CAs that match the NACE codes of the selected JOBs or sectors. If two JOBs are associated with one CA, the information about that CA will be matched to both JOBs. In section 4, where we lay out our proposal for the coding of parameters for the database, we add a unique identifier to each CA to ensure that indicators calculated from the data will not double count the estimated number of employees covered in such double-linked CAs.

Table 5: Online publication of collective agreements metadata by national registries

Country	Name	Start date	End date	Signatories	Number of employees covered	NACE	Full texts	Annexes	Expired/ outdated CAs
Austria	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Belgium	Yes	Yes	Yes	Yes	No	Yes	n.a	Yes	Yes
Bulgaria*	No	No	No	No	No	No	No		
Cyprus	Yes	Yes	Yes	Yes	No	No	Yes, in scanned format	Yes	Yes
Czechia	Yes	Yes	Yes	Yes	No	Yes	Yes, for sector-level agreements	Yes	Yes
Estonia	Yes	Yes	Yes	Yes	Yes	Yes, NACE	Yes	Yes	Yes
Finland	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
France	Yes	Yes	Yes	Yes	No	IDCC	Yes, full texts	No	Yes
Germany*	No	No	No	No	No	No	No	No	No
Greece	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes
Hungary	Yes	Yes	Yes	Yes	No	Yes	No	No	Yes
Ireland				No	No	No	No	No	No
Italy	Yes	Yes	Yes	Yes	Yes (INPS)	Yes	Yes	Yes	Yes
Latvia	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Lithuania	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
Luxembourg	Yes	Yes	Yes	Yes	No	No	Yes, in PDF	Yes	No
Netherlands	Yes	Yes	Yes	No, in PDF	No	Yes	Yes	Yes	Yes

Preliminary framework for a Eurofound Collective Agreements Database
on collectively bargained minimum pay rates

Country	Name	Start date	End date	Signatories	Number of employees covered	NACE	Full texts	Annexes	Expired/ outdated CAs
Poland*	No	No	No	No	No	No	No	No	No
Portugal	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Slovakia	Yes	Yes	Yes	Yes	No	No	Yes, for sector-level agreements	Yes	Yes
Slovenia	Yes	Yes	Yes	Yes	No	No	Yes, in HTML	Yes	Yes
Spain	Yes	Yes	Yes	Yes	No	n.a.	Yes	Yes	Yes
Sweden	Yes	Yes	Yes	Yes	No	n.a.	Yes	Yes	Yes

Notes: n.a., not available. No information available for Croatia, Denmark, Malta and Romania * Bulgaria, Germany and Poland: only aggregate data available online

Source: Authors' own elaboration

3.3 Selecting collective agreements in different bargaining contexts

Finally, sampling strategies for CAs will be different across countries due to differences in the modalities of wage bargaining, including the dominant level of wage bargaining and the degree of articulation between different levels of bargaining. In step 3, we therefore aim to fine-tune the selection of CAs from the lists with linked job–CA or sector–CA matches that were produced in the previous step. Some JOBS or sectors in some countries will not be matched to any CA, while in other JOBS/sectors and other countries many CAs will be matched. We propose that, if the full list of CAs linked to one job/sector in a country includes 10 CAs or fewer, all CAs be coded for the selected job/sector, meaning we code the full universe.¹⁶ We choose this threshold for coding the full population based on feasibility in regard to workload for coders of the database.

In cases where more than 10 CAs are linked to one job/sector in a country, coding the full universe potentially places a disproportionate burden on coders, especially if a large share of the CAs cover relatively few workers. We therefore propose a rule-based selection method in cases where more than 10 CAs are linked to one job/sector in a country. For example, in the case of Bulgaria, which is shown in Table 6, JOBS in the construction (F) and commerce (G) sectors are associated with 10 and 9 CAs respectively and for these JOBS all available CAs would be coded. In arts, entertainment and recreation (R), on the contrary, 94 enterprise-level and 4 sector-level CAs are registered and a selection will be made. As only NACE codes and no ISCO codes are available, as is expected to be the case in most countries, this matching is done at NACE code level and some selected CAs may not contain pay rates for the job in question. As the assignment of ISCO codes is done in the process of coding the CAs, described in section 4, this means that all CAs identified on the basis of NACE code are included and some may be excluded in the next step. In the remainder of this section, we outline our proposals for selecting CAs in different bargaining contexts.

Table 6: Number of collective agreements in force at 'enterprise or industry/branch level' and sector level by economic activity in Bulgaria in 2020

NACE sector	Enterprise	Industry/branch
A Agriculture, forestry and fishing	75	2
B Mining and quarrying	13	1
C Manufacturing	99	3
D Electricity, gas, team and air conditioning supply	22	1
E Water supply, sewerage, waste management and remediation	44	1
F Construction	8	2
G Wholesale and retail trade, repair of motor vehicles and motorcycles	9	0
H Transportation and storage	34	1

¹⁶ The threshold of 10 CAs is an initial threshold. It may change if the total number of CAs to be coded exceeds the scheduled workload. If the total number of CAs to be coded is lower than expected, we can set a higher threshold.

I Accommodation and food service activities	10	1
J Information and communication	8	
K Financial and insurance activities	3	
L Real estate activities	19	
M Professional, scientific and technical activities	34	
N Administrative and support service activities	14	
O Public administration and defence; compulsory social security	122	
P Education	603	1
Q Human health and social work activities	280	1
R Arts, entertainment and recreation	94	4
S Other service activities	33	
Total:	1,524	18

Source: *National Institute for Conciliation and Arbitration, Bulgaria*

The ICTWSS and Eurofound data collections provide data on bargaining contexts at country level (or at the level of the metal sector in cases of large within-country variation in bargaining patterns). For the Eurofound Collective Agreements Database, however, drawing a sampling grid of CAs representative of the bargaining patterns in the selected sector for each country would be preferable. We propose to use the SES to do so. The SES contains data on two-digit ISCO and one-digit NACE codes and can therefore be directly linked to the JOBS selected in the previous step using the EU-LFS¹⁷; or to the more detailed NACE codes from the Eurostat secure data centre if a sector approach is chosen. In the 2014 SES, these data are available at job level for 19 of the 23 Member States participating in the survey.¹⁸ In countries where no breakdown of bargaining by level can be calculated for the selected JOBS using the SES, we propose to ask the Eurofound correspondents to provide this division based on other data collections. Where no breakdowns are available from the SES or national data collections, we propose to use national-level data from the European Company Survey.

In cases where over 80% of the labour force in a selected job/sector of a country is covered by **sector-level agreements** according to SES data, we propose to select sector-level CAs only, acknowledging that this is the predominant level at which bargaining takes place.¹⁹ In these

¹⁷ In this working paper, we use the 2014 Structure of Earnings Survey, which is the last wave that has been publicly released. We do propose to use the 2018 wave for the sampling of the Eurofound Collective Agreements Database.

¹⁸ Data for Germany, Italy, Netherlands and Sweden are insufficiently detailed to use in the 2014 SES. No 2018 SES data is available at the time of writing of this working paper working paper (Q4, 2022).

¹⁹ This is an initial threshold that might be adapted to match the level of Eurofound resources and if the number of selected sector-CAs turns out to be lower or higher.

contexts, we argue that the sector-level CAs set the wage floors that workers in the sector rely on. We propose to use a stop rule for the selection of CAs in which, first, the 10 sector-level CAs covering the largest share of the labour force are selected to reach at least a number comparable to cases in which the full universe is coded. Afterwards, the next largest CA is selected for coding until the workers (estimated to be covered) under these CAs reaches 75% of the population covered by all CAs associated with the JOB country cell, or the number of selected CAs reaches 25.²⁰

In cases where bargaining in a job/sector of a country only takes place at **enterprise level** according to SES data, the Eurofound Collective Agreements Database will sample from these agreements exclusively. This is expected to be a common occurrence in Estonia, Hungary, Ireland, Latvia, Lithuania, Malta and Poland (OECD and AIAS, 2021). Since no extant sector-level agreements were identified in Romania in our own research, an enterprise CA sampling approach will most likely have to be adopted there too. For the selection of CAs, we use the metadata from the national registries to identify JOB-country or sector-country cells where more than one CA is matched to a job. Using a stop rule, we propose to first select the 10 enterprise CAs covering the largest share of the labour force. For example, if one job is associated with hundreds of CAs, we aim to identify the 10 largest CAs (see next section for identification of number of employees covered per CA). Afterwards, the next largest CA is selected for coding until the estimated covered population reaches 30% or the number of selected CAs reaches 25.²¹

In JOB-country or sector-country cells where **both enterprise- and sector-level bargaining** are conducted and enterprise-level bargaining covers more than 20% of the labour force, we propose to select both sector and enterprise CAs proportional to the distribution found in the SES. We use the metadata from the national registries, relying on multiple archives if no single one contains the full universe of CAs, to identify both MEB and SEB CAs covering the relevant job or sector.

For these cells, we use the detailed (two-digit) NACE data to explore the CA configuration in order to identify nested agreements: notably, whether the MEB CAs overlap with the SEB CAs or are adjacent. This check can be cross-checked with analyses using the European Company Survey's last wave, which allows the identification of enterprises covered by more than one CA. Where sector- and enterprise-level CA NACE codes do not overlap, we propose to consider the CAs not to be nested and use a straightforward selection of CAs, matching the estimated workers covered per CA to the proportion of enterprise- versus sector-level bargaining in the JOB-country or sector-country cell as well as possible. For example, if in the case of NACE code 56 'Food and beverage service activities' one sector-level CA is identified for cafes and pubs, and an enterprise CA for a restaurant chain, both are coded and the workers covered by each are understood to be non-overlapping.

In cases where the NACE codes of sector-level and enterprise-level CAs linked to a JOB-country cell overlap, we propose to assume the pay scales found in enterprise CAs are valid for the workforce of the enterprise in question. Therefore, we propose to select all CAs with overlapping NACE codes so the estimated number of workers covered by enterprise CAs can be subtracted from the estimated number of workers covered by the sector-level CA when calculating indicators. This will address both nested CA structures and opt-outs. Where needed we propose to manually control for many-to-one,

²⁰ This is an initial threshold that might be adapted to match the level of Eurofound resources and if the number of selected sector-CAs turns out to be lower or higher.

