

Subnational Level Data to Measure Gender (in)Equality in the EU: Opportunities and Limitations of Official Datasets



Enrico di Bella and Fabrizio Culotta

1 Introduction

Data-driven decision-making is a process in which metrics and data guide policymakers' strategies and actions. Empirical evidence plays a crucial role in policymaking since it offers substantial ground for planning, evaluating, and challenging policies. The correct interpretation of a complex phenomenon requires information that data, and their analysis, can offer. Wider data availability allows a richer interpretation of the social phenomena under study and a better decision-making process. Larger data availability can translate into at least two forms: a larger number of variables available or a deeper territorial level detail. In this chapter, we focus mainly on this second point, particularly on the role of regional data for policymaking and the availability of data at subnational levels in the main European datasets.

Country-level analyses can mask territorial disparities, particularly when synthetic indicators are used. The national average score can hide the territorial heterogeneity of a phenomenon, compensating for the low scores of some regions and high scores for others. Analyses at the regional level are therefore necessary if the dispersion of a phenomenon within a country hints at disparities that should be addressed through policymaking. Territorial differences within countries are due to numerous causes (e.g. Kaasa et al., 2014; Marelli, 2007), whose origins are often controversial and lost in history. In Italy, for instance, it has been repeatedly verified that most social phenomena have a marked territorial difference between the north and south of the country as a consequence of its relatively recent unification in 1861 and unequal socioeconomic growth in the subsequent one and half century. In Germany, the separation between the Federal Republic of Germany (West Germany)

E. di Bella (✉) · F. Culotta

Department of Political and International Sciences, University of Genoa, Genoa, Italy

e-mail: enrico.dibella@unige.it; fabrizio.culotta@edu.unige.it

© The Author(s) 2023

E. di Bella et al. (eds.), *Measuring Gender Equality*, Social Indicators Research Series 87, https://doi.org/10.1007/978-3-031-41486-2_5

119

and the German Democratic Republic (East Germany) from 1949 to 1990 resulted in significant social differences, traces of which are still visible in many socioeconomic analyses at the subnational level. Even countries that emerged as modern states at the end of the Middle Ages and have a long tradition of national identity, such as France or Spain, record some differences at the territorial level. Generally, no country has some form of internal heterogeneity. Regardless of the degree of territorial analysis, it is possible to identify forms of inequality or segregation based on the territory's economic, social, or environmental characteristics.

These different origins of territorial disparities make regional studies relevant because they can integrate the conclusions drawn at the national level by revealing divergences in socioeconomic outcomes and unmasking the role of local factors affecting the performance of that territory. They can also help identify clusters, that is, the presence of multiple but homogeneous groups of regions. The findings from a regional analysis can inform policymakers of the possible risks and opportunities related to implementing economic policies. For example, the sectorial distribution of firms can reveal the possibility of asymmetric effects from industrial policies designed at the national level. Indeed, the labour market is one of the domains that most interact with others; work choices depend on education; and it influences income and work–life balance with further effects on the other domains of gender (in)equality.

Similarly, information on the geographical distribution of workers' skills can provide information on the population targeted at the local level. A detailed set of information at the local level can also reveal the role of cultural institutions in shaping regional disparities (Fortin, 2005; Alesina et al., 2013; Kushi & McManus, 2018).

Having information at the regional level is also important for verifying the conditions implicitly assumed when nationwide policies are implemented. As is often the case, national policies disregard geographical heterogeneity by embracing a space-neutral approach (Altavilla & Caroleo, 2013; Iammarino et al., 2019). This policy paradigm implicitly assumes that the economic structure of a country is uniformly reproduced at the subnational level. The validity of this one-size-fits-all approach cannot be questioned, at least in the long run, when territorial differences are expected to level out. Forces driving economic growth, such as innovation, usually originate in some developed core areas and then propagate towards developing peripheral regions (Blanchard & Katz, 1992). Of course, the (speed of) propagation strictly depends on the mobility of productive factors. Thus, the assumption of factor mobility is central to the correct implementation of space-neutral policies (Iammarino et al., 2019). Accordingly, information on the mobility conditions of workers, capital, and entrepreneurship can reveal mobility frictions that influence the spread of new economic opportunities. Violations of mobility assumptions would invalidate, at least partially, their original scope. For example, when dealing with gendered issues in the labour market (see Chapter “Gender-Responsive Regional Fiscal Policies: The Labour Market”), the empirical literature highlights the role of different commuting attitudes in explaining gender gaps (Perales & Vidal, 2015; Gimenez-Nadal & Molina, 2016; Nisic, 2017; Petrongolo

& Ronchi, 2020; Fuchs et al., 2021; Le Barbanchon et al., 2021). Compared to men, women value the geographical closeness of job opportunities because a higher share of family responsibilities ties women to the area where they live.¹ Under this limiting condition, information about vacancies and workers at the regional level is crucial for the effectiveness of gender-responsive policies. If the effects of a policy correlate with economic performance at the regional level, it is likely that place-neutral policies will exacerbate regional disparities.

