RUCH PRAWNICZY, EKONOMICZNY I SOCJOLOGICZNY
Rok LXXXVI – zeszyt 2 – 2024
https://doi.org/10.14746/rpeis.2024.86.2.12

MARCIN SZYMKOWIAK,[a] KAMIL WILAK[b]

SPREE ESTIMATION OF THE NUMBER OF DISABLED PEOPLE IN TERMS OF ECONOMIC ACTIVITY[1]

ESTYMACJA TYPU SPREE LICZBY OSÓB Z NIEPEŁNOSPRAWNOŚCIAMI POD WZGLĘDEM ICH AKTYWNOŚCI ZAWODOWEJ

The creation of equal opportunities in the labour market for people with disability remains a challenge in many countries around the world. The impact of disability, especially when it comes to work opportunities for disabled people, cannot be properly understood without access to the relevant statistics. Information about working-age disabled people is crucial in the development of labour market policy. Such information should be available not only at the national level but also at lower levels of spatial aggregation. The main aim of this article is to propose a way of producing reliable estimates of key labour market indicators (economic activity rate, employment rate and unemployment rate) for working-age disabled people in Poland at the province level. The authors apply SPREE estimation and use data from the Labour Force Survey 2011–2019 (LFS) and the 2011 Census to produce estimates characterized by better precision and stability over time than what can be achieved by applying direct estimation. While the scale of changes in the labour market situation of working-age disabled people was found to be similar to trends observed in the whole working-age population, there are differences in spatial patterns associated with these two groups.

Keywords: Labour Force Survey; disability; economic activity; small area estimation (SAE); Structure Preserving Estimation (SPREE)
JEL: C13, J60

Stworzenie równych szans dla osób z niepełnosprawnościami na rynku pracy pozostaje wyzwaniem dla wielu krajów na całym świecie. Lepsze zrozumienie zjawiska niepełnosprawności, zwłaszcza w odniesieniu do sytuacji osób z niepełnosprawnościami na rynku pracy, wymaga po-

[a] Poznań University of Economics and Business; Statistical Office in Poznań, Poland / Uniwersytet Ekonomiczny w Poznaniu; Urząd Statystyczny w Poznaniu, Polska marcin.szymkowiak@ue.poznan.pl, https://orcid.org/0000-0003-3432-4364

[b] Poznań University of Economics and Business; Statistical Office in Poznań, Poland / Uniwersytet Ekonomiczny w Poznaniu; Urząd Statystyczny w Poznaniu, Polska kamil.wilak@ue.poznan.pl, https://orcid.org/0000-0002-4305-6202

[1] Funding: (1) The work of Marcin Szymkowiak has been developed as part of a project entitled 'Indirect estimation of disability in the 2011 census', which was financed by Poland's National Science Centre from a grant awarded by virtue of decision no. DEC-2013/11/B/HS4/01472. (2) Financed by the Minister of Science under the 'Regional Initiative for Excellence' Programme.

RPEiS 86(2), 2024: 237–258. *ISSN (Online) 2543-9170 ISSN (Print) 0035-9629*

siadania odpowiednich danych statystycznych. W kontekście formułowania polityki rynku pracy kluczową rolę odgrywa informacja na temat osób z niepełnosprawnościami w wieku produkcyjnym. Ważne jest, aby taka informacja była dostępna nie tylko na poziomie całego kraju, ale również na niższym poziomie agregacji przestrzennej. Głównym celem artykułu jest uzyskanie wiarygodnych oszacowań najważniejszych wskaźników rynku pracy (współczynnik aktywności zawodowej, wskaźnik zatrudnienia i stopa bezrobocia) osób z niepełnosprawnościami w wieku produkcyjnym w Polsce na poziomie województw. W tym kontekście wykorzystano estymację typu SPREE na podstawie danych z Badania Aktywności Ekonomicznej Ludności 2011–2019 (BAEL) oraz Narodowego Spisu Powszechnego Ludności i Mieszkań 2011. Wykorzystanie metody SPREE pozwoliło na uzyskanie szacunków cechujących się wysoką precyzją i stabilnością w czasie, znacznie wyższą niż w przypadku klasycznego podejścia, jakim jest estymacja bezpośrednia. Uzyskane szacunki wskazują, że dynamika zmian sytuacji na rynku pracy osób z niepełnosprawnością w wieku produkcyjnym jest zgodna z tendencjami obserwowanymi w całej populacji osób w wieku produkcyjnym, natomiast nie zaobserwowano takich samych wzorców w ujęciu przestrzennym.