²¹ This is an initial threshold that might be adapted to match the level of Eurofound resources and if the number of selected sector-CAs turns out to be lower or higher.

one-to-many and overlapping or nested matches, which ensures that this will be done by coders with maximum knowledge of the local bargaining context.

3.4 Strategies for selecting collective agreements

In summary, we propose to use metadata from the national registries to select CAs in countries where this is possible, and to implement a snowball sample in countries where no archives are available. Where registries maintain archives but no structured metadata are available, we propose to manually code the NACE sectors of CAs. This is an investment, but will only be necessary for the first wave of data collection and might persuade national registries to join the project as well as improve their archiving methods. We then select CAs according to the distribution of CAs across bargaining levels in the selected JOBs in each county, using the SES to calculate the proper ratio. We propose to collect the full texts of the selected CAs from the registries. If those are not available, we propose to collect the full texts from the signatories, via the WageIndicator network or via Eurofound correspondents. Table 7 shows our estimated strategy for each Member State. In the implementation phase of the project, each country strategy will be written out in more detail and national registries and Eurofound correspondents will be asked to give feedback towards the adjustment of the sampling strategy for any yet unaccounted for local factors.

Table 7: Overview of proposed country strategies to select collective agreements

Country	Sampling strategy	Metadata (use available data/make API)	Full text (online/offline from registry, correspondent)	Bargaining levels to sample from (SES 2014)
Austria	Registry	Available	Registry	SES 2014 not available
Belgium	Registry	Available	DK	National, industry, enterprise
Bulgaria	Registry	Available	Registry	Industry and enterprise
Croatia	Registry	DK	DK	SES 2014 not available
Cyprus	Registry	Available	Correspondent (registry CAs are scans)	Industry and enterprise
Czechia	Registry for MEB, snowballing for SEB	Available for MEB, API for SEB	Registry for MEB, correspondent for SEB	Industry and enterprise
Denmark	Snowball sampling from trade unions and employer organisations	API	Correspondent	Industry
Estonia	Registry	Available	Registry	Industry and enterprise
Finland	Registry	Available	Registry	National, industry, enterprise
France	Registry	Available	Registry	Industry and enterprise
Germany	Registry	API	Correspondent	Information on bargaining levels not included in SES 2014
Greece	Registry	Available	Registry	SES 2014 not available
Hungary	Registry	Available	Correspondent	Industry and enterprise
Ireland	Snowball sampling from IRN, ICTU	Unstructured	Correspondent	SES 2014 not available
Italy	Registry	Available	Registry	Information on bargaining levels not included in SES 2014

Preliminary framework for a Eurofound Collective Agreements Database
on collectively bargained minimum pay rates

Country	Sampling strategy	Metadata (use available data/make API)	Full text (online/offline from registry, correspondent)	Bargaining levels to sample from (SES 2014)
Latvia	Gazette publications (tbd)	API	Correspondent	Industry and enterprise
Lithuania	Registry	Available	Registry	Industry and enterprise
Luxembourg	Registry	Available	Registry	National, industry, enterprise
Malta	Registry (tbd) Potentially full population	API	Correspondent	Industry and enterprise
Netherlands	Registry	Available	Registry	No detailed SES breakdown by level
Poland	Registry	API	Correspondent	National and enterprise
Portugal	Registry	Available	Registry	National, industry, enterprise
Romania	Registry	DK	DK	National, industry, enterprise
Slovakia	Register for MEB, snowballing for SEB	Available for MEB, API for SEB	Registry for MEB, correspondent for SEB	Industry and enterprise
Slovenia	Register for MEB, snowballing for SEB	Available for MEB, API for SEB	Registry for MEB, correspondent for SEB	National and industry
Spain	Registry	Available	Registry	Industry and enterprise
Sweden	Medlingsinstitutet and Unionen	Available	Unionen	Information on bargaining levels not included in SES 2014

Note: DK, don't know.

Source: Authors' own elaboration

4. Coding parameters for the database

4.1 JOBS versus sectors: expected data quality and potential indicators

In the final stage of the construction of the Eurofound Collective Agreements Database, pay in the selected CAs (section 3) must be linked to the sampled JOBS or sectors (section 2). In order to do so, the pay system of the CAs must be studied. This includes the pay structure, which is reflected in the total of negotiated pay rates in all pay scale tables found in the CA, as well as premium pay (e.g. for performance or for night work). Pay structures vary substantially across countries and CAs. In general, CA pay structures set out either the minimum pay rates (wage floors) or exact pay rates that workers covered by the agreement are entitled to. Minimum pay rates may be applicable to all workers in the CA or vary by group, but bear a relatively loose relation to earned wages as more experienced or higher skilled workers are expected to earn well in excess of the minimum rates. Alternatively, CAs may contain pay structures including pay rates that are more prescriptive in the sense that the negotiated rates largely correspond to earned wages. Pay structures typically contain a series of pay rates that are stratified through a few or many pay grades (also known as pay bands) that set out which pay rates apply to different jobs, job families, career families and so forth (Armstrong, 2007).

The applicable pay rate, i.e. the single amount reflecting the (minimum) pay level for an employee in the job or sector, must then be extracted from this pay system. Strategies for doing so differ depending on whether the Eurofound Collective Agreements Database uses a JOB-based or sector-based approach. In the sector approach, the applicable rate is considered to be the lowest pay rate for workers in the selected sector. Thus, the lowest pay rate found in the CA is selected (excluding subminimum pay rates such as youth scales). The great advantage of this approach is that it makes almost no assumptions about the shape pay structures take, making it easily applicable across EU Member States and limiting the time spent on coding each CA. If information is available in the CA about the specific occupations that are linked to the lowest pay rate, this occupational information can be coded as well, but no ex-ante selection of occupations is made. It would thus be feasible to describe the occupations that pay rates refer to, but these would be different occupations in the respective countries, and comparisons between countries would thus be made at sector level. Comparability of pay rates might be hampered by the diversity of occupations to which the lowest pay rates refer. For example, the lowest pay rate found in the ICT sector in one country might apply to an entry-level programmer, while in others it might refer to cleaning jobs.

In the JOBS approach, the applicable pay rate is considered to be the pay range for workers in the selected JOBS. From previous experiences in coding the pay negotiated in CAs, we know that in practice it is rare for only one pay rate to apply to a job, and many CAs will contain multiple pay rates that could be linked to a job. This series of pay rates is referred to as the applicable 'pay range' in this paper.

Once the relevant pay range for a job has been identified using the JOBS approach, the consortium proposes to code the minimum, median and maximum rates within the applicable pay range. The minimum payable rate measures the wage floor set in the CA, which is of direct interest to the aim of the database to monitor collectively agreed minimum pay rates applicable to a selection of low-

(and medium-) paid JOBs. However, the minimum payable rate will underestimate the wages earned by workers covered by a CA, since it tends to apply to workers without any years of tenure. The median pay rate in the pay range, often referred to as the 'midpoint', is assumed in the human resources management literature to represent the target pay rate, which is the pay level that represents the value employers are willing to pay a worker who is fully qualified to carry out the job (Armstrong, 2007). Finally, the highest payable rate represents the pay ceiling and can be used to quantify the potential for pay progression within a job for the CA in question.

The advantages of the JOBs approach are its intuitive link to a comparable list of JOBs and its ability to incorporate the full pay range linked to this job, rather than just a minimum rate. If pay rates are linked to both sectors and occupations and these JOBs have been pre-selected ex ante (i.e. the same JOBs have been coded across countries), this enables Eurofound to gather comparable data. Since the JOBs approach offers the coder a clear frame, the minimum, median and maximum pay rates for the selected job(s) can all be coded without leaving the scope of low- and medium-paid jobs. This means that we have a range of pay rates for each job in countries where this is applicable, and can not only refer to the lowest pay rate workers in these JOBs are entitled to, but also give estimates about the negotiated pay range as a whole, with lower and upper boundaries. Doing this minimises the risk of coding pay rates that may be even lower but that are not effectively used (empty pay scales) as minimum pay rates.

The main disadvantages of this JOBs approach relate to its relative complexity and the difficulty of fitting it to diverse CA formats across EU Member States. The JOBs approach takes as its point of departure that CAs will set pay rates for different occupations and, in so doing, takes the most complex pay systems as its standard. Some countries will not negotiate pay ranges for JOBs, or will just set one minimum pay rate for all jobs covered by the CA. Therefore, data yielded following the JOBs approach might be far more detailed in some countries than others, and statistical analyses will need to take this uneven nature of the data into account. Finally, the JOBs approach puts a relatively heavy burden on coders, who will need to link pay rates to job classification systems, and therefore it requires more time for coding per CA than a less complex approach.

Applicable pay rates to be coded in the Eurofound Collective Agreements Database thus differ depending whether a sector or JOBs approach is chosen. In essence, this stage thus aims to answer the question 'What is the pay **range** negotiated for workers in the sampled job in this CA?' or 'What is the minimum pay **rate** negotiated for workers in the sampled sector in this CA?' This question should be answered for each job or sector when coding a CA, potentially more than once if multiple pay scale tables set out pay ranges for different time periods during the validity of the CA. The resulting database will take the CA as the unit of observation and will contain variables including the lowest, median and highest pay rates of the pay range for each job linked to the CA or the lowest pay rate for each sector linked to the CA.²²

4.2 Measuring metadata

The first group of parameters to be coded in the Eurofound Collective Agreements Database are the metadata used for the selection of CAs that can be linked to the sampled JOBs or sectors. Therefore,

²² The CA is taken as the unit of observation because this is the most stable unit and because it is the sampled unit. Statistical techniques allow researchers to switch between 'long' and 'wide' data formats easily according to the needs of the analyses.

these parameters need to be collected for all CAs archived in the national registries. In countries where snowball sampling is implemented for all or some of the selected CAs, these parameters will need to be coded manually.

The parameters included in the CA metadata are **CA name**, **bargaining level**, (estimated) **number of employees covered**, **signatories**, **start date**, **end date and duration** (computed), and **NACE** code, or **ISCO** code in the case of an occupational CA. In addition, the coding scheme assigns a **unique identifier number** to each coded CA. In the future, this will allow Eurofound to identify whether newly added CAs are renewals of a CA that is already in the database and, if so, of which CA. Finally, to be able to estimate the number of workers covered by each CA, we propose to code the **exclusion/inclusion** of specific groups of employees in the CA.