Currently, regional disparities remain larger than before the 2008 crisis (European Commission, 2022). Difficulties in triggering developmental paths in less-developed regions are responsible for the persistence of territorial differences. Space-based policies cope better with persistent regional differentials because the intervention is based on regional demographic and economic characteristics (Iammarino et al., 2019). The compensative effect between developed and developing regions, typical of space-neutral policies, no longer operates in the context of place-based policies. Regional policies are called actions to reduce structural disparities and restore regional convergence. Regional cohesion policies, as promoted by the EU, will play a crucial role in this respect, as they stress the role of local communities in reaching national and community-wide goals. However, while space-neutral policies generally do not require information at the subnational level (if not, *ex post*, in the monitoring phase), space-based policies (in general, and European cohesion policies in particular) require a rich set of information at the subnational level to be fully implemented and effective.

Despite the rich availability of gender (in)equality indicators at the national level (see Chapter “The Main Indicators of Gender (in)Equality”), studies on gender equality indexes at the regional level are scarce due to limited data availability. Early examples were provided for Norwegian municipalities (Kjeldstad & Kristiansen, 2001). Not surprisingly, Scandinavian countries, such as Norway and Sweden, have a long tradition of regional policies (Rönblom, 2005). Other studies have focused on Spanish regions (Peinado & Céspedes, 2004; Bericat Alastuey & Sánchez Bermejo, 2008; Fernández-Sáez et al. 2016; Gil-Lafuente et al., 2019) and provinces (Martin and Garvi, 2009). Regarding Italian territory, see Amici and Stefani (2013) and Costantini and Monni (2006). More recently, Hippe and Perrin (2017) studied gender inequality in human capital across the EU regions (NUTS 1²)

¹The so-called household responsibility hypothesis (Gimenez-Nadal & Molina, 2016), for which a disproportionate burden of household responsibility on women requires shorter commute times. This makes more difficult for them to work any distance away from home compared to men.

²The European Union has established a common classification of territorial units for statistics, known as ‘NUTS’, in order to facilitate the collection, development, and publication of harmonised regional statistics in the EU. This hierarchical system is also used for socioeconomic analyses of the regions and the framing of interventions in the context of EU cohesion policy. The NUTS classification is hierarchical in that it subdivides each Member State into three levels: NUTS 1, NUTS 2, and NUTS 3. The second and third levels are subdivisions of the first and second levels. A Member State may decide to add further levels to the hierarchy by subdividing NUTS level 3. Some of the existing administrative units used for the requirements of the hierarchical NUTS classification are listed in Annex II to Regulation (EC) No 1059/2003.

over 1900–1960. Castellano and Rocca (2019) constructed an index of the gender gap in the labour market at the NUTS 1 level for 31 European countries. The results from 2013 highlight a strong regional variability across various labour market outcomes. Dijkstra et al. (2019) measured gender gaps in terms of female disadvantages and achievements for over 270 NUTS 2 in the European Union. The results confirm that gender gaps are more prominent in regions with higher unemployment rates, whereas female achievement is greater in regions with a higher GDP per capita. The EU Regional Gender Equality Monitor at the Joint Research Center (Norlen et al., 2019) provides a regional-level analysis of the UE. Finally, di Bella et al. (2021) adapted the multifaceted GEI indicator (EIGE, 2017) to Italy's regional context. See Cascella et al. (2022) for an extension of the GEI at the regional level. The results from the R-GEI confirmed that gender equality varies substantially within European member countries. Finally, Perrin (2021) represents the first attempt to extend the regional analysis over a long time horizon by building an indicator at the county level for France since the 1850s. Aimed at investigating the long relationship between the gender gap and economic growth, the work concludes that France's north-south divide already existed at that time. Gender equality is positively associated with economic performance. Northern counties, where the gender gap is narrower, experience higher economic growth rates. Persistent regional differences in economic outcomes can thus lead to persistent sex differences between women and men.

2 Datasets for Gender (in)Equality Measurement: The Eurostat Collection

Eurostat provides several microdata sources for analysing gender (in)equality in the EU. Microdata are survey records that contain primary information on individuals, households, and business entities. The increasing availability of microdata satisfies the need for more detailed information, bringing considerable advantages compared to secondary data published in tabular form on the Eurostat portal. Being the records of individual questionnaires, suitably anonymised in such a way as to preserve the privacy of the respondents, microdata make it possible to conduct complex statistical analyses that would otherwise be impossible. Hypotheses and theories can be tested on different subsets of the population, whether distinguished according to their territorial area of reference or according to some socioeconomic criteria. Using a 'rich-in-details dataset' allows researchers to analyse data concerning the specific socioeconomic aspects of society. This occurs when a given survey questionnaire is augmented with ad hoc modules to retrieve information on a specific phenomenon. Another case is when a given survey increases territorial representativeness by considering subnational levels. A further extension of a standard survey is the inclusion of a longitudinal dimension, which allows one to track the dynamics of individual entities over time. The various innovations that statistical institutes make

in the data provision process move in a fourth direction: a cross-country dimension. The harmonisation of national surveys and the conducting of the same surveys across member countries provides a broader perspective of researchers and other stakeholders. The resulting comparative analysis sheds light on the similarities and differences that converge and diverge across countries.