Słowa kluczowe: Badanie Aktywności Ekonomicznej Ludności; niepełnosprawność; aktywność ekonomiczna; statystyka małych obszarów; estymatory typu SPREE
JEL: C13, J60

I. INTRODUCTION

Nowadays, disability is regarded as a major social problem. Almost everyone is likely to temporarily or permanently experience disability at some point in life. According to the World Health Organization (2016), over one billion people across the world suffer from various dysfunctions that would fall under the definition of a disabled person. It is also estimated that nearly 200 million people experience severe disability. To make matters worse, globally, the number of people with disability is expected to increase. This trend is mainly due to population ageing. The seriousness of this issue has been recognized in the Millennium Development Goals report, where disability is identified as a problem which will require complex solutions (United Nations [UN], 2015a).

The problem of disability and its consequences are also addressed by the European Union. Statistical data indicate that around 16–17% of the European Union's residents have a slight to severe disability, which means that nearly 80 million people experience various functional limitations in their social life (European Commission [EC], 2010). Poland is no exception in this regard: according to the 2011 and 2021 censuses, around 12% and 14.3% of the population respectively suffer from disability.

In light of the above, in March 2021, the European Commission (2021) adopted the Strategy for the Rights of Persons with Disabilities 2021–2030. It is based on the results of the previous European Disability Strategy 2010–2020 aimed at improving the quality of life of people with disability, as well as their social inclusion and wellbeing, thus enabling them to fully exercise their rights in various areas of life related to health or working conditions (EC, 2010). The main objective of the strategy is to help eliminate barriers for people with disability by launching initiatives concerning eight key areas:

(1) Accessibility – people with disability must have access to goods, services, assistive devices, transport, facilities, information and communication technologies in the same way as able-bodied people;

(2) Participation – people with disability must be able to fully exercise their fundamental rights;

(3) Equality – active policies must be implemented in all EU Member States to promote equality and ensure legislation which counteracts discrimination based on disability;

(4) Employment – increase in the employment of disabled people, collaboration with social partners to foster intra-job mobility, improving the quality of jobs;

(5) Education and training – disabled pupils and students must benefit from an accessible education system and lifelong learning programmes;

(6) Social protection – measures to mitigate income inequalities, risks of poverty and social exclusion to which people with disability are subject;

(7) Health – people with disability must benefit from equal access to services and health facilities, including mental health facilities;

(8) External action – promotion of the rights of people with disability at the international level.

In order to implement this strategy the European Union and its Member States will have to improve the collection and processing of statistical data. More detailed information will also be required about disability at lower levels of spatial aggregation, and for more specific domains, to meet the growing demand for information from local authorities, the labour market and government institutions responsible for supporting disabled people in different areas. In particular, there is much need for detailed statistics about the economic activity of disabled people to inform numerous programmes and policies designed to facilitate access to employment and post-secondary education and increase the labour market participation of persons with disabilities (Turcotte, 2014).

Individuals with disabilities find it difficult to pursue sustainable careers (Peijen & Wilthagen, 2022). Moreover, disability has been shown to have a negative effect on employment and earnings, and to cause dramatic differences in labour market outcomes (Jones, 2008). People with disability are often excluded from the labour market, either because they are held back from accessing jobs or because they find it difficult to find and retain a job (EC, 2016). The socio-economic debate has also stressed the need to empower people with disability in the labour market (Mussida & Sciulli, 2015). In addition to severely limiting the labour market participation of disabled people themselves, disability can, in many cases, affect the employment opportunities of other household members (Calegari et al., 2022).

People with disability not only have low participation rates in employment but also in vocational education and training (Cocks et al., 2015). It is therefore necessary to promote the employment of people with disability and reduce social and economic disparities experienced by disabled people (Lo, 2012; Hanga et al., 2015; Wittenburg et al., 2013). Initiatives in this regard

are increasingly undertaken by local and government labour market agencies, which require statistics about the economic activity of disabled people at a relatively low level of aggregation or more detailed breakdowns in order to achieve these objectives. The need for disability statistics of sufficient quality to inform decisions and monitor the level of this phenomenon was recognized, among others, by 2030 Sustainable Development Agenda (UN, 2015b).

Given that the majority of regular statistical surveys that collect information about the economic activity of disabled people involve relatively small samples and do not guarantee the required quality when processed using traditional methods of estimation, it is necessary to resort to more advanced methods of small area estimation (SAE). The same problem exists in the case of the Polish Labour Force Survey (LFS), which is the main source of information about disabled people's participation in the labour market. Because sample sizes for relevant breakdowns are too small to apply classical estimation methods, disability statistics are usually published at a relatively high level of aggregation (at country or province level) or for very broadly defined domains. More detailed information at lower levels, including additional variables (labour market activity, level of education, etc.) are generally not available. A possible remedy can be found in the methodology of small area estimation, which provides techniques for estimating the parameters of more detailed domains (Rao & Molina, 2015).

II. LITERATURE REVIEW

The literature devoted to the analysis of disability using small area estimation techniques is quite rich. The National Center for Health Statistics in the US pioneered the use of synthetic estimation for developing state estimates of disability and other health characteristics from the National Health Interview Survey (National Center for Health Statistics, 1968). It was necessary to use synthetic estimators because sample sizes in most states were too small to provide reliable direct state estimates.