Here we propose to follow the coding scheme developed for the WageIndicator CA database (Ceccon and Medas, 2022). This coding scheme has been used for the coding of 532 CAs in the EU27 and has proven to be able to capture the variety in bargaining contexts, and in cases of nested CAs it allows the related CAs to be identified. In this coding scheme the variables are coded as follows.

CA name is a text field that includes the title of the CA and the years covered. For example, a CA name could be 'IKEA Netherlands 2018-2020'.

The **identifier number** will be a numeric indicator including the country code, number of the CA (in order to include in the database) and a rank ID, which starts at 01 and counts upwards to CAs that are linked. To be able to add renewed CAs, coders are asked 'Is this a renewal of an existing (but expired) CA included in this database?' and, if so, are asked to indicate to which CA the renewal should be linked. To be able to link separate appendices to existing CAs, coders are asked 'Is this CA an additional part of an existing CA?' and they can link the appendix to the relevant ID number.

Start date, end date and duration are measured by two items. The first two ask 'Is an operative start (end) date mentioned?' Coders enter the start and end dates (dd/mm/yyyy) where relevant. Duration of the CA in months can be computed from these two items.

Bargaining level E is coded using a number of items that ask whether the CA covers a single enterprise or multiple enterprises. As Table 8 shows, the coding scheme asks whether the CA was concluded with a single employer, multiple employers, one or more employers' associations, one or more trade unions, one or more professional associations and any other **signatories** (e.g. works council). Coders reply yes or no to each of these items. A variable distinguishing SEB and MEB can then be constructed.

Table 8: Questions to ascertain bargaining level and names of signatories

Is the agreement concluded with a single employer/company/institution or multiple employers?	1 'a single employer/company/institution' 2 'multiple employers'
Which of the following actors are signatories to the agreement? (tick all applicable options)	0 'None' 1 'one or more single employer' 2 'one or more employers' associations' 3 'one or more trade unions' 4 'one or more professional associations' 5 'one or more works councils' 6 'other, namely...'

We propose to collect data about the **number of employees covered** by the selected CAs for the years in question. Some registries keep records of the (estimated) number of workers covered by a CA, which can be used if available and are the preferred source of data. A small number of CAs include references to the number of workers covered. However, we assume that for quite a few selected CAs this information will be missing, when neither the CA text nor the registries provide this data. In the case of sector CAs, estimates can be made centrally using the EU-LFS or SES data, but these cannot provide reliable estimates for enterprise-level CAs. Therefore, we propose that additional efforts be made to find this information from the enterprise’s annual reports, from public or commercial registries, or via Eurofound correspondents. To reflect the different data sources used for this variable, we propose to add two variables that capture the source of the data and whether the data is based on exact information or an estimate.

The **exclusion/inclusion** of specific groups of employees from the scope of the CA is measured using the CA text (Table 9). We propose to ask specifically for the groups of workers that are most often excluded from CA provisions (part-time workers, temporary agency workers, apprentices, students and mini-jobs) and, if so, whether these groups are excluded from the wage provisions. A fifth item contains a text field for coders to add any other groups excluded from the scope of the CA. These data can help improve estimates of the number of workers covered under a CA, where these are derived from the EU-LFS or from the SES.

Table 9: Questions to ascertain excluded groups from the collective agreements

Are any groups of workers explicitly excluded from the pay rates determined by the CA? Tick all that apply

	Excluded from any provision of the CA	Excluded from wage provisions of the CA
0 'No'		
1 'Yes, part-time workers'		
2 'Yes, temporary agency workers'		
3 'Yes, apprentices'		
4 'Yes, mini-jobs or student jobs'		
5 'Other, namely...'		

Sector is measured using the **NACE** industry classification as the most detailed level available, but at least at the two-digit level required for the matching of CAs to JOBS (Table 10). For this step, we propose to use the NACE API with lists of NACE Rev. 2 coded industries, available in all EU languages, freely available from the survey codings website : [Articles \(surveycodings.org\)](https://surveycodings.org)). The detailed NACE coding will make it possible to identify nested CAs (hence MEB and related SEB CAs) and resolve issues of favourability and opt-outs based on the CA texts or on country-level information about bargaining contexts. In addition, we propose to add one item to be able to distinguish public sector CAs from private sector CAs. These data can help improve estimates of the number of workers covered under a CA, where these are derived from the EU-LFS or SES.

Table 10: Questions to ascertain which industry is addressed in the collective agreements

What is the industry covered by the CA?	NACE API
What is the more detailed industry (NACE2004) covered by the CA?	NACE API
Is the private or the public sector covered by the CA?	1 'Private' 2 'Public' 3 'Semi-public' 4 'Other incl. non-profit, co-op' 5 'I don't know'

4.3 Measuring occupations

The degree of importance of coding occupations differs between the sector-based and JOBs-based approaches, as is explained in section 4.4, but both measure occupation in the same manner (ISCO classification). In order to link the pay scales in the CAs to the sampled JOBs, occupations need to be coded. In the sector-based approach, occupations are coded post hoc, as a way of providing additional information on the occupations covered by the lowest pay rate, where available. In the sector approach, coding occupations thus serves to provide insight into the types of workers who are entitled to the lowest identified pay rate, whereas in the JOBs approach occupations are essential for identifying the pay rates to be coded.

Metadata linking CAs to ISCO codes are unlikely to be available from the national registries, with the exception of occupational CAs, which are rare and only exist in a few countries. Occupational CAs can be identified using two items that are similar to the measurement of the inclusion or exclusion of specific groups of workers. First, the coders are asked whether any occupational groups are explicitly included or excluded from the scope of the CA. If so, coders are asked which occupations. For this last step, we propose to use the ISCO API in all EU languages with long lists of ISCO coded occupational titles, freely available from the survey codings website: <https://www.surveycodings.org/occupation-measurement>). These items serve to match entire CAs to JOBs.

Table 11: Questions to ascertain which occupation(s) is/are addressed in the collective agreements

Does the agreement explicitly INCLUDE specific occupations for coverage?	1 'Yes' 0 'No'
Which occupations?	ISCO API
Does the agreement explicitly EXCLUDE specific occupations for coverage?	1 'Yes' 0 'No'
Which occupations?	ISCO API

In all other CAs, which we estimate will be almost all CAs, pay scales will need to be matched to JOBs or to the lowest pay rates in a CA manually. This will only be done for CAs where pay scales are

related to a job classification system, if these systems identify occupational titles or jobs, such as drivers or checkout operators, and thus go beyond enterprise specific abbreviations or numbering. If the CA pay rates cannot be ISCO classified because job descriptions are absent or not codable, the JOBs –CA matching will be based on NACE codes only. To determine whether occupation-level matching is possible, coders are asked whether a job classification system or job titles are included in the CA or its appendices.

Table 12: Questions to ascertain whether job descriptions are addressed in the collective agreements

JOBs approach	
Occupational title of job	Pre-filled ISCO-08
Does the CA include job titles or job descriptions, or refer to a job classification system?	1 'Yes' 0 'No'
Is a pay rate or pay range specified for the [occupational component job]? (N.B. question used for routing)	0 'No, no match with any listed occupations' 1 'No, occupation is listed but not linked to pay scales' 2 'Yes, pay rate' 3 'Yes, range'
Sector approach	
Does the CA include job titles or job descriptions, or refer to a job classification system?	1 'Yes' 0 'No'
Are any job titles or job descriptions linked to the lowest pay rate?	1 'Yes' 0 'No'
For which jobs?	ISCO API

We propose, for the relevant CAs, to undertake a manual ISCO-08 two-digit coding of the occupational titles specified in the job classification system, where possible based on the job descriptions provided in the CA. The consortium can help to develop text recognition techniques for automatic ISCO coding of job titles and descriptions, where needed supplemented with manual coding. The consortium will build on Python-based techniques used to classify job titles in vacancies into ISCO-08 (Tijdens and Kaandorp, 2019), as well as other Python-based techniques that are yet to be developed to classify job titles from a large pool of job descriptions in almost all EU27 languages. This job–ISCO mapping may result in one CA mapping to more than one job.

4.4 Measuring pay rates

In the next step, JOBs need to be matched to specific pay rates in the CA. This stage aims to detail all selected CAs with respect to the wages agreed in the CAs. This section in the coding scheme includes the parameters related to wages in collective agreements. In the wage section, we propose to code the inclusion of a minimum wage amount (wage floor) in the CA, in the case of a sector-based

approach, and the minimum, median and maximum pay rates associated with a job, in a JOBs-based approach. The coding scheme also needs to capture other pay elements, such as holiday pay or 13th/14th month.

First, we propose to code a number of general indicators regarding wages in CAs (Table 13). We code whether CAs include any **clauses on wages** at all. This variable is primarily used for routing. Any CAs that do not will receive missing values on all other wage-related variables. Second, coders are asked to specify whether the CA provides any information on **where wages are determined**. This will help identify cases of opt-outs, articulated bargaining and favourability rules. Third, we code whether the CA sets a **wage floor**, either explicitly or in reference to the statutory minimum wage (in countries where one exists). In CAs without pay scale tables, this wage will be coded as the lowest pay rate for the selected job or sector in the CA. Fourth, **wage increases** are coded.

Table 13: Questions to indicate whether and how wages are included in the collective agreements

Does the agreement have clauses on wages?	1 'Yes' 0 'No'
According to the CA, where are wages determined?	1 'Individual contracts' 2 'Industry/sectoral level' 3 'National framework agreement' 4 'State or regional level' 5 'Elsewhere' 6 'Company level' -7 'Insufficient data'
Does the agreement set a minimum/lowest pay rate?	1 'Yes' 0 'No'

A second set of parameters on wages code the pay scale tables of the CA, in those cases where pay scale tables are included (Table 14). First, we code whether **one or multiple pay scale tables** are included in the CA. Our past research indicates that, in CAs that include pay scales, there is often more than one table. Coders identify these CAs where multiple tables are included and indicate how the tables are different from each other. In our past research, we found there are four main ways in which multiple pay scales in a single CA tend to differ from each other.