Eurostat has a specific mandate and long tradition of coordinating survey activities across EU member states, candidates, and EFTA countries. In doing so, it ensures standardisation in the definitions and measurements of the countries involved in the sample surveys so that the data are comparable. Access to these data is possible through a special request for scientific purposes.³ Different Eurostat surveys focus on different aspects of individuals' economic and social activities, such as income structure, health conditions, labour market outcomes, vocational training, time use in daily life, and adoption of ICT devices.

In this chapter, to be consistent with what will be discussed in the next section, we discuss only the Eurostat surveys used in EIGE's GEI framework (EIGE, 2022). The selection comprises the European Consumer Household Panel (EHCP) and the European Union Survey on Income and Living Conditions (EU-SILC) for household income and living conditions, the European Union Labour Force Survey (EU-LFS) for labour market statistics, the Structure of Earning Survey (SES) for the structure of earnings, the Harmonised European Time Use Surveys (HETUS) for the use of daily time, and the European Health Interview Survey (EHIS). Eurofound data, although important for gender analysis in Europe and used by the EIGE at the national level, are not listed here because these surveys are not sufficiently representative at any subnational level because of their limited sample size.

2.1 *ECHP and EU-SILC*

In 1991, Eurostat set up a task force to obtain information on households and individual income from national registers and household income surveys, and to check whether the available output could be harmonised ex post. The failure of this approach led to the launch of a specific EU survey—the ECHP. It was conducted for eight yearly waves, between 1994 and 2001. As of 2003, the European Statistics on Income and Living Conditions (EU-SILC) inherited the characteristics of the former ECHP. Currently, the EU-SILC represents the primary data source covering household income and living conditions in the EU, from labour market status to poverty and social activities. Similar to the ECHP, the EU-SILC is a panel survey consisting of interviews held with households and individuals year after year. The interviews covered various topics, ranging from sociodemographic information to household

³These are based on the original microdata which are confidential information accessible by means of direct identifiers. Once anonymization procedures are applied, the original survey data become non-confidential and can be released publicly.

income and finance, working life, housing, social relations, and health. The survey was conducted at the European level and involved 14 Member States.⁴ The survey was conducted both across and within the households. The household module contains demographic information, income and financial conditions, children, accommodations, and durables. The personal file contains a section for the job search activity if unemployed, for the training activity if employed, and for social relations and life satisfaction. Furthermore, owing to the longitudinal dimension of the sampling design, it is possible to extract information on the social dynamics of the surveyed units repeatedly over time.

Regarding the gender gap, the EU-SILC represents an essential source of information concerning the characteristics of women within the household environment, especially for their focus on labour and housing conditions, deprivation, and social activities. These additional aspects can provide new insights into gender-related differences. The information provided by the EU-SILC is reported at the regional level (NUTS 2). This feature offers a detailed level of analysis. Thus, the EU-SILC is a valuable source of microdata for regional analysis, and it represents an important source for the construction of the GEI across countries since it is used in the entire construction of the money domain, both for economic resources and financial resources subdomains, and the health Access subdomain.

2.2 *EU-LFS*

The European Labor Force Survey (EU-LFS) is a collection of national LFS gathered by Eurostat, and it is the data source on which the official labour market statistics for employment, unemployment, and inactivity of individuals aged 15+ years across European countries since 1983.

As a collection of national LFS, the EU-LFS interviewed household members to capture information about their current and past employment experiences. They gather information on demographics, labour market status, employment characteristics of the main job as well as secondary employment, atypical work, previous work experience and employment status, working hours, job search activities and methods, education and training, and income decile. Regarding the gender gap, the EU-LFS provides information not only about the different incidences of activity and inactivity between men and women, but also about employment characteristics for the employed population and job-search activities for unemployed individuals. Thus, the EU-LFS represents the primary data source for analysing the gender gap in the labour market across European countries. The EU-LFS is adopted in

⁴The first wave of 1994 comprised the following twelve countries: Belgium, Denmark, France, Germany, Greece, Ireland, Italy, the Netherlands, Portugal, Spain, and the UK. Austria and Finland joined in 1995 and 1996, respectively. Sweden joined in 1997 despite the ECHP having been derived from the Swedish Living Conditions Survey itself.

constructing the GEI, mainly for the domains of work, time, and knowledge. The EU-LFS captures the one-year-before-employment situation. This information is usually rearranged in terms of transitional probabilities for the ins and outs of unemployment, employment, and inactivity. Thus, gender-related differences are expected to emerge in such transitional contexts. Moreover, it provides information up to the provincial level (NUTS-3) and is therefore suitable for regional analysis.