Statistics New Zealand[2] has prepared a very interesting Shiny application in R showing disability estimates for small areas. Most of them were produced by applying small area estimation techniques to data from two sources. The first one was the 2013 Household Disability Survey, which provides information about the prevalence of disability. The second one was the 2013 Census, which provides the official population count for New Zealand and population breakdowns by age, sex, and the place of usual residence. By combining data from these two sources and using SAE models, it was possible to produce estimates for areas smaller than regions and estimate the likelihood that an individual person was disabled. SAE techniques were used because the number of disabled respondents in the Household Disability Survey 2013 was too small

[2] https://statisticsnz.shinyapps.io/Disability/

to provide reliable estimates for areas smaller than regions. Estimates produced by means of the SAE methods were presented in breakdowns defined by five-year age groups, region and the type of disability (hearing, physical, psychological and vision).

An important application of a hierarchical latent class model for predicting small area disability counts from survey data was described by Fabrizi et al. (2015). The authors considered the estimation of the number of severely disabled people based on data from the Italian survey on 'Health conditions and appeal to Medicare'. One of the main challenges raised by the authors was the fact that the survey measures disability indirectly by a set of categorical questionnaire items concerning the respondent's ability to accomplish everyday tasks, which means that the variable of interest is not observed directly. In their article, the authors adopted a full Bayesian approach and small area estimation based on a latent class model, in which the probability of a person belonging to each latent class changes depending on the three covariates: sex, marital status and age.

Another very important approach to the problem of disability was used by Maples (2017), who proposed a small area model to improve the estimates of state-level disability from the Survey of Income and Program Participation (SIPP) conducted by the US Census Bureau. He extended the work on the logistic-normal unit-level small area model developed by Jiang and Lahiri (2001, 2006) to incorporate Kim and Rao's proposal (2012), which recommended using an independent large-scale survey that contains a common set of covariates to improve model prediction estimation when those covariates are not known at the level of population. The SIPP data were also used by You et al. (2014), who applied a bivariate Fay-Herriot model and a measurement error model to data from the American Community Survey (ACS) to improve disability estimates from the SIPP. ACS and Behavioral Risk Factor Surveillance System data were also used by Xie et al. (2020), who applied Bayesian hierarchical regression in the context of small area estimation to estimate the prevalence of disability for all U.S. counties.

Polish researchers have also conducted studies aimed at estimating the prevalence of disability using SAE methods. For instance, Klimanek et al. (2018) applied the SAE methodology to estimate the percentage of disabled people, in the legal and biological sense, at the level of districts (NUTS 4/LAU 1 units) in the Wielkopolskie Province by the level of education. This methodological exercise was based on data from the 2011 census and involved selected techniques of indirect estimation. Estimates obtained in the study revealed patterns of spatial variation in disability for the target domains with a greater degree of precision than could be achieved by direct estimation. Gołata and Dehnel (2021) assessed the quality of estimates of the number of disabled persons produced from the 2011 census data. The purpose of this comparative study was to identify similarities and discrepancies between the estimates, and to determine the size and source of these discrepancies. In their analysis the authors accounted for such aspects as the measurement methods, the definitions and criteria of disability, the fact that a response to the question was optional, and the quality of disability information obtained from various sources.

As demonstrated by the examples above, attempts to measure disability using small area estimation techniques are undertaken to meet the growing demand for information on this social phenomenon at low levels of spatial aggregation and for more detailed domains. On the whole, existing sources of statistical information in this area are relatively modest. This is also true for Poland, where disability statistics are published mainly at country level and for a limited number of breakdowns.

The following study was an attempt to assess the labour market participation of people with disability at a lower level of territorial division, namely at the province level.[3] More precisely, its goal was to produce reliable estimates of key labour market indicators (economic activity rate, employment rate and unemployment rate) of disabled people of working age (18–59 for women and 18–64 for men) in Poland at the province level. To do that, SPREE-type estimation was applied to data from the Labour Force Survey for 2011–2019 and the 2011 Census. The resulting estimates were characterized by better precision and were more stable over time than direct estimates produced by Statistics Poland.

III. DATA SOURCES

Data from the LFS for 2011–2019 and the 2011 Census were used to estimate selected characteristics of disabled people of working age at the level of province with additional breakdowns including the labour market status.

The LFS is one of the main sources of statistical data about the participation of disabled people in the labour market. The survey is designed to produce estimates of employment, unemployment and economic inactivity for the target population aged 15 and above. The sample includes people with a certified degree of disability or incapacity for work. Publications prepared in Poland that present statistics based on LFS data also contain disability statistics at the national level, broken down by labour market status (working, unemployed and economically inactive) and other variables. There are also counts of disabled people depending on the degree of disability (significant, moderate, small degree of disability or equivalent), household type and selected demographic characteristics. Given the sample size and the scope of the LFS, estimates for the subpopulation of people with disability of working age are not published by Statistics Poland at the province level or for more detailed domains. Another problem that needs to be taken into account in the estimation process[4] is that the question about disability is optional.