- Some pay scale tables only differ with regard to the payment period (e.g. pay rates per hour, per month, for a 36-hour working week, for a 38-hour working week, etc). In this case, we propose to select a single table for coding because the information is equivalent.
- Some CAs will include multiple pay scale tables associated with consecutive pay increases negotiated within the duration of the CA. For example, a pay rate might be set at €1,500 from 1 January 2020, at €1,515 from 1 July 2020 and at €1,530 from 1 January 2021. In these

cases, we propose to code the pay rate data from all tables and their operative start and end dates with the aim of creating a time series²³.

- Some pay scale tables will include subminimum pay rates (e.g. for young workers, long-term unemployed people or apprentices). We propose not to code these tables in the first wave of coding for the Eurofound Collective Agreements Database, since past research indicates they are only included in a small number of countries. Since the coding grid does identify CAs with subminimum pay rates, back-coding will be possible in later waves if deemed appropriate.
- A last set of differentiations is based on substantive differences (e.g. different operations of the enterprise such as distribution and sales activities or different occupational groups). These differentiations contain crucial information for the linking of CA pay rates to JOBs and should be manually matched to the selected job. In the case of a sector-based approach, pay rates from all these pay scale tables should be considered when identifying the lowest pay rate.

In a sector-based approach, coders will need to select the lowest pay rate in the CA from all pay scale tables that are substantially different from each other (e.g. considering pay scale tables for both the distribution and store career families in a supermarket CA) and ignore pay scale tables or pay grades containing only subminimum rates, as well as those tables containing the same pay rate data expressed in different pay period units (e.g. hourly and monthly pay rates). In cases where multiple pay scale tables are valid at different time periods, the selected lowest pay rate needs to be identified and coded for each time period, including data on the start and end dates of each part of the time series.

In a JOBs-based approach, matching might be to the pay rates within a pay scale table or across pay scale tables. When matching pay rates to JOBs in the CA pay system, we ask coders to consider **pay grades differentiated** by occupation, job title/type, skill level, tenure and geographical region. This can refer to differentiations within a single pay scale table or between multiple pay scale tables. Differentiations by occupation, job, skill level, education and subsector can be used to manually match the appropriate pay rate(s) to the selected job. Differentiations based on tenure can be included in upper and lower range estimates of pay rates. In regard to all other differentiations, we propose to code the pay rates that apply to regular, adult workers in medium-sized or large firms.

In the coding scheme, we propose to code the differentiation of pay scale tables identified in the CA for the purpose of consistency checks within and across different waves of the database. In a second wave of data collection, coders can follow the choices of the first wave coders. Furthermore, recording these choices makes it possible to introduce routing in the coding scheme: coders who indicate that a CA contains more than one table and that pay rates in different tables are valid at different dates will be asked how many different time periods were found and will be presented with the appropriate number of fields to code the pay rates as a time series.

²³ Coding all three tables, as well as their start and end dates is possible but will require more effort.

Table 14: Questions to ascertain whether and how pay scales are differentiated

Are wages determined in the CA by means of pay scales?	<p>1 'Yes, in one table'</p> <p>2 'Yes, in more than one table'</p> <p>3 'Yes, but there are only indices (no wages)'</p> <p>0 No</p>
In what respect are pay scale tables different? (multiple answers possible)	<p>1 'Age groups'</p> <p>2 'Geographical regions '</p> <p>3 'Validity at different dates '</p> <p>4 'Specific groups of workers (such as apprentices, disabled people or previously unemployed) '</p> <p>5 'Related to workers' educational level '</p> <p>6 'Sub-sector or units in companies'</p> <p>7 'Related to job types or activities'</p> <p>8 'Related to workers' years of tenure'</p> <p>9 'Related to firm size '</p> <p>10 'Related to skills required for an activity'</p> <p>11 'Differentiated otherwise (please specify how)'</p>
How many different time periods to the pay scale tables refer to? (if pay scale tables are valid at different dates)	Number field

Once the relevant pay range or pay rate has been identified, the consortium proposes to code the minimum, median and maximum rates within the applicable pay range associated with a job, or the minimum pay rate associated with a sector (Table 15). The minimum pay rate measures the wage floor set in the CA, which is of direct interest to the aim of the database to monitor 'collectively agreed minimum wage rates applicable to a selection of low- (and medium-) paid JOBs.

If pay scales tables or pay grades are broken down for groups of employees, only the relevant pay rates are taken into account. For example, when coding the pay rates of cleaners in the cleaning industry, rates for cleaners with little and much experience are coded, but the pay rates for managers and fore(women) are excluded. If no data are available on the differentiation within the pay scale table, the minimum, median and maximum pay rates across the pay scale table are coded. In CAs that fix a wage floor but do not set out pay scales, the wage floor will be coded as the lowest wage. In addition, the start and end dates of the time period to which the pay rate data refers and the calculation period (e.g. hourly, monthly wages) are coded. If a choice is made to code this information for multiple pay scale tables with different start and end dates, the coding described in

this paragraph is repeated for each time period. In the sector-based approach, information about the job characteristics associated with the lowest pay rate is coded when available in the CA text.

Table 15: Questions to ascertain the minimum, median and maximum amounts in the pay range for the relevant job or sector

Coding parameters for JOBs-based approach	
What is the lowest payable rate for adults in [job]?	number field
What is the median payable rate for adults in [job]?	number field
What is the highest payable rate for adults in [job]?	number field
To which dates does this pay range refer?	dd/mm/yy start date dd/mm/yy end date (if applicable)
What is the calculation basis for the wages in the pay range?	1 'Days' 2 'Weeks' 3 'Two weeks' 4 'Years' 5 'Months' -1 'Hours'
Coding parameters for sector-based approach	
What is the lowest payable rate for adults in [job]?	number field
Is this pay rate explicitly associated with a specific job type in the CA text? (tick all that apply) Prompt to coder in cases 4 and 6: Please ignore subminimum pay rates when selecting pay rate	1 'Occupation/job title/job type' 2 'Skill levels' 3 'Tenure (years of experience)' 4 'Age' 5 'Geographical regions' 6 'Specific groups of workers (such as, apprentices, disabled people or previously unemployed people)' 7 'Educational level' 8 'Subsectors' 9 'Firm size' 10 'Other' 11 'No information in CA'
Which occupation? (if selected above)	ISCO API

What skill level? (if selected above)	ISCO skill levels API
How many years of tenure? (if selected above)	Number field
Which region? (if selected above)	Region API
What educational level? (if selected above)	Education API
Which subsector? (if selected above)	NACE API
What firm size? (if selected above)	[determine labels]
To which dates does this pay rate refer?	dd/mm/yy start date dd/mm/yy end date (if applicable)
What is the calculation basis for the wages in the pay rate?	1 'Days' 2 'Weeks' 3 'Two weeks' 4 'Years' 5 'Months' -1 'Hours'

Finally, to be able to make reliable estimates of wages determined in the CAs, several other elements of wages are coded. These include compulsory annual bonuses, such as holiday pay, compulsory annual bonuses, such as 13th/14th month, and meal vouchers' (Table 16).

Table 16: Questions to ascertain whether extra payments are agreed

Is an extra payment for the paid annual leave agreed?	1 'Yes' 0 'No'
How is the extra payment given for a worker with one year of experience?	1 'Percentage of regular wage' 2 'Amount' 3 'Both percentage of regular wage and amount' 4 'Percentage of official minimum wage' 5 'Number of days of regular wage' 6 'Hours/days off' -7 'Insufficient data' -10 'The CA explicitly refers to the law'
Extra payment (% of basic wage):	number field
Extra payment (amount):	number field

Number of working days:	number field
Are meal allowances agreed?	1 'Yes' 0 'No'
What is the minimum amount of the allowance for one meal?	number field
Is an extra payment at the end of the year agreed (13th/14th month)?	1 'Yes' 0 'No'
Extra payment (% of basic wage):	number field

4.5 Measuring working hours

For comparability, an hourly wage has to be computed once the pay ranges have been matched to the JOBS. The final section in the coding scheme therefore captures the length of the working week. We propose to capture the length of the regular working week, as well as the number of paid days of annual leave plus paid public holidays, to compute the number of annual working days per CA (Table 17).

Table 17: Questions to ascertain standard working hours

Does the agreement have clauses on standard working hours, schedules, holidays and/or days of leave?	1 'Yes' 0 'No'
Does the CA have clauses on standard working hours per day?	1 'Yes' 0 'No'
Does the CA have clauses on standard working hours per week?	1 'Yes' 0 'No'
Does the CA have clauses on standard working hours per month?	1 'Yes' 0 'No'
Does the CA have clauses on standard working hours per year?	1 'Yes' 0 'No'
Does the CA have clauses on standard working days per week?	1 'Yes' 0 'No'
How many working hours per day are agreed in the CA? (1–12)	number field
How many working hours per week are agreed in the CA? (1–60)	number field
How many working hours per month are agreed in the CA? (1–258)	number field
How many working hours per year are agreed in the CA? (1–3,096)	number field

How many working days per week are agreed in the CA? (1–7)	number field
Does the CA have clauses on paid annual leave?	1 'Yes' 0 'No'
How many days of paid annual leave are agreed for a worker with one year of service in the CA? (1–100)	number field
How many weeks of paid annual leave are agreed for a worker with one year of service in the CA? (1–20)	number field
Is paid leave agreed for specific public holidays in the CA?	1 'Yes' 0 'No'
For which public holidays?	API

These parameters will be used to compute the annual working days. If data on working hours are not present in the CA, we propose to use the length of the standard working week in the country and the standard number of holidays in the country. These data can be added centrally for all CAs where coders indicate data on working time are missing. Eurofound correspondents can be asked to provide these figures for their country once, so they do not need to enter them in every coded CA. The data collected for the above-mentioned parameters will allow us to assess for how many CAs we are able to compute the hourly wages agreed in the CAs. Based on the assessment of the working hours, comparable minimum, median and maximum hourly rates of the relevant pay scales can be computed.

5. Discussion and conclusions

5.1 Two approaches for the Eurofound Collective Agreements Database

In summary, we propose to collect data on collectively negotiated pay rate in low- (and medium-) paid JOBS for a CA database that is sustainable, scalable, representative and comparable. We propose using a three-stage sampling strategy to sample low- (and medium-) paid JOBS or sectors, select matching CAs, and link pay ranges to JOBS or sectors. Based on exploratory analyses, both the JOBS approach and the sector approach present feasible strategies for the development of the Eurofound Collective Agreements Database.