2.3 *EU-SES*

The European Structure of Earning Survey (EU-SES) represents a European survey of enterprises and employee characteristics, with the corresponding earning structure. Enterprises were sampled only if they operated in the private sector with at least ten employees. It provides detailed information on wage levels and registers the benefits and taxes connected to the employment contract. The set of variables considered by the EU-SES distinguishes among firm-specific (e.g. employee size, economic sector), worker-specific (e.g. education, professional type, seniority), and other contract-specific variables (including holidays, social contributions, taxes, and allowances). The informative content of the EU-SES helps trace the dynamics of remuneration components, the effects of employment policies, and labour costs across (countries' fiscal regimes of) European countries.

The survey was conducted at the NUTS 1 level, that is, across territorial macro areas, which implies severe limitations for using these data in a regional analysis. Two additional GEI reasons do not make EU-SES suitable for the regional transposition of EIGE's GEI: time frequency and the sampling scheme. Relating to the first point, unlike data from EU-SILC and EU-LFS that are available yearly, the EU-SES has a four-year frequency. Relating to the sampling scheme, we note that since enterprises are sampled only if operating in the private sector with at least ten employees, this hinders the possibility of tracking the dynamics of the gender gaps in those regions characterised by a relatively high presence of (very) small firms. The larger the portion of women employed in these firms, the larger their underrepresentation in the EU-SES dataset. If a firm's size is unevenly distributed across regions, the estimated regional dispersion in gender gaps is biased. This consideration is more substantial if one considers that local employment opportunities are women usually showing lower (geographical) commuting profiles than men (Gimenez-Nadal & Molina, 2016; Petrongolo & Ronchi, 2020; Le Barbanchon et al., 2021; Fuchs et al., 2021).

2.4 *HETUS*

The Harmonised European Time Use Survey (HETUS) is a collection of national time use surveys. HETUS represents a unique dataset provided by Eurostat regarding the time use of European citizens, made recently in 2017, available for scientific

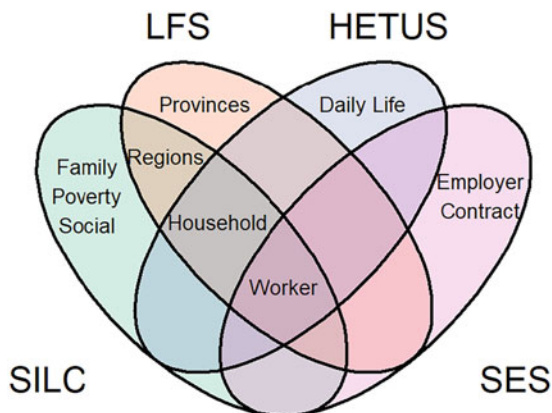
purposes. The main objective of this survey was to quantify the time spent on various activities at daily frequency. The overall HETUS questionnaire contains, in fact, around two thousand variables. Starting in 2000, covering the years 1998–2006, the survey was conducted once every 10 years. The second wave of 2010 was conducted from 2008 to 2015 across 15 EU and three non-EU countries.⁵ The sample of participants comprises 0.047% of the European resident population, corresponding to 120 thousand households, of which more than one-third are from Poland (27 thousand) and Italy (17 thousand). Among the surveyed countries, Germany and the Netherlands showed the lowest values in terms of sample size. The survey instrument is divided into three main parts: the household questionnaire, the individual questionnaire, and the time diary, registering activities in ten-minute time slots. Individual activities are classified according to the activity coding list (ADL), that is, the official 2-digit⁶ classification adopted by Eurostat to characterise the daily activities of individuals.

HETUS is an important source of information for understanding social phenomena. Given its focus on individuals' daily activities, the HETUS dataset allows for a deeper analysis of the gender gap in terms of time use. Thus, HETUS content can be exploited to identify further dimensions in which men and women exhibit different attitudes. Not surprisingly, the HETUS dataset was adopted to show different commuting attitudes between women and men. See Anxo et al. (2011) and Gimenez-Nadal and Alberto (2022) for the multicountry studies. See also Gimenez-Nadal and Molina (2014, 2016) for a regional study focusing on Spain and the Netherlands at the subnational level. They concluded that women commute less than men do because of their larger share of family responsibilities, limiting the possibility of finding employment over a longer search radius. In addition, information extrapolated from HETUS can also help to characterise in more detail the inactivity among the working-age population, the distribution of tasks within households (e.g. cooking and family care), and the impact of cultural differences within the European population. Despite the detailed set of information regarding individual daily activities, wave 2010 does not present any regional breakdowns, with countries being the only possible territorial partition (at NUTS 0 level) of the population. Therefore, it is not of interest if the analysis of the gender gap is conducted only at the national level.

⁵The 15 EU countries, representing the entire first wave (2000) of HETUS, are Austria, Belgium, Germany, Estonia, Greece, Finland, France, Hungary, Italy, Luxembourg, the Netherlands, Poland, Romania, Spain, and the UK. The three extra-EU countries are Norway, Serbia, and Turkey.