---

[3] Province, also called voivodship, is the highest administrative division in Poland, corresponding to NUTS 2 according to the nomenclature used in official statistics until 2017.

[4] For more information about the Polish LFS and its methodology see Statistics Poland (2012).

The second data source used in our analysis was the 2011 census, in which data were collected using a mixed-mode design, namely from administrative sources (registers and information systems) and directly from respondents in a sample survey and by means of complete enumeration. Data in the complete enumeration survey were collected via an online form or obtained from administrative sources and included basic socio-demographic and address information about people that was not included in the sample survey, the census of collective accommodation facilities or the survey of homeless people. The sample survey conducted during the 2011 census provided data which could not be obtained from registers and information systems. The survey involved a random sample of approximately 20% of all dwellings in the country. The survey was designed to collect information regarding six main thematic areas: socio-demographic characteristics, economic activity, internal and foreign migration, nationality and religious affiliation, households and families, buildings and dwellings. Within these thematic areas, more detailed surveys of specifically defined populations were also carried out. The category of socio-demographic information included a survey of the population of people with disability, which was divided into two groups: people with a valid certificate of disability issued by a competent body ('legal disability') and disabled people without a certificate but experiencing a completely or severely limited ability to perform basic activities typical for their age ('biological disability'). Since the LFS only contains information about people with 'legal disability', our analysis is limited to the labour market status of this group of disabled people.

IV. SPREE METHODOLOGY

SPREE estimators are a general class of synthetic estimators which rely on information from direct estimates (Rao & Molina, 2015). They are based on a well-known technique of reweighting in contingency tables, which is referred to as iterative proportional fitting (IPF). It involves correcting cell counts in a multi-dimensional contingency table so that the corrected values add up to the known marginal totals. Original cell counts in the contingency table can come from the last census, while the marginal totals come from reliable direct estimates, usually produced from survey data. SPREE estimators can be used to estimate totals for small areas in intercensal periods. Hence, they constitute a special class of estimators which can be used for presenting census results either for intercensal periods at low levels of spatial aggregation or for fine-grained domains (Józefowski & Szymkowiak, 2013).

Let us assume we are interested in producing a two-dimensional contingency table with r rows and c columns. It means that the finite population U is cross-classified into $r \times c$ cells. Moreover, let $N_{i,j}$ denote the total number of units in the (i, j)-th cell, where $i = 1, ..., r, j = 1, ... , c$. These numbers are assumed to be known from the last census or from an administrative register. Let $\hat{N}_{i.}$ and $\hat{N}_{j}$ denote known marginal totals which were estimated by applying the

the calibration estimator to data from the latest sample survey (e.g. the LFS). In other words, we assume that the marginal totals are reliable direct estimates. This case is shown in Table 1.

Table 1

Cell counts in an $r \times c$ contingency table

Level of X	Level of Y				Estimates
	Y_1	Y_2	...	Y_c	
X_1	$N_{1,1}$	$N_{1,2}$	...	$N_{1,c}$	$\hat{N}_{1.}$
X_2	$N_{2,1}$	$N_{2,2}$	...	$N_{2,c}$	$\hat{N}_{2.}$
...	...	...	...	...	...
X_r	$N_{r,1}$	$N_{r,2}$	...	$N_{r,c}$	$\hat{N}_{r.}$
Estimates	$\hat{N}_{.1}$	$\hat{N}_{.2}$	...	$\hat{N}_{.c}$	

Source: the authors' elaboration.

Since counts $N_{i,j}$ in Table 1 do not add up to the estimated marginal totals $\hat{N}_{i.}$ and $\hat{N}_{.j}$, $i = 1, ..., r, j = 1, ..., c$, they need to be adjusted to them. One of the most popular techniques to ensure consistency inside the contingency table is the IPF procedure. The main idea is to look for new cell counts $m_{i,j}$ inside the contingency table such that the following constraints are satisfied:

$$\widehat{N}_{i.} = \sum_{j=1}^{c} m_{i,j}\,, \quad i = 1, \ldots, r,$$

$$\widehat{N}_{.j} = \sum_{i=1}^{r} m_{i,j}\,, \quad j = 1, \ldots, c.$$

IPF involves a sequence of iteration cycles each consisting of two steps. In each step the algorithm ensures that internal cell counts add up to marginal values in rows and then in columns. This process is repeated until internal cell counts add up to all the marginal values. The following formulas are used to produce the adjusted values $m_{i,j}$ in Table 1 at the k-th step of the algorithm:

$$m_{i,j}^{(2k-1)} = \frac{m_{i,j}^{(2k-2)} \widehat{N}_{i.}}{\sum_{j=1}^{c} m_{i,j}^{(2k-2)}},$$

$$m_{i,j}^{(2k)} = \frac{m_{i,j}^{(2k-1)} \widehat{N}_{.j}}{\sum_{i=1}^{r} m_{i,j}^{(2k-2)}},$$

where $m_{i,j}^{(0)} = N_{i,j}$. This iterative algorithm is repeated until the internal cell counts are consistent with the marginal totals, assuming that $m_{i,j} = \lim_{k\to\infty} m_{i,j}^{(k)}$. The generalization of IPF for more complex contingency tables is quite straightforward.[5]