In the JOBS approach, CAs are selected by sampling low-paid, and potentially medium-paid, JOBS (i.e. occupations nested in sectors, for example a cleaner in the cleaning industry or a shelf stacker in the retail sector) using microdata from the EU-LFS and SES. CAs covering the sector component of the job are selected using national CA archives, and applicable pay rates/ranges are coded from among all the pay rates in the agreement based on the occupational component of the job. In the sector approach, low-paid (sub)sectors are sampled based on EU microdata, CAs covering those same (sub)sectors are selected using national CA archives and the lowest pay rates found in those CAs are coded.

To sample JOBS or sectors, an EU ranking of JOBS/sectors with the lowest average income decile is constructed by averaging the national rankings of JOBS/sectors weighted by the employment size. In the JOB approach, we then propose to select 10 out of 200 JOBS for inclusion in the Eurofound Collective Agreements Database, and we suggest these be selected based on the share of workers and the collective bargaining coverage in these JOBS. Calculations in this step of the sampling strategy can be performed based on reliable datasets (EU-LFS and SES) and, due to the decision to sample the same JOBS across countries for comparability, missing country data constitute a minor obstacle. For the purpose of the conceptual note, we chose to include 10 out of a list of 200 low- to medium-paid that represents 30% of employment in the EU; however, this threshold may be adapted to match the resources and ambitions of the Eurofound Collective Agreements Database.

In the second stage, each selected job or sector in each EU Member State is matched to CAs with the use of CA metadata or snowball sampling. Since the matching in the JOBS approach is based on the sector component of the job, this stage is the same across the two potential strategies. The sampling of CAs to measure the development of collectively agreed wage floors in Europe will have to be differentiated by country for the foreseeable future. In a number of countries, representative or near-representative sampling of CAs is already a possibility. These include countries with well-developed archives maintained by national registries as well as countries with small numbers of CAs, facilitating the coding of all CAs. In the short term, any CA database will have to contend with these disparities. In the long term, cooperation between Eurofound and the national registries should adopt as one of its aims to improve and normalise the data collection, coding of metadata and archiving of CAs in the Member States, which would increase the number of countries where representative sampling strategies may be implemented.

Due to data limitations, we propose to implement a snowball sampling strategy for the selection of Danish, Maltese and Irish CAs, as well as for SEB CAs from Czechia, Latvia, Slovakia and Slovenia. In

all other countries, we recommend selecting CAs based on structured, or if necessary unstructured, metadata from national CA registries. This recommendation is dependent on access to said archives, which is the aim of work package 2 of this project, and may therefore be subject to change.

The metadata from the registries are used to identify and link CAs to sectors or to JOBcountry cells. First, the start and end dates, bargaining level (sectoral/enterprise CA²⁴), (estimated) number of employees covered, NACE code and, when applicable, ISCO²⁵ code of these CAs are required to be able to link the CAs in the registry to the selected JOBS. Second, the name of the CA and, where possible, names of signatories are required to identify the selected CA and retrieve its full text.

To select CAs from all linked CAs, in case these are too many to code, we propose a selection strategy based on bargaining levels in the different countries and sectors. In countries where over 80% of the labour force in a job or sector is covered by **sector-level agreements** according to SES data, we propose to select sector-level CAs only. In countries where bargaining in a job or sector only takes place at enterprise level according to SES data, the Eurofound Collective Agreements Database will sample from these agreements exclusively. In cases where **both enterprise- and sector-level bargaining** are conducted and enterprise-level bargaining covers more than 20% of the labour force, we propose to select both sector and enterprise CAs proportional to the distribution found in the SES.

We propose to code four groups of parameters: metadata needed to link CAs to JOBS; occupation data to link pay rates to JOBS; pay scale tables and other wage data needed to estimate the level of pay; and data on working hours to be able to calculate comparable pay rates. Parameters to be coded in the Eurofound Collective Agreements Database are the metadata used for the selection of CAs that can be linked to the sampled JOBS. The parameters included in the CA metadata are **CA name, bargaining level, (estimated) number of employees covered, signatories, start date, end date and duration** (computed), and **NACE** code, or **ISCO** code in the case of an occupational CA. Occupational CAs can be identified using two items that are similar to the measurement of the **inclusion or exclusion** of specific groups of workers.

Pay rates or pay ranges will need to be matched to JOBS manually in the case of the JOBS approach. Occupations are coded for CAs where pay rates are related to a job classification system, if these systems identify occupational titles. In the section on measuring pay rates, we propose to code the minimum, median and maximum pay rates in the pay range linked to the job, holiday pay and compulsory annual bonuses, such as 13th/14th month ; or the minimum pay rate for the sector-based approach, and any information on job characteristics associated with this pay rate (e.g. occupational title). The final section in the coding scheme captures the length of the working week, including items that measure the length of the regular working week, as well as the number of paid days of annual leave plus paid public holidays, to compute the number of annual working days per CA. The collected data for the above-mentioned parameters will allow us to assess for how many CAs we are able to compute the hourly wages agreed in the CAs. Based on the assessment of the

²⁴ These are abbreviated as respectively multi-employer bargaining (MEB) CAs and single-employer bargaining (SEB) CAs.

²⁵ Note that most CAs will be NACE-based CAs, covering a sector or an enterprise. However, there will also be ISCO-based CAs, covering one or more occupational groups, e.g. airline pilots or domestic workersCACACA

working hours, comparable minimum, median and maximum hourly rates of the relevant pay rates can be computed.

5.2 Indicators derived from the two approaches

From a technical point of view, both the sector and JOBs approaches are feasible. Both approaches would result in a database that is scalable (i.e. more JOBs or sectors can be added if greater resources are available). Both would also be able to produce a database addressing the aim of the envisaged Eurofound Collective Agreements Database: to monitor collectively agreed minimum wage rates applicable to a selection of low- (and medium-) paid JOBs over time. This would thus result in indicators measuring collectively negotiated pay in a range of occupations or sectors (e.g. what minimum pay is negotiated for shelf stackers, plumbers, kindergarten minders, etc., or for workers in supermarkets, the cleaning industry, etc).

Both approaches allow for the calculation of the lowest agreed pay rate, which can be used to generate insights into the relation between it and the statutory minimum wage. Coding the lowest pay rate can help assess in which countries negotiated and statutory wage floors are set at the same levels, where negotiated pay rates start higher, whether the lowest pay rate varies much within countries, and whether there is a de facto inter-industry wage floor in countries without a statutory minimum wage. Furthermore, the lowest pay rates in JOBs or sectors can be compared with living wages or at-risk-of-poverty thresholds. The time series of pay rates may also be used to calculate pay increases in a comparable way. Pay increases are vital information on collective bargaining and may also be used as an indicator in the reporting on the Eurofound Collective Agreements Database.

Both the JOBs approach and the sector approach also have their advantages and drawbacks. Crucially, they would result in different types of indicators. The JOBs approach would result in estimates of the collectively bargained pay levels of workers in specific JOBs (occupations nested in sectors), including upper and lower boundaries where available and weighted by the number of workers covered by each CA where multiple CAs cover one job. The sector approach would result in estimates of the minimum pay levels all workers in a sector are entitled to, without being able to estimate how many workers this rate is applicable to. If occupations linked to that pay rate are included in the CA and are coded, these could be reported illustratively, but not in a systematic way.

Even more important, however, is that the choice of the sector or JOBs approach should depend on the data quality it yields in practice. Within the framework of this pilot project for the Eurofound Collective Agreements Database, two piloting phases are planned at the time of writing this working paper (Q1 2022): in the first half of 2022, the project planning envisages a small pilot of the methodology coding a small number of JOBs (or sectors) in five countries. After this initial, small-scale pilot, a second pilot is planned, in which CAs from all EU Member States are coded (to be finished in the first quarter of 2023). The exact shape of the second piloting phase (e.g. how many JOBs/CAs) is dependent on the results of the first small-scale pilot.

We propose to use the first small-scale pilot to gain more insights into the difference in data quality using the JOBs approach versus the sector approach. We would do so in the following way.

- Eight countries would be selected for the small-scale pilot based on their different bargaining practices.

- We would use EU-LFS and SES data to sample four low-paid JOBS across four different sectors following the JOBs approach.
- CAs linked to those JOBs would be identified, based on the sector component of the job, as described in the methodological note, and would therefore be similar across the two approaches.
- Coders would then code the CAs according to both approaches at once: in each CAs the coders would (1) select the minimum, median and maximum pay rate linked to the job based on the occupational component of the job, as proposed in the methodological note, and (2) select the lowest pay rate in the CA, excluding subminimum pay rates, as proposed in the sector approach. CAThey would code the occupational title(s) linked to the lowest pay rate in the CA, if available in the CA text.

This would result in two sets of data points: one following the JOBs approach and one following the sector approach. This would allow us to evaluate the feasibility of linking pay rates to specific JOBs across countries, while also evaluating the comparability of the lowest pay rates in CAs on a sector basis only. Following the first small-scale pilot, the data from the two approaches will be compared, as will the sampling/coding effort required and the feasibility of both. The results from the small-scale pilot will inform the strategy for the larger pilot.