⁶The specific domains (one-digit) of ADL are personal care, employment, study, household and family care, voluntary work and meeting, social life and entertainment, sports and outdoor activities, hobbies and computing, mass media, travel, and, residually, unspecified time. The new version of 2018, used for the third wave of 2020, further expands the two-digit ADL-2008 to a three-digit structure to classify the daily activities of individuals.

Fig. 1 Eurostat surveys (EU-SILC, EU-LFS, SES, HETUS) intersecting elements. Source: authors' elaboration



2.5 *EHIS*

The European Health Interview Survey (EHIS) consists of four modules on health status, healthcare use, health determinants, and socioeconomic background variables for EU citizens coded into three categories (natives, EU citizens, and others) and by five-year bands (starting with 15–17, 18–19). The EHIS targets a population aged at least 15 years and living in private households. The regional detail is limited in this survey, being only available in Wave 3 (2019) for Belgium, Bulgaria, Czechia, Denmark, Germany, Estonia, Ireland, Spain, Croatia, Hungary, Italy, Lithuania, Austria, Poland, Portugal, Slovenia, Slovakia, Finland, Sweden, and Serbia; removed for the other countries in Wave 3; and not available in previous waves. Health data belong to special (sensitive) categories. Access to microdata on health requires justification of substantial public interest.

2.6 *On the Integration Among Eurostat Data Sources*

A set of shared variables across these surveys highlights the opportunity for the joint use of different data sources. However, while EU-SILC and EU-LFS can be exploited for regional analysis and SES allows analysis across macro areas, HETUS can only be used for a cross-country comparison. Figure 1 shows the different elements of the surveys (dimensions, territorial levels, units of measurement). Each survey does not exclusively cover a single element, but some are common to multiple surveys, such as education and worker conditions. Overall, what emerges from Fig. 1 shows that efforts to include the regional dimension in SES and HETUS are likely to enrich future regional analysis. Additional information about employment relations and daily life can enrich the variables considered when constructing synthetic indicators of gender equality.

3 Limitations of Official Datasets for Gender (in)Equality Measurement at the Regional Level in the EU

The first obvious prerequisite for a regional analysis of gender (in)equality is the availability of data collected at the regional level. Although it is seemingly obvious, this prerequisite has two implications. First, if a survey is not structured to provide regionally representative estimates, its use is limited and only possible after imputing regional values according to specific statistical criteria. For instance, if the statistical representativeness of the sample is built at the NUTS 1 level, it is valid only at the territorial level. The joint analysis of data at the regional (NUTS 2) and macro area (NUTS 1) levels must be accommodated with ad hoc statistical methods. For instance, as a simpler solution, NUTS 2 values can be equalled to the NUTS 1 average for all regions belonging to a specific macro area, or more appropriately, region-specific values can be estimated using more sophisticated small-area estimation models. However, an analysis conducted at the macro-area level would remain valid only on a potential, ideal ground because, in most countries, the level of practical decision-making corresponds to the national or regional level. Furthermore, official statistics are often commented on as if they come from census data and are free of sampling errors. However, this is not the case, and it must always be borne in mind that most of these statistics are estimates that originate from survey data, and that they have a degree of uncertainty due to their sampling nature. Estimates at the regional level have precision (think of the standard errors of the estimates) that is necessarily lower than at higher hierarchical levels, at least because of the smaller sample size. Therefore, moving from a national to a regional analysis requires an evaluation of the consistency of estimates to assess whether territorial differences can be considered statistically significant. In addition, evaluating the completeness of information may result in the selection of alternative measures because a given indicator may be inappropriate when applied to a regional context. This may happen, for example, in the case of an extension of a variable originally designed at the national level to the regional level. For instance, the indicator of women on boards of listed companies makes sense in a national analysis but not from a regional perspective. The economic reference area of a listed company cannot be ascribed to a specific region; in any case, this is hardly the region where the company has its head office. Datasets used to construct the nationwide indicator may not include, by design, information at the subnational level (e.g. HETUS). In addition, different privacy regimes for microdata protection may reduce the number of countries involved in cross-country analysis. Instead, the problem of time misalignment among different datasets, which occurs when the construction of a synthetic indicator involves different data sources, may alter the reference year of some (groups of) variables.

A similar discussion is valid for differences in time frequency. Table 1 provides an overview of the distinctive features of the Eurostat survey data. The four selected datasets cover different time spans. In addition, the time frequencies differed. It ranges from yearly to 4 (SES) or 10 years (HETUS). Ideally, one would like to merge

Table 1 Comparison of selected Eurostat survey datasets

Survey	Population	Content	Years (frequency)	Max NUTS detail	Notes
ECHP/ SILC	Households members	Income and living conditions	1994–2001/ 2004–2020 ^a (yearly)	NUTS 1 ^b NUTS 2	
LFS	Individuals aged 15+	Labour market status	1983–2020 (yearly)	NUTS 3	
HETUS	Households members	Daily activity	2010 ^c (ten years)	NUTS 0	
SES	Employees	Structure of earnings	2002–2018 (four years)	NUTS 1	

Notes: ^a Panel for 2005–2019; ^b Germany; ^c collected in 2008–2015

the contents of different datasets that share the same time span, time frequency, and territorial disaggregation. Failures in doing so lead to the inappropriateness of some Eurostat datasets or variables originally included in the design of the synthetic indicator. At the extreme, it can lead to structural changes in terms of domains and subdomains. Modifications in the structure of the indicator, for example, of the R-GEI (di Bella et al., 2021), are not immediate and straightforward tasks. Instead, they require full scrutiny of the content of the various survey questionnaires and, when replacements occur, coherence with the scope of the indicator.