The relative estimation error (REE) of SPREE, which is a measure of precision, can be estimated by applying resampling methods, such as bootstrap.

V. THE ESTIMATION PROCEDURE

The main objective of the study was to estimate the number of working-age people with 'legal disability' across Poland's 16 provinces in terms of economic activity status (in employment, unemployed and economically inactive) by applying the SPREE approach to data from the LFS for 2011–2019 and the 2011 Census. The estimated totals were used to calculate the key labour force indicators (activity rate, employment rate and unemployment rate) for the population of disabled people at province level. Because we are interested in estimates broken down by region (province) and category (labour market status), they can be presented in the form of a 16 × 3 contingency table (see Table 2).

Table 2

The structure of contingency tables for SPREE estimation

Province	Labour market status			Estimates
	Employed	Unemployed	Inactive	
Dolnośląskie (DS)	$N_{1,1}$	$N_{1,2}$	$N_{1,3}$	$\hat{N}_{1.}$
Kujawsko-Pomorskie (KP)	$N_{2,1}$	$N_{2,2}$	$N_{2,3}$	$\hat{N}_{2.}$
Lubelskie (LU)	$N_{3,1}$	$N_{3,2}$	$N_{3,3}$	$\hat{N}_{3.}$
Lubuskie (LB)	$N_{4,1}$	$N_{4,2}$	$N_{4,3}$	$\hat{N}_{4.}$
Łódzkie (LD)	$N_{5,1}$	$N_{5,2}$	$N_{5,3}$	$\hat{N}_{5.}$
Małopolskie (MA)	$N_{6,1}$	$N_{6,2}$	$N_{6,3}$	$\hat{N}_{6.}$
Mazowieckie (MZ)	$N_{7,1}$	$N_{7,2}$	$N_{7,3}$	$\hat{N}_{7.}$
Opolskie (OP)	$N_{8,1}$	$N_{8,2}$	$N_{8,3}$	$\hat{N}_{8.}$
Podkarpackie (PK)	$N_{9,1}$	$N_{9,2}$	$N_{9,3}$	$\hat{N}_{9.}$
Podlaskie (PD)	$N_{10,1}$	$N_{10,2}$	$N_{10,3}$	$\hat{N}_{10.}$
Pomorskie (PM)	$N_{11,1}$	$N_{11,2}$	$N_{11,3}$	$\hat{N}_{11.}$
Śląskie (SL)	$N_{12,1}$	$N_{12,2}$	$N_{12,3}$	$\hat{N}_{12.}$

[5] For instance, the method of finding new counts in a three-way contingency table using the IPF algorithm can be found in Rao and Molina (2015).

Table 2 (*continued*)

Province	Labour market status			Estimates
	Employed	Unemployed	Inactive	
Świętokrzyskie (SK)	$N_{13,1}$	$N_{13,2}$	$N_{13,3}$	$\hat{N}_{13.}$
Warmińsko-Mazurskie (WN)	$N_{14,1}$	$N_{14,2}$	$N_{14,3}$	$\hat{N}_{14.}$
Wielkopolskie (WP)	$N_{15,1}$	$N_{15,2}$	$N_{15,3}$	$\hat{N}_{15.}$
Zachodniopomorskie (ZP)	$N_{16,1}$	$N_{16,2}$	$N_{16,3}$	$\hat{N}_{16.}$
Estimates	$\hat{N}_{.1}$	$\hat{N}_{.2}$	$\hat{N}_{.3}$	

Source: the authors' elaboration.

The marginal estimates $\hat{N}_{i.}$ and $\hat{N}_{j}$ ($i = 1, \ldots, 16, j = 1,2,3$) were obtained from the LFS for 2011–2019, separately for all nine years, while the internal counts come from the 2011 Census and are the same for each year of the reference period. In order to obtain the marginal estimates in Table 2, the LFS datasets required some additional processing. First, observations with item non-response were excluded so that the final dataset only contained individuals with known disability status. As a result, the calibration weights in each quarterly dataset had to be adjusted separately. New calibration weights were computed by applying the method originally used in the Polish LFS.[6] Quarterly estimates were obtained by summing up the calibration weights within quarterly datasets. Annual estimates were obtained by averaging the corresponding quarterly estimates. Estimates obtained in this way, later referred to as direct estimates, are denoted as $\hat{N}_{i.}$ and $\hat{N}_{j}$ in Table 2. The same procedure was used to produce marginal estimates for all nine years of the reference period.