References

- Armstrong, M. (2007), *A handbook of employee reward management and practice*, 2nd ed., Kogan Page Limited, London.
- Charni, K., Greenan, N., Bruno, A. S., Cardoso, J., Cursi, Y. and Tanguy, J. (2020), 'Hunting for fairness: A text mining approach to assess gender equality efforts in firm and industry-level collective bargaining agreements', 2020 SASE Annual Conference, 21 July, Amsterdam. <https://sase.confex.com/sase/2020/meetingapp.cgi/Paper/15525>
- Ceccon, D., Medas, G. (2022). *Codebook Wage Indicator Collective Agreements Database – Version 5*, Wage Indicator Foundation, Amsterdam
- Eurofound and European Commission Joint Research Centre (2019), *European JOBS Monitor 2019: Shifts in the employment structure at regional level*, European JOBS Monitor series, Eurofound, Dublin. <https://www.eurofound.europa.eu/nb/publications/report/2019/european-jobs-monitor-2019-shifts-in-the-employment-structure-at-regional-level>
- Eurofound (2022). Summary of national databases and archives for Eurofound Collective Agreements Database.
- Garnero, A., Rycx, F. and Terraz, I. (2020), 'Productivity and wage effects of firm-level collective agreements: Evidence from Belgian linked panel data', *British Journal of Industrial Relations*, Vol. 58, No. 4, pp. 936–972.
- Nicolitsas, D. (2020), 'Wage-setting practices in the metalworking sector in Greece', *Employee Relations*, Vol. 42, No. 6, pp. 1441–1463.
- OECD and AIAS (2021), *OECD/AIAS ICTWSS database*, OECD Publishing.
- Tijdens, K. (2021), *Feasibility of a Europe-wide data collection of collective agreements*, Colbar-Europe Report 10, WageIndicator Foundation, Amsterdam.
- Tijdens, K. and Kaandorp, C. (2019), *Classifying job titles from job vacancies into ISCO-08 and related job features – The Netherlands*, Deliverable 8.5 of the Seriss project funded under the European Union's Horizon 2020 Research and Innovation Programme GA No. 654221. Available at: www.seriss.eu/resources/deliverables
- Tijdens, K., Besamusca, J., Ceccon, D., Cetrulo, A., Van Klaveren, M., Medas, G. and Szüdi, G. (2022), 'Comparing the content of collective agreements across the European Union: Is Europe-wide data collection feasible?' *E-Journal of International and Comparative LABOUR STUDIES*, Vol. 11,(No. 02).
- van Klaveren, M. and Gregory, D. (2019), *Restoring multi-employer bargaining in Europe: Prospects and challenges*, ETUI, Brussels.
- Yerkes, M. and Tijdens, K. (2010), 'Social risk protection in collective agreements: Evidence from the Netherlands', *European Journal of Industrial Relations*, Vol. 16, No. 4, pp. 369–383.

Appendix I: Full list of JOBs

Rank by low-income pay	NACE	NACE label	ISCO	ISCO label	CA coverage (%)	Group as a % share of total employment	% of group with earnings in lowest 3 deciles	% of group with earnings in lowest 5 deciles
1	17	Human health and social work activities	82	Assemblers	31	0.02	0.94	0.97
2	5	Water supply, sewerage, waste management and remediation activities	92	Agricultural, forestry and fishery labourers	85	0.00	0.97	0.99
3	7	Wholesale and retail trade; repair of motor vehicles and motorcycles	92	Agricultural, forestry and fishery labourers	45	0.02	0.73	0.95
4	17	Human health and social work activities	93	Labourers in mining, construction, manufacturing and transport	84	0.07	0.81	0.92
5	17	Human health and social work activities	61	Market-oriented skilled agricultural workers	70	0.02	0.82	0.93
6	19	Other service activities	53	Personal care workers	75	0.04	0.74	0.93
7	8	Transportation and storage	92	Agricultural, forestry and fishery labourers	12	0.00	1.00	1.00
8	17	Human health and social work activities	92	Agricultural, forestry and fishery labourers	76	0.01	0.75	0.94
9	14	Administrative and support service activities	95	Street and related sales and service workers	4	0.00	0.71	1.00
10	16	Education	94	Food preparation assistants	57	0.03	0.81	0.98
11	7	Wholesale and retail trade; repair of motor vehicles and motorcycles	95	Street and related sales and service workers	33	0.01	0.54	0.72
12	6	Construction	92	Agricultural, forestry and fishery labourers	66	0.00	0.95	0.97
13	17	Human health and social work activities	75	Food processing, wood working, garment and other craft and related trades workers	83	0.02	0.79	0.88
14	19	Other service activities	91	Cleaners and helpers	69	0.07	0.62	0.88
15	19	Other service activities	92	Agricultural, forestry and fishery labourers	87	0.00	0.67	0.83
16	9	Accommodation and food service activities	94	Food preparation assistants	47	0.35	0.58	0.90
17	14	Administrative and support service activities	91	Cleaners and helpers	58	0.56	0.55	0.87

Preliminary framework for a Eurofound Collective Agreements Database
on collectively bargained minimum pay rates

Rank by low-income pay	NACE	NACE label	ISCO	ISCO label	CA coverage (%)	Group as a % share of total employment	% of group with earnings in lowest 3 deciles	% of group with earnings in lowest 5 deciles
18	17	Human health and social work activities	91	Cleaners and helpers	72	0.50	0.57	0.90
19	15	Public administration and defence, compulsory social security	92	Agricultural, forestry and fishery labourers	44	0.05	0.74	0.88
20	17	Human health and social work activities	94	Food preparation assistants	74	0.08	0.55	0.85
21	14	Administrative and support service activities	94	Food preparation assistants	53	0.00	0.29	0.94
22	7	Wholesale and retail trade; repair of motor vehicles and motorcycles	91	Cleaners and helpers	50	0.10	0.58	0.87
23	3	Manufacturing	92	Agricultural, forestry and fishery labourers	53	0.01	0.60	0.86
24	19	Other service activities	96	Refuse workers and other elementary workers	88	0.01	0.55	0.81
25	16	Education	91	Cleaners and helpers	55	0.33	0.52	0.89
26	16	Education	81	Stationary plant and machine operators	49	0.00	0.61	0.84
27	18	Arts, entertainment and recreation	91	Cleaners and helpers	57	0.02	0.63	0.90
28	13	Professional, scientific and technical activities	91	Cleaners and helpers	66	0.00	0.61	0.78
29	12	Real estate activities	91	Cleaners and helpers	65	0.02	0.60	0.81
30	19	Other service activities	51	Personal service workers	62	0.49	0.58	0.83
31	9	Accommodation and food service activities	52	Sales workers	54	0.16	0.55	0.84
32	16	Education	93	Labourers in mining, construction, manufacturing and transport	82	0.00	0.48	0.75
33	9	Accommodation and food service activities	91	Cleaners and helpers	47	0.22	0.49	0.80
34	14	Administrative and support service activities	92	Agricultural, forestry and fishery labourers	64	0.05	0.50	0.83
35	12	Real estate activities	42	Customer services clerks	62	0.01	0.51	0.73
36	9	Accommodation and food service activities	81	Stationary plant and machine operators	55	0.02	0.57	0.87
37	19	Other service activities	81	Stationary plant and machine operators	90	0.06	0.48	0.83
38	9	Accommodation and food service activities	53	Personal care workers	72	0.00	0.38	0.71

Preliminary framework for a Eurofound Collective Agreements Database
on collectively bargained minimum pay rates

Rank by low-income pay	NACE	NACE label	ISCO	ISCO label	CA coverage (%)	Group as a % share of total employment	% of group with earnings in lowest 3 deciles	% of group with earnings in lowest 5 deciles
39	17	Human health and social work activities	96	Refuse workers and other elementary workers	69	0.03	0.53	0.82
40	3	Manufacturing	91	Cleaners and helpers	57	0.10	0.53	0.81
41	19	Other service activities	93	Labourers in mining, construction, manufacturing and transport	61	0.01	0.49	0.83
42	17	Human health and social work activities	81	Stationary plant and machine operators	72	0.03	0.64	0.85
43	7	Wholesale and retail trade; repair of motor vehicles and motorcycles	94	Food preparation assistants	47	0.01	0.46	0.85
44	14	Administrative and support service activities	53	Personal care workers	61	0.01	0.46	0.74
45	16	Education	75	Food processing, wood working, garment and other craft and related trades workers	90	0.00	0.45	0.68
46	5	Water supply, sewerage, waste management and remediation activities	61	Market-oriented skilled agricultural workers	77	0.00	0.64	0.90
47	3	Manufacturing	94	Food preparation assistants	52	0.01	0.54	0.85
48	15	Public administration and defence, compulsory social security	91	Cleaners and helpers	48	0.18	0.42	0.82
49	16	Education	53	Personal care workers	76	0.27	0.53	0.82
50	9	Accommodation and food service activities	92	Agricultural, forestry and fishery labourers	83	0.01	0.33	0.70
51	15	Public administration and defence, compulsory social security	94	Food preparation assistants	94	0.02	0.43	0.89
52	16	Education	96	Refuse workers and other elementary workers	50	0.07	0.53	0.83
53	9	Accommodation and food service activities	61	Market-oriented skilled agricultural workers	41	0.00	0.39	0.83
54	8	Transportation and storage	91	Cleaners and helpers	81	0.02	0.50	0.80
55	15	Public administration and defence, compulsory social security	82	Assemblers	49	0.00	0.76	0.80
56	14	Administrative and support service activities	93	Labourers in mining, construction, manufacturing and transport	41	0.08	0.38	0.75

Preliminary framework for a Eurofound Collective Agreements Database
on collectively bargained minimum pay rates

Rank by low-income pay	NACE	NACE label	ISCO	ISCO label	CA coverage (%)	Group as a % share of total employment	% of group with earnings in lowest 3 deciles	% of group with earnings in lowest 5 deciles
57	7	Wholesale and retail trade; repair of motor vehicles and motorcycles	93	Labourers in mining, construction, manufacturing and transport	53	0.70	0.39	0.75
58	12	Real estate activities	96	Refuse workers and other elementary workers	65	0.02	0.41	0.78
59	8	Transportation and storage	61	Market-oriented skilled agricultural workers	74	0.00	0.75	0.88
60	9	Accommodation and food service activities	51	Personal service workers	56	2.09	0.39	0.74
61	17	Human health and social work activities	72	Metal, machinery and related trades workers	68	0.02	0.62	0.76
62	7	Wholesale and retail trade; repair of motor vehicles and motorcycles	96	Refuse workers and other elementary workers	45	0.07	0.37	0.72
63	17	Human health and social work activities	53	Personal care workers	79	1.90	0.36	0.71
64	11	Financial and insurance activities	91	Cleaners and helpers	66	0.00	0.51	0.76
65	6	Construction	91	Cleaners and helpers	59	0.01	0.46	0.77
66	7	Wholesale and retail trade; repair of motor vehicles and motorcycles	61	Market-oriented skilled agricultural workers	56	0.01	0.49	0.75
67	18	Arts, entertainment and recreation	96	Refuse workers and other elementary workers	51	0.03	0.40	0.78
68	18	Arts, entertainment and recreation	32	Health associate professionals	87	0.00	0.52	0.74
69	9	Accommodation and food service activities	96	Refuse workers and other elementary workers	50	0.02	0.45	0.72
70	14	Administrative and support service activities	81	Stationary plant and machine operators	54	0.04	0.41	0.74
71	8	Transportation and storage	81	Stationary plant and machine operators	80	0.04	0.44	0.75
72	5	Water supply, sewerage, waste management and remediation activities	91	Cleaners and helpers	73	0.01	0.52	0.76
73	6	Construction	96	Refuse workers and other elementary workers	41	0.01	0.39	0.67
74	14	Administrative and support service activities	52	Sales workers	52	0.08	0.50	0.78
75	14	Administrative and support service activities	96	Refuse workers and other elementary workers	64	0.10	0.46	0.68
76	4	Electricity, gas, steam and air conditioning supply	91	Cleaners and helpers	81	0.00	0.73	0.91