4 From GEI to R-GEI in a Cross-Country Setting

As specified in the Introduction, this book stems from the EU Erasmus+ programme-funded project ReGem (Regional Gender Equality Measurement in the EU⁷), inspired by the work by di Bella et al. (2021) to investigate gender equality at the regional level in Italy, France, Spain, and Germany. Instead of proposing a brand-new gender equality index based on additional statistical sources than those proposed by the EIGE (Norlen et al., 2019), the ReGem project sought to understand to what extent the EIGE's GEI could be declined at the regional level by using the same data sources of GEI by constructing what di Bella et al. (2021) called Regional-GEI or R-GEI, but for a broader study area. In their work limited to Italy, di Bella et al. (2021) identified that only 10 out of 31 variables of EIGE's GEI could also be measured at the regional level using the same definitions and data as the original nationwide GEI. Other variables could not be measured at the regional level because they were meaningless if rescaled at that territorial level or were based on surveys not designed for regional representativeness. The authors identified 14 alternative variables using the same data sources as the GEI or data from the Italian Institute of Statistics (Istat). However, the authors did not find a suitable substitution for these

⁷Project website: regem.unige.it

seven variables. The domain that required the most relevant revisions was power because, from a regional perspective, most of the original indicators lost their relevance. For instance, the national-level indicator ‘percentage of women ministers or having a seat in parliament’ power subdomains has been replaced by a regional/local variant ‘share of women in regional boards’ or ‘share of municipal and regional women assessors and women city mayors’ using data from Italian administrations.

In the ReGem project, it was decided to make a replication of the variable selection process that di Bella et al. (2021) conducted for Italy for the four countries in the study area, but without adding country-specific data sources to the Eurostat databases. The final result (Table 2) substantially confirmed the results by di Bella et al. (2021), leading to the identification of 12 local-national equivalent variables, nine ‘close to the original’ alternatives derived from Eurostat databases, and 9 variables for which it was impossible to find, among Eurostat databases, valuable substitutes. Unfortunately, accessing micro-data from the EHIS survey was impossible due to the sensitive nature of the respondents’ health data and the availability of regional-level data limited to Italy and Germany. Furthermore, for Germany, most datasets, including the LFS and EU-SILC data, are only available at the NUTS 1 level; therefore, in the analyses presented in the next chapter, the country was excluded from the analysis.

5 Conclusions

Regional analysis is an important approach for identifying areas within a country in which the gender gap is most pronounced. Limiting the analysis to a national assessment may mask the extreme situations behind an intermediate value. However, going from a national to a regional analysis of gender inequality is a complex process, especially when conducting a cross-country analysis. The reformulation of variables measured at the national level at the regional level may run into issues that are not immediately resolvable. In some cases, it may happen that the national indicator does not make sense when defined at the local level or that the data used for the national indicator are not representative at the regional level. It is then necessary to assess whether a suitable variable exists to replace the original variable within the alternative datasets. This is often possible by using specific national surveys, but without any guarantee that, in a cross-country analysis, there is homogeneity of definition and measurement between different countries. It may also happen that no suitable variables are found to replace the original ones, and thus, there is only partial coverage of the regional index compared to the national one.

The ReGem project focused on four key countries of the European Union, verifying that, although in a context coordinated by Eurostat, there are differences between surveys and inhomogeneity of territorial detail between countries within the same survey. At present, if one wants to remain within the logic of the GEI, the reference indicator for comparing gender equality between EU countries, it seems inevitable to integrate Eurostat sources with data from national surveys, as done by