Unit-level data from the survey conducted during the 2011 Census were prepared in a similar way and were used as internal cell counts $N_{i,j}$. Individuals without a defined disability status were also removed from the dataset and the calibration weights were re-adjusted using the original method of calibration.[7] Internal cell counts $N_{i,j}$ broken down by province and labour market status were then calculated as the sum of calibration weights.

The SPREE procedure involving the IPF algorithm was then applied to the contingency table containing the internal cell counts $N_{i,j}$ and the marginal estimates $\hat{N}_{i.}$, $\hat{N}_{j}$ ($i = 1, \ldots, 16, j = 1,2,3$) to estimate the number of disabled people in each province and for each category of the labour market status (48 domains) independently for each of the nine years. Finally, the key labour market indicators: the activity rate, the employment rate and the unemployment rate, were calculated using estimated totals $m_{i,j}$, later on referred to as SPREE estimates. These SPREE estimates were compared with the direct

[6] The calibration weights were constructed to ensure that quarterly samples correctly represented the number of individuals by age (12 age groups), sex, and the place of residence (rural vs urban) at country level up to 2015, and starting from 2016 – at province level. More details can be found in Statistics Poland (2017).

[7] The calibration procedure used in the 2011 Census is explained in detail in Szymkowiak (2019).

estimates obtained for the domains defined as interactions of province and labour market status, in order to see if there is an improvement in the precision of the estimates.

To assess estimation quality, precision measures were computed using the design-based bootstrapping technique, which involves re-sampling from the survey sample using the same settings as those in the original sample. The bootstrap procedure with 500 replications, which is used in the Polish LFS, was applied to the SPREE estimates estimation.[8] It should be noted that the bootstrap procedure accounts for the randomness of the estimator resulting from the random nature of the LFS sample, but does not account for the randomness in the sample data collected during the 2011 Census.

VI. RESULTS

As pointed out earlier, the number of people with 'legal disability' in each province by labour market status, and the key labour market indicators, were obtained using data from the Labour Force Survey for 2011–2019 and the National Census of Population and Housing 2011.

As described in the previous section, in order to conduct the SPREE estimation, direct estimates from the LFS are used as the marginal totals in Table 2, namely counts of people with disability in each province are used as row marginals $\hat{N}_{i\cdot}$, while counts of people with disability for each category of labour market status are used as column marginals $\hat{N}_{\cdot j}$. The method of direct estimation used in the study was the same as that used by Statistics Poland to produce official estimates, including estimates of labour market indicators for people with disability. Because there are no official estimates at the aggregation level chosen for our study, direct estimates of labour market indicators were also produced to serve as a benchmark for the SPREE estimates.

Table 3 shows descriptive statistics regarding the number of respondents with 'legal disability' in the LFS sample for each reference year and each category of labour market status (quarterly data for each year were aggregated).

The numbers in Table 3 should be interpreted as follows: 177 in the first row represents the smallest number of disabled people in employment recorded in one province in the annual LFS sample collected in 2011. In other words, the number of respondents in this category sampled in other provinces was higher: its median value was 305 and the maximum was 426. As can be seen, the number of disabled respondents included in the annual LFS samples was the smallest for the unemployed category, and the largest for the group of economically inactive. Moreover, the number of respondents with disability in each category of the labour market status declines from year to year. This is not due to a decline in the prevalence of disability but the result of a falling response rate and the sensitive nature of the question about legal disability. As can be expected,

[8] The bootstrap procedure used in the Polish LFS is explained in Statistics Poland (2017).

this fact has an impact on the quality of the estimates shown in Table 4, which contains descriptive statistics for the relative estimation errors of direct estimates and SPREE estimates of the three labour market indicators.

Table 3

Number of respondents in annual samples by labour market status

Year	Min	Q_1	Me	Q_3	Max
Employed					
2011	177	238	305	423	426
2012	170	275	306	403	443
2013	164	215	281	342	400
2014	173	207	259	306	378
2015	123	198	230	260	335
2016	129	170	208	280	425
2017	132	150	197	255	521
2018	101	140	171	213	490
2019	73	129	159	203	436
Unemployed					
2011	26	47	52	66	77
2012	34	47	53	69	82
2013	20	45	59	63	80
2014	23	34	45	53	72
2015	13	25	32	36	46
2016	16	20	28	33	38
2017	11	15	19	27	34
2018	7	11	15	18	23
2019	3	9	13	15	24
Inactive					
2011	790	948	1,038	1,113	1,443
2012	736	933	966	1,112	1,427
2013	709	824	919	984	1,303
2014	674	732	866	958	1,206
2015	632	704	775	850	1,066
2016	533	629	684	770	865
2017	363	429	589	658	910
2018	336	359	516	582	888
2019	273	303	466	561	778

Source: the authors' elaboration based on LFS.