Preliminary framework for a Eurofound Collective Agreements Database
on collectively bargained minimum pay rates

Rank by low-income pay	NACE	NACE label	ISCO	ISCO label	CA coverage (%)	Group as a % share of total employment	% of group with earnings in lowest 3 deciles	% of group with earnings in lowest 5 deciles
77	10	Information and communication	96	Refuse workers and other elementary workers	31	0.00	0.41	0.74
78	16	Education	52	Sales workers	68	0.00	0.51	0.77
79	9	Accommodation and food service activities	93	Labourers in mining, construction, manufacturing and transport	41	0.00	0.48	0.78
80	18	Arts, entertainment and recreation	94	Food preparation assistants	67	0.00	0.36	0.80
81	19	Other service activities	61	Market-oriented skilled agricultural workers	76	0.01	0.35	0.71
82	19	Other service activities	75	Food processing, wood working, garment and other craft and related trades workers	86	0.02	0.41	0.74
83	9	Accommodation and food service activities	83	Drivers and mobile plant operators	38	0.02	0.46	0.76
84	7	Wholesale and retail trade; repair of motor vehicles and motorcycles	52	Sales workers	56	5.08	0.40	0.71
85	15	Public administration and defence, compulsory social security	96	Refuse workers and other elementary workers	32	0.19	0.42	0.73
86	10	Information and communication	91	Cleaners and helpers	63	0.00	0.62	0.92
87	6	Construction	34	Legal, social, cultural and related associate professionals	54	0.01	0.35	0.68
88	17	Human health and social work activities	51	Personal service workers	63	0.25	0.34	0.70
89	18	Arts, entertainment and recreation	51	Personal service workers	65	0.09	0.35	0.67
90	14	Administrative and support service activities	61	Market-oriented skilled agricultural workers	65	0.24	0.32	0.67
91	16	Education	54	Protective services workers	56	0.01	0.50	0.80
92	19	Other service activities	52	Sales workers	75	0.01	0.39	0.71
93	3	Manufacturing	62	Market-oriented skilled forestry, fishery and hunting workers	67	0.00	0.53	0.70
94	5	Water supply, sewerage, waste management and remediation activities	93	Labourers in mining, construction, manufacturing and transport	71	0.04	0.35	0.71

Preliminary framework for a Eurofound Collective Agreements Database
on collectively bargained minimum pay rates

Rank by low-income pay	NACE	NACE label	ISCO	ISCO label	CA coverage (%)	Group as a % share of total employment	% of group with earnings in lowest 3 deciles	% of group with earnings in lowest 5 deciles
95	7	Wholesale and retail trade; repair of motor vehicles and motorcycles	53	Personal care workers	21	0.01	0.14	0.72
96	13	Professional, scientific and technical activities	42	Customer services clerks	48	0.06	0.33	0.59
97	17	Human health and social work activities	71	Building and related trades workers, excluding electricians	89	0.04	0.33	0.65
98	9	Accommodation and food service activities	75	Food processing, wood working, garment and other craft and related trades workers	83	0.04	0.37	0.69
99	16	Education	72	Metal, machinery and related trades workers	78	0.01	0.51	0.76
100	14	Administrative and support service activities	51	Personal service workers	50	0.13	0.34	0.67
101	2	Mining and quarrying	91	Cleaners and helpers	91	0.00	0.35	0.70
102	14	Administrative and support service activities	54	Protective services workers	36	0.67	0.36	0.66
103	14	Administrative and support service activities	42	Customer services clerks	58	0.27	0.34	0.66
104	17	Human health and social work activities	52	Sales workers	59	0.01	0.40	0.66
105	14	Administrative and support service activities	75	Food processing, wood working, garment and other craft and related trades workers	84	0.02	0.32	0.68
106	14	Administrative and support service activities	71	Building and related trades workers, excluding electricians	74	0.10	0.27	0.66
107	9	Accommodation and food service activities	42	Customer services clerks	68	0.27	0.30	0.67
108	3	Manufacturing	61	Market-oriented skilled agricultural workers	58	0.02	0.22	0.70
109	15	Public administration and defence, compulsory social security	53	Personal care workers	95	0.10	0.35	0.69
110	18	Arts, entertainment and recreation	52	Sales workers	64	0.03	0.34	0.72
111	13	Professional, scientific and technical activities	96	Refuse workers and other elementary workers	56	0.00	0.36	0.64

Preliminary framework for a Eurofound Collective Agreements Database
on collectively bargained minimum pay rates

Rank by low-income pay	NACE	NACE label	ISCO	ISCO label	CA coverage (%)	Group as a % share of total employment	% of group with earnings in lowest 3 deciles	% of group with earnings in lowest 5 deciles
112	18	Arts, entertainment and recreation	71	Building and related trades workers, excluding electricians	75	0.02	0.29	0.61
113	16	Education	71	Building and related trades workers, excluding electricians	56	0.02	0.33	0.65
114	7	Wholesale and retail trade; repair of motor vehicles and motorcycles	75	Food processing, wood working, garment and other craft and related trades workers	66	0.28	0.32	0.65
115	13	Professional, scientific and technical activities	82	Assemblers	71	0.00	0.11	0.80
116	5	Water supply, sewerage, waste management and remediation activities	54	Protective services workers	74	0.00	0.55	0.75
117	16	Education	73	Handicraft and printing workers	63	0.00	0.32	0.62
118	6	Construction	93	Labourers in mining, construction, manufacturing and transport	62	0.54	0.33	0.63
119	6	Construction	73	Handicraft and printing workers	65	0.00	0.29	0.69
120	16	Education	51	Personal service workers	53	0.22	0.34	0.66
121	3	Manufacturing	75	Food processing, wood working, garment and other craft and related trades workers	55	1.85	0.35	0.65
122	18	Arts, entertainment and recreation	54	Protective services workers	68	0.04	0.30	0.64
123	6	Construction	42	Customer services clerks	61	0.01	0.21	0.45
124	3	Manufacturing	52	Sales workers	56	0.39	0.39	0.67
125	12	Real estate activities	51	Personal service workers	79	0.07	0.26	0.71
126	16	Education	61	Market-oriented skilled agricultural workers	73	0.00	0.55	0.76
127	18	Arts, entertainment and recreation	92	Agricultural, forestry and fishery labourers	80	0.00	0.26	0.57
128	17	Human health and social work activities	42	Customer services clerks	55	0.12	0.23	0.58
129	8	Transportation and storage	93	Labourers in mining, construction, manufacturing and transport	74	0.34	0.27	0.61

Preliminary framework for a Eurofound Collective Agreements Database
on collectively bargained minimum pay rates

Rank by low-income pay	NACE	NACE label	ISCO	ISCO label	CA coverage (%)	Group as a % share of total employment	% of group with earnings in lowest 3 deciles	% of group with earnings in lowest 5 deciles
130	9	Accommodation and food service activities	71	Building and related trades workers, excluding electricians	58	0.02	0.19	0.50
131	18	Arts, entertainment and recreation	42	Customer services clerks	62	0.10	0.32	0.56
132	3	Manufacturing	96	Refuse workers and other elementary workers	46	0.08	0.31	0.65
133	7	Wholesale and retail trade; repair of motor vehicles and motorcycles	51	Personal service workers	49	0.10	0.35	0.67
134	9	Accommodation and food service activities	54	Protective services workers	43	0.02	0.31	0.60
135	3	Manufacturing	93	Labourers in mining, construction, manufacturing and transport	55	1.31	0.31	0.63
136	17	Human health and social work activities	14	Hospitality, retail and other services managers	58	0.00	0.01	0.31
137	7	Wholesale and retail trade; repair of motor vehicles and motorcycles	54	Protective services workers	46	0.02	0.43	0.72
138	10	Information and communication	93	Labourers in mining, construction, manufacturing and transport	40	0.00	0.26	0.66
139	5	Water supply, sewerage, waste management and remediation activities	96	Refuse workers and other elementary workers	73	0.25	0.31	0.61
140	17	Human health and social work activities	54	Protective services workers	75	0.03	0.25	0.68
141	7	Wholesale and retail trade; repair of motor vehicles and motorcycles	42	Customer services clerks	53	0.13	0.24	0.59
142	8	Transportation and storage	96	Refuse workers and other elementary workers	88	0.12	0.28	0.59
143	6	Construction	75	Food processing, wood working, garment and other craft and related trades workers	70	0.14	0.23	0.61
144	12	Real estate activities	41	General and keyboard clerks	73	0.04	0.26	0.63
145	9	Accommodation and food service activities	41	General and keyboard clerks	93	0.02	0.29	0.60
146	18	Arts, entertainment and recreation	41	General and keyboard clerks	87	0.05	0.27	0.61
147	19	Other service activities	42	Customer services clerks	78	0.02	0.24	0.59

Preliminary framework for a Eurofound Collective Agreements Database
on collectively bargained minimum pay rates