Table 2 R-GEI variables in EIGE'S GEI framework

RGEI Source at NUTS2*	Same as GEI NA	NA	Same as GEI NA	NA	EU-LFS (Ad hoc module 2018)	NA	EU-SILC	EU-LFS (Ad hoc module 2019)		
R-GEI description (measure, population)	Same as GEI NA	NA	Same as GEI NA	NA	People caring for and educating their children, elderly, or disabled people (%; 18+ population)	NA	Workers regularly attending leisure activities outside their home (cinema, concerts, sports) (%; 16+ workers)	Workers involved in voluntary or charitable activities during the last 12 months (%; 16+ population) (%; 16+ workers)		
GEI source	EU-LFS	EU-LFS	EU-LFS	EWCS	EWCS	EWCS	EWCS	EWCS		
GEI description (measure, population)	Full-time equivalent employment rate (%; 15+ population)	Duration of working life (years; 15+ population)	Employed people in education, human health, and social work activities (%; 15+ employed)	Ability to take one or two working hours off to take care of personal or family (%; 15+ workers)	Career prospects index (points 0-100)	People caring for and educating their children or elderly, or disabled people every day (%; 18+ population)	Workers doing sport, cultural or leisure activities outside their home several times a week or at least daily (%; 15+ workers)	Workers involved in voluntary or charitable activities at least once a month (%; 15+ workers)		
Sub-domain	Participation			Segregation and quality			Social activities			
Domain	Work			Time						
RGEI Source at NUTS2*	Same as GEI	Same as GEI	NA	NA	Same as GEI	Same as GEI	EU-SILC	Same as GEI	Same as GEI	Same as GEI
RGEI description (RGEI measure)	Same as GEI	Same as GEI	NA	NA	Same as GEI	Same as GEI	Monthly earnings (mean, working population 15-64)	Same as GEI	Same as GEI	Same as GEI
GEI source	EU SILC	EU mortality tables	EU-SILC	EHIS	EHIS	EU-SILC	EU-SES	EU-SILC	EU-SILC	EU-SILC
GEI description	Self-perceived good or very good	Life expectancy at birth	Healthy life expectancy at birth	People not smoking and not	People doing physical	Population without unmet needs for	Monthly earnings (mean, working)	Equivalised net income	Not-at-risk-of-poverty as 60% median income	Income distribution as S20/S80 %;

(continued)

Table 2 (continued)

RGEI Source at NUTS2*	Same as GEI	Same as GEI	Same as GEI	Same as GEI	NA	EU-LFS (Ad hoc module 2018)	NA	EU-SILC population (aged <65)	EU-LFS (Ad hoc module 2019) (mean, 16+ population)	EU-LFS (Ad hoc module 2019) (% 16+ population)	16+ population
(GEI measure)	NA (years, population)	involved in harmful drinking (% 16+ population)	activities and/or consuming fruits and vegetables (% 16+ population)	medical examination 16+ population	NA	for dental examination (% 16+ population)	EU-SILC population (aged <65)	EU-LFS (Ad hoc module 2019) (mean, 16+ population)	EU-LFS (Ad hoc module 2019) (% 16+ population)	16+ population	
Sub-domain	Status	Behaviour	Access	Financial resources	Economic resources						
Domain	Health	Money	Money	Money	Money						
RGEI source at NUTS2*	Same as GEI	Same as GEI	NA	EU-LFS	EU-LFS						
RGEI description (RGEI measure)	Same as GEI	Same as GEI	NA	Women in supervisory positions (large companies, 50+ employees)	Women in supervisory positions (finance and insurance sectors/employees)						
GEI source	EU-LFS	EU-LFS	EIGE database	EIGE database	EIGE database						
GEI description (GEI measure)	Graduates of tertiary education (% 15+ population)	People participating in formal and non-formal education and training (% 15+ population)	Share of tertiary students in the fields of education, health and welfare, humanities and art (% 15+ population)	Share of ministers (% women and men)	Share of regional assemblies (% women and men)						
Sub-domain	Attainment and participation	Segregation	Political	Economic	Social						
Domain	Knowledge	Power	Power	Economic	Social						

di Bella et al. (2021). However, the task seems very complex for a complete comparison across the 27 EU countries.