Two main conclusions can be drawn from the analysis of the data in Table 4. Firstly, compared to the direct estimator, the SPREE estimator is characterized by a higher precision, as evidenced by smaller relative estimation errors. This is true for all reference years and for all three labour market indicators.

Secondly, this gain in precision is particularly evident in the case of the unemployment rate. In fact, direct estimates of the unemployment rate are unreliable, as indicated by excessively high values of REEs, which often exceed the threshold of 10 or even 20%. In contrast, the relative estimation errors of all SPREE estimates do not exceed the level of 10%, and in many cases are at 3–6%. In other words, only SPREE estimates of labour market indicators can be regarded as acceptable.

Table 4

Relative estimation errors (REE) for labour market indicators (in %)

Year	REE of the direct estimator					REE of the SPREE estimator				
	Min	Q_1	Me	Q_3	Max	Min	Q_1	Me	Q_3	Max
Activity rate										
2011	3.9	4.4	4.7	5.2	6.9	1.2	1.2	1.2	1.3	1.3
2012	3.8	4.4	4.7	5.2	7.1	1.2	1.2	1.3	1.3	1.4
2013	4.1	4.6	5.0	5.4	7.6	1.2	1.2	1.2	1.3	1.3
2014	4.2	4.6	5.4	5.8	6.3	1.2	1.3	1.3	1.3	1.4
2015	4.4	5.3	5.6	6.2	8.4	1.4	1.5	1.5	1.6	1.7
2016	4.1	5.4	5.8	6.5	9.5	1.6	1.6	1.7	1.7	1.8
2017	3.9	5.4	6.1	7.1	9.1	1.5	1.6	1.6	1.6	1.7
2018	4.2	5.9	6.5	7.5	9.6	1.7	1.8	1.8	1.8	2.0
2019	4.8	6.2	6.7	7.8	10.4	1.8	1.9	1.9	2.0	2.1
Employment rate										
2011	4.0	4.9	5.3	5.8	8.0	1.3	1.4	1.4	1.4	1.5
2012	4.3	4.7	5.4	5.8	7.9	1.3	1.4	1.4	1.4	1.5
2013	4.7	5.0	5.8	6.1	8.2	1.4	1.4	1.4	1.5	1.5
2014	4.6	5.1	5.8	6.6	6.9	1.3	1.4	1.4	1.5	1.5
2015	4.8	5.7	6.1	6.6	9.0	1.6	1.7	1.7	1.7	1.8
2016	4.3	5.9	6.4	7.1	10.1	1.8	1.8	1.9	1.9	2.0
2017	4.0	5.7	6.6	7.5	9.7	1.6	1.7	1.7	1.7	1.8
2018	4.3	6.3	7.0	7.8	9.7	1.7	1.8	1.8	1.9	2.0
2019	4.8	6.5	7.2	8.5	11.5	1.9	2.0	2.1	2.1	2.2

Table 4 (*continued*)

Year	REE of the direct estimator					REE of the SPREE estimator				
	Min	Q_1	Me	Q_3	Max	Min	Q_1	Me	Q_3	Max
Unemployment rate										
2011	10.0	11.6	12.4	14.1	18.7	3.2	3.2	3.3	3.4	3.5
2012	10.0	10.6	12.3	13.1	16.3	2.9	3.0	3.0	3.1	3.2
2013	9.7	11.7	12.4	13.8	23.5	3.2	3.3	3.4	3.5	3.5
2014	10.5	12.9	13.9	15.9	21.5	3.5	3.5	3.6	3.7	3.8
2015	13.1	15.6	16.8	18.1	27.8	4.5	4.6	4.7	4.8	4.8
2016	14.3	17.5	18.9	21.1	27.1	5.2	5.2	5.3	5.4	5.5
2017	16.4	18.6	23.2	25.3	33.5	5.9	6.0	6.0	6.1	6.2
2018	21.0	25.1	27.7	30.5	39.2	7.2	7.2	7.3	7.4	7.5
2019	20.7	25.0	28.0	31.7	59.3	7.8	7.8	7.9	8.0	8.1

Source: the authors' elaboration based on LFS and the 2011 Census.