Rank by low-income pay	NACE	NACE label	ISCO	ISCO label	CA coverage (%)	Group as a % share of total employment	% of group with earnings in lowest 3 deciles	% of group with earnings in lowest 5 deciles
148	18	Arts, entertainment and recreation	61	Market-oriented skilled agricultural workers	63	0.02	0.35	0.64
149	17	Human health and social work activities	41	General and keyboard clerks	83	0.29	0.22	0.61
150	16	Education	42	Customer services clerks	79	0.04	0.19	0.43
151	15	Public administration and defence, compulsory social security	71	Building and related trades workers, excluding electricians	96	0.12	0.23	0.57
152	14	Administrative and support service activities	41	General and keyboard clerks	77	0.16	0.26	0.61
153	12	Real estate activities	83	Drivers and mobile plant operators	66	0.00	0.40	0.44
154	17	Human health and social work activities	83	Drivers and mobile plant operators	68	0.09	0.27	0.57
155	15	Public administration and defence, compulsory social security	81	Stationary plant and machine operators	45	0.01	0.57	0.71
156	19	Other service activities	71	Building and related trades workers, excluding electricians	81	0.01	0.25	0.58
157	19	Other service activities	54	Protective services workers	88	0.01	0.30	0.59
158	7	Wholesale and retail trade; repair of motor vehicles and motorcycles	32	Health associate professionals	32	0.24	0.23	0.57
159	7	Wholesale and retail trade; repair of motor vehicles and motorcycles	72	Metal, machinery and related trades workers	52	1.00	0.24	0.53
160	15	Public administration and defence, compulsory social security	61	Market-oriented skilled agricultural workers	76	0.09	0.27	0.57
161	10	Information and communication	42	Customer services clerks	57	0.06	0.30	0.58
162	4	Electricity, gas, steam and air conditioning supply	51	Personal service workers	55	0.00	0.26	0.43
163	8	Transportation and storage	44	Other clerical support workers	96	0.44	0.21	0.58
164	16	Education	41	General and keyboard clerks	76	0.24	0.21	0.58
165	16	Education	83	Drivers and mobile plant operators	68	0.01	0.29	0.56

Preliminary framework for a Eurofound Collective Agreements Database
on collectively bargained minimum pay rates

Rank by low-income pay	NACE	NACE label	ISCO	ISCO label	CA coverage (%)	Group as a % share of total employment	% of group with earnings in lowest 3 deciles	% of group with earnings in lowest 5 deciles
166	13	Professional, scientific and technical activities	71	Building and related trades workers, excluding electricians	77	0.01	0.26	0.48
167	19	Other service activities	73	Handicraft and printing workers	66	0.00	0.15	0.62
168	7	Wholesale and retail trade; repair of motor vehicles and motorcycles	41	General and keyboard clerks	70	0.41	0.21	0.55
169	17	Human health and social work activities	73	Handicraft and printing workers	80	0.00	0.50	0.56
170	7	Wholesale and retail trade; repair of motor vehicles and motorcycles	34	Legal, social, cultural and related associate professionals	40	0.03	0.24	0.56
171	14	Administrative and support service activities	73	Handicraft and printing workers	56	0.00	0.19	0.68
172	7	Wholesale and retail trade; repair of motor vehicles and motorcycles	83	Drivers and mobile plant operators	42	0.54	0.23	0.54
173	6	Construction	41	General and keyboard clerks	87	0.16	0.24	0.54
174	6	Construction	71	Building and related trades workers, excluding electricians	69	2.90	0.21	0.52
175	7	Wholesale and retail trade; repair of motor vehicles and motorcycles	71	Building and related trades workers, excluding electricians	36	0.13	0.21	0.49
176	18	Arts, entertainment and recreation	34	Legal, social, cultural and related associate professionals	77	0.19	0.31	0.54
177	12	Real estate activities	52	Sales workers	47	0.00	0.10	0.39
178	13	Professional, scientific and technical activities	41	General and keyboard clerks	83	0.26	0.22	0.53
179	18	Arts, entertainment and recreation	44	Other clerical support workers	75	0.03	0.19	0.50
180	17	Human health and social work activities	43	Numerical and material recording clerks	67	0.10	0.24	0.56
181	7	Wholesale and retail trade; repair of motor vehicles and motorcycles	82	Assemblers	28	0.03	0.17	0.53

Preliminary framework for a Eurofound Collective Agreements Database
on collectively bargained minimum pay rates

Rank by low-income pay	NACE	NACE label	ISCO	ISCO label	CA coverage (%)	Group as a % share of total employment	% of group with earnings in lowest 3 deciles	% of group with earnings in lowest 5 deciles
182	13	Professional, scientific and technical activities	93	Labourers in mining, construction, manufacturing and transport	59	0.02	0.22	0.53
183	3	Manufacturing	82	Assemblers	61	1.15	0.24	0.54
184	8	Transportation and storage	75	Food processing, wood working, garment and other craft and related trades workers	88	0.01	0.25	0.58
185	9	Accommodation and food service activities	31	Science and engineering associate professionals	71	0.02	0.16	0.53
186	9	Accommodation and food service activities	43	Numerical and material recording clerks	78	0.03	0.20	0.50
187	17	Human health and social work activities	34	Legal, social, cultural and related associate professionals	49	0.48	0.23	0.51
188	12	Real estate activities	61	Market-oriented skilled agricultural workers	54	0.00	0.26	0.66
189	15	Public administration and defence, compulsory social security	73	Handicraft and printing workers	40	0.01	0.24	0.55
190	12	Real estate activities	71	Building and related trades workers, excluding electricians	67	0.01	0.15	0.47
191	13	Professional, scientific and technical activities	73	Handicraft and printing workers	51	0.02	0.22	0.50
192	8	Transportation and storage	42	Customer services clerks	93	0.16	0.16	0.47
193	16	Education	34	Legal, social, cultural and related associate professionals	72	0.41	0.24	0.50
194	13	Professional, scientific and technical activities	51	Personal service workers	48	0.01	0.24	0.53
195	19	Other service activities	83	Drivers and mobile plant operators	73	0.01	0.17	0.53
196	14	Administrative and support service activities	72	Metal, machinery and related trades workers	44	0.04	0.18	0.49
197	7	Wholesale and retail trade; repair of motor vehicles and motorcycles	81	Stationary plant and machine operators	53	0.20	0.28	0.58

Preliminary framework for a Eurofound Collective Agreements Database
on collectively bargained minimum pay rates

Rank by low-income pay	NACE	NACE label	ISCO	ISCO label	CA coverage (%)	Group as a % share of total employment	% of group with earnings in lowest 3 deciles	% of group with earnings in lowest 5 deciles
198	13	Professional, scientific and technical activities	32	Health associate professionals	64	0.07	0.25	0.48
199	16	Education	43	Numerical and material recording clerks	72	0.04	0.23	0.52
200	6	Construction	54	Protective services workers	63	0.00	0.44	0.63

Appendix II: Sources for collective agreements universe count

Country	URL of source
Austria	http://www.kvssystem.at/servlet/ContentServer?pagename=KVS/Page/KVS_Index&n=KVS_2.3
Belgium	https://emploi.belgique.be/fr/themes/commissions-paritaires-et-conventions-collectives-de-travail-cct/conventions-collectives-3
Bulgaria	https://www.nipa.bg/%D0%B1%D0%B0%D0%B7%D0%B0-%D0%B4%D0%B0%D0%BD%D0%BD%D0%B8/signed/2020-level/
Croatia	https://www.worker-participation.eu/National-Industrial-Relations/Countries/Croatia/Collective-Bargaining#_ftn2
Cyprus	Ηλεκτρονικό Αρχείο Συλλογικών Συμβάσεων Τμήμα Εργασιακών Σχέσεων (mlsi.gov.cy)
Czechia	https://www.mpsv.cz/kolektivni-smlouvy-vyssiho-stupne-zavazne-pro-dalsi-zamestnavatele
Denmark	https://www.danskindustri.dk/vi-radgiver-dig-ny/personale/overenskomster-og-det-fagretlige-system/hvad-er-en-overenskomst/
Estonia	https://klak.sm.ee/leping/filter/23/load.html
Finland	https://www.finlex.fi/fi/viranomaiset/tyoehto/
France	https://www.legifrance.gouv.fr/liste/idcc?init=true%20as%20in%2015/04/2021
Germany	https://www.bmas.de/DE/Arbeit/Arbeitsrecht/Tarifvertraege/allgemeinverbindliche-tarifvertraege.html
Germany	https://www.bundesanzeiger.de/pub/de/suchen2?5
Greece	https://ypergasias.gov.gr/category/sillogikes_rythmiseis_ergasias/page/23/
Hungary	https://www.worker-participation.eu/National-Industrial-Relations/Countries/Hungary/Collective-Bargaining#_ftn2
Hungary	http://www.mkir.gov.hu/ksznylv.htm
Ireland	https://www.worker-participation.eu/National-Industrial-Relations/Countries/Ireland/Collective-Bargaining
Italy	https://www.cnel.it/Archivio-Contratti
Latvia	https://www.worker-participation.eu/National-Industrial-Relations/Countries/Latvia/Collective-Bargaining
Lithuania	https://socmin.lrv.lt/lt/paslaugos/administracines-paslaugos/kolektyvinu-sutarciu-registras-ir-kolektyvinu-sutarciu-registravimo-tvarka
Luxembourg	Conventions collectives - Conditions de travail - Inspection du travail et des mines - Luxembourg (public.lu)
Malta	https://dier.gov.mt/en/Pages/home.aspx
Netherlands	https://www.uitvoeringarbeidsvoorwaardenwetgeving.nl/mozard/document/docnr/1034411
Poland	http://www.dialog.gov.pl/dialog-krajowy/uklady-zbiorowe-pracy/stan-rejestru-prowadzonego-przez-ministra-rodziny-pracy-i-polityki-spoecznej/
Portugal	Instrumentos de Regulamentação Coletiva de Trabalho Publicados – DGERT
Romania	http://dialogsocial.gov.ro/sector-de-activitate/
Romania	http://dialogsocial.gov.ro/grup-de-unitati/
Slovakia	https://www.employment.gov.sk/sk/praca-zamestnanost/vztah-zamestnanca-zamestnavateľa/kolektivne-pracovnopravne-vztahy/kolektivne-zmluvy/zoznam-kolektivnych-zmluv-vyssiho-stupna/

Slovenia	http://www.arhiv-spletisc.gov.si/
Spain	https://www.mites.gob.es/estadisticas/cct/welcome.htm
Sweden	https://www.mi.se/app/uploads/AR20_bok_ori_skm.pdf

Source: *Authors' own elaboration*

The European Foundation for the Improvement of Living and Working Conditions (Eurofound) is a tripartite European Union agency established in 1975. Its role is to provide knowledge in the area of social, employment and work-related policies according to Regulation (EU) 2019/127.