References

- Alesina, A., Giuliano, P., & Nunn, N. (2013). On the origins of gender roles: Women and the plough. *The Quarterly Journal of Economics*, 128(2), 469–530.
- Altavilla, C., & Caroleo, F. E. (2013). Asymmetric effects of national-based active labour market policies. *Regional Studies*, 47(9), 1482–1506.
- Amici, M., & Stefani, M. L. (2013). A gender equality index for the Italian regions. *Bank of Italy Occasional Paper*, 190.
- Anxo, D., Mencarini, L., Pailhé, A., Solaz, A., Tanturri, M. L., & Flood, L. (2011). Gender differences in time use over the life course in France, Italy, Sweden, and the US. *Feminist Economics*, 17(3), 159–195.
- Bericat, E. (2012). The European gender equality index: Conceptual and analytical issues. *Social Indicators Research*, 108(1), 1–28.
- Bericat Alastuey, E., & Sánchez Bermejo, E. (2008). *Balance de la desigualdad de género en España. Un sistema de indicadores sociales*. Centro de Estudios Andaluces.
- Blanchard, O. J., & Katz, L. (1992). Regional evolutions. *Brookings Papers on Economic Activity*, 1992, 1–75.
- Cascella, C., Williams, J., & Pampaka, M. (2022). An extended regional gender gaps index (eRGGI): Comparative measurement of gender equality at different levels of Regionality. *Social Indicators Research*, 159(2), 757–800.
- Castellano, R., & Rocca, A. (2019). Analysing the gender gap in European labour markets at the NUTS-1 level. *Cogent Social Sciences*, 5(1), 1595294.
- Costantini, V., & Monni, S. (2006). *Gender disparities in Italy from a human development perspective*. Roma Tre University, Rome.
- di Bella, E., Leporatti, L., Gandullia, L., & Maggino, F. (2021). Proposing a regional gender equality index (R-GEI) with an application to Italy. *Regional Studies*, 55(5), 962–973.
- Dijkstra, L. L., Papadimitriou, E., & Norlén, H. (2019). *The regional gender equality monitor measuring female disadvantage and achievement in EU regions*. Publications Office of the European Union.
- EIGE. (2017). *Gender equality index 2017: Methodological Report*. European Institute for Gender Equality.
- EIGE. (2022). *Gender equality index 2022: The COVID-19 pandemic and care*. European institute for gender equality.
- European Commission. (2022). *Cohesion in Europe towards 2050—8th cohesion report*. Publications Office of the European Union.
- Fernández-Sáez, J., Ruiz-Cantero, M. T., Guijarro-Garvi, M., Rodenas-Calatayud, C., Martí-Sempere, M., & Jiménez-Alegre, M. D. (2016). Tiempos de equidad de género: Descripción de las desigualdades entre comunidades autónomas, España 2006-2014. *Gaceta Sanitaria*, 30, 250–257.
- Fortin, N. M. (2005). Gender role attitudes and the labour market outcomes of women across OECD countries. *Oxford Review of Economic Policy*, 21(3), 416–438.
- Fuchs, M., Rossen, A., Weyh, A., & Wydra-Somaggio, G. (2021). Where do women earn more than men? Explaining regional differences in the gender pay gap. *Journal of Regional Science*, 61(5), 1065–1086.
- Gil-Lafuente, A. M., Torres-Martinez, A., Amiguet-Molina, L., & Boria-Reverter, S. (2019). Gender equality index of the autonomous communities of Spain: A multidimensional analysis. *Technological and Economic Development of Economy*, 25(5), 915–933.

- Gimenez-Nadal, J. I. M., & Alberto, J. (2022). The gender gap in time allocation. *IZA World of Labor*, 497.
- Gimenez-Nadal, J. I., & Molina, J. A. (2014). Regional unemployment, gender, and time allocation of the unemployed. *Review of Economics of the Household*, 12(1), 105–127.
- Gimenez-Nadal, J. I., & Molina, J. A. (2016). Commuting time and household responsibilities: Evidence using propensity score matching. *Journal of Regional Science*, 56(2), 332–359.
- Hippe, R., & Perrin, F. (2017). Gender equality in human capital and fertility in the European regions in the past. *Investigaciones de Historia Económica*, 13(3), 166–179.
- Iammarino, S., Rodriguez-Pose, A., & Storper, M. (2019). Regional inequality in Europe: Evidence, theory and policy implications. *Journal of Economic Geography*, 19(2), 273–298.
- Kaasa, A., Vadi, M., & Varblane, U. (2014). Regional cultural differences within European countries: Evidence from multi-country surveys. *Management International Review*, 54(6), 825–852.
- Kjeldstad, R., & Kristiansen, J. E. (2001). Constructing a regional gender equality index: Reflections on a first experience with Norwegian data. *Statistical Journal of the United Nations Economic Commission for Europe*, 18(1), 41–49.
- Kushi, S., & McManus, I. P. (2018). Gender, crisis and the welfare state: Female labor market outcomes across OECD countries. *Comparative European Politics*, 16(3), 434–463.
- Le Barbanchon, T., Rathelot, R., & Roulet, A. (2021). Gender differences in job search: Trading off commute against wage. *The Quarterly Journal of Economics*, 136(1), 381–426.
- Marelli, E. (2007). Specialisation and convergence of European regions. *The European Journal of Comparative Economics*, 4(2), 149–178.
- Martín, R. D., & Garvi, M. G. (2009). Gender inequality and economic growth in Spain: An exploratory analysis. *Review of Regional Studies*, 39(1), 23–48.
- Nisic, N. (2017). Smaller differences in bigger cities? Assessing the regional dimension of the gender wage gap. *European Sociological Review*, 33(2), 292–3044.
- Norlen, H., Papadimitriou, E. & Dijkstra, L. (2019). The regional gender equality monitor. Publications Office of the European Union, EUR 29679 EN.
- Peinado, J. M., & Céspedes, G. C. (2004). Gender and regional inequality in human development: The case of Spain. *Feminist Economics*, 10(1), 37–64.
- Perales, F., & Vidal, S. (2015). Looking inwards: Towards a geographically sensitive approach to occupational sex segregation. *Regional Studies*, 49(4), 582–598.
- Perrin, F. (2021). Can the historical gender gap index deepen our understanding of economic development? *Journal of Demographic Economics*, 88(3), 370–417.
- Petrongolo, B., & Ronchi, M. (2020). Gender gaps and the structure of local labor markets. *Labour Economics*, 64, 101819.
- Rönblom, M. (2005). Letting women in? Gender mainstreaming in regional policies. *Nordic Journal of Women's Studies*, 13(3), 164–174.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