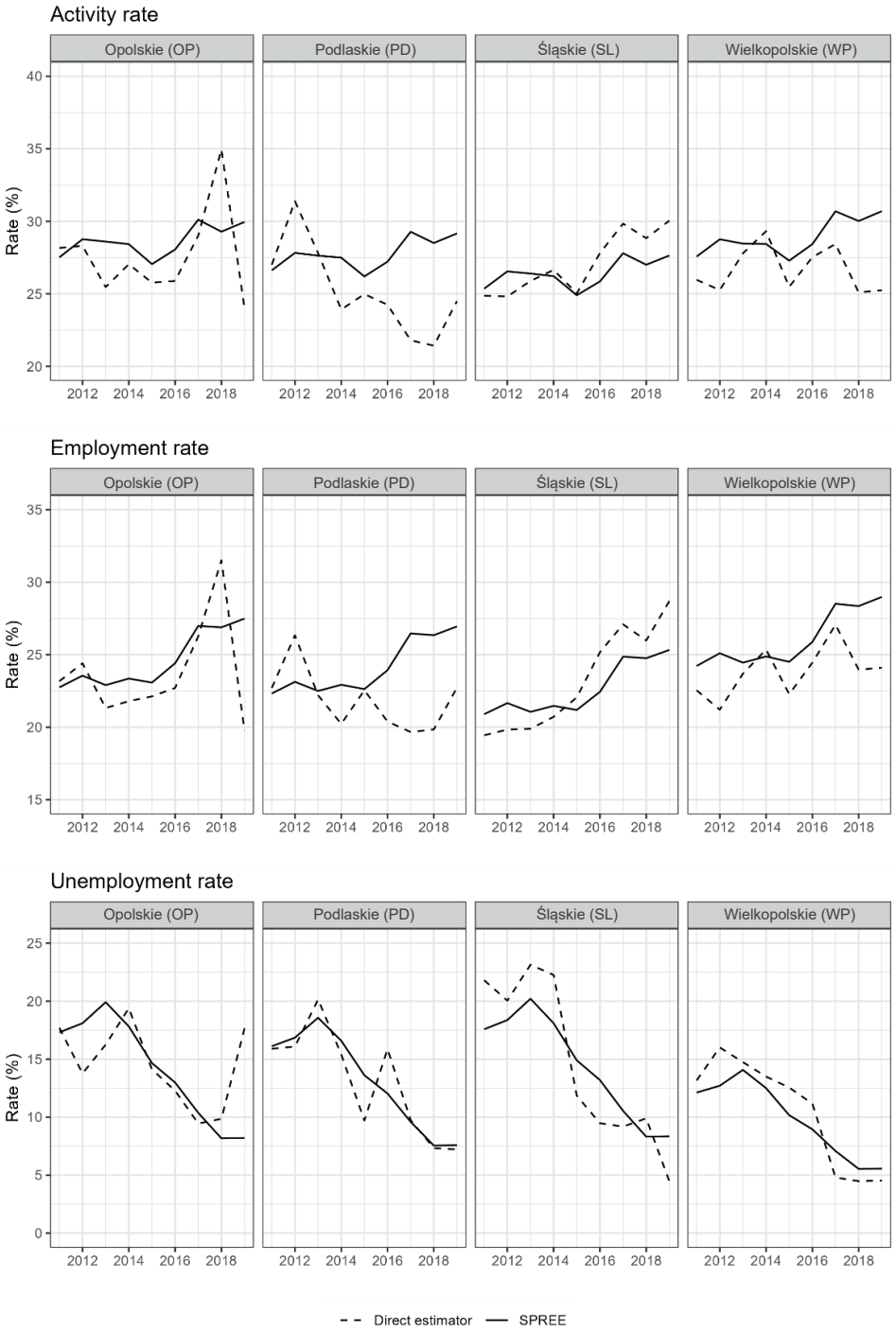
Figures 1 and 2 present how each labour market indicator changes over the course of the reference period. Figure 1 presents direct and SPREE estimates of the three indicators for four selected provinces throughout the reference period.

As indicated earlier, the direct estimates, in addition to their poor quality, particularly in the case of the unemployment rate, are also very unstable over time, in contrast to the SPREE estimates, which are highly stable over time and represent a clear upward trend in the activity rate and the employment rate, and a decline in the unemployment rate. Such results indicate an improvement in the labour market situation of people with legal disability. This conclusion is confirmed by the changing levels of SPREE estimates of all three labour market indicators shown in Figure 2. The choropleth maps represent changes in the labour market participation of people with legal disability that took place across the provinces between 2011 and 2019. As can be seen, after eight years the values of the economic activity rate and the employment rate increased while those of the unemployment rate fell across all provinces.

No spatial patterns can be observed within the three indicators. For example, in the case of the economic activity rate and the employment rate, there is a large disparity between the values of these indicators in two neighbouring provinces, namely Lubuskie (LB) and Zachodniopomorskie (ZP), located in the western part of the country, with the highest and lowest values respectively. A similar situation can be observed in Eastern Poland, where there is a notable disparity in the levels of both indicators between Podkarpackie (PK) and Lubelskie (LU). Differences between the values of these indicators in the rest of the country are much less pronounced. In the case of the unemployment rate, the lowest level can be observed in Wielkopolskie (WP).

Figure 1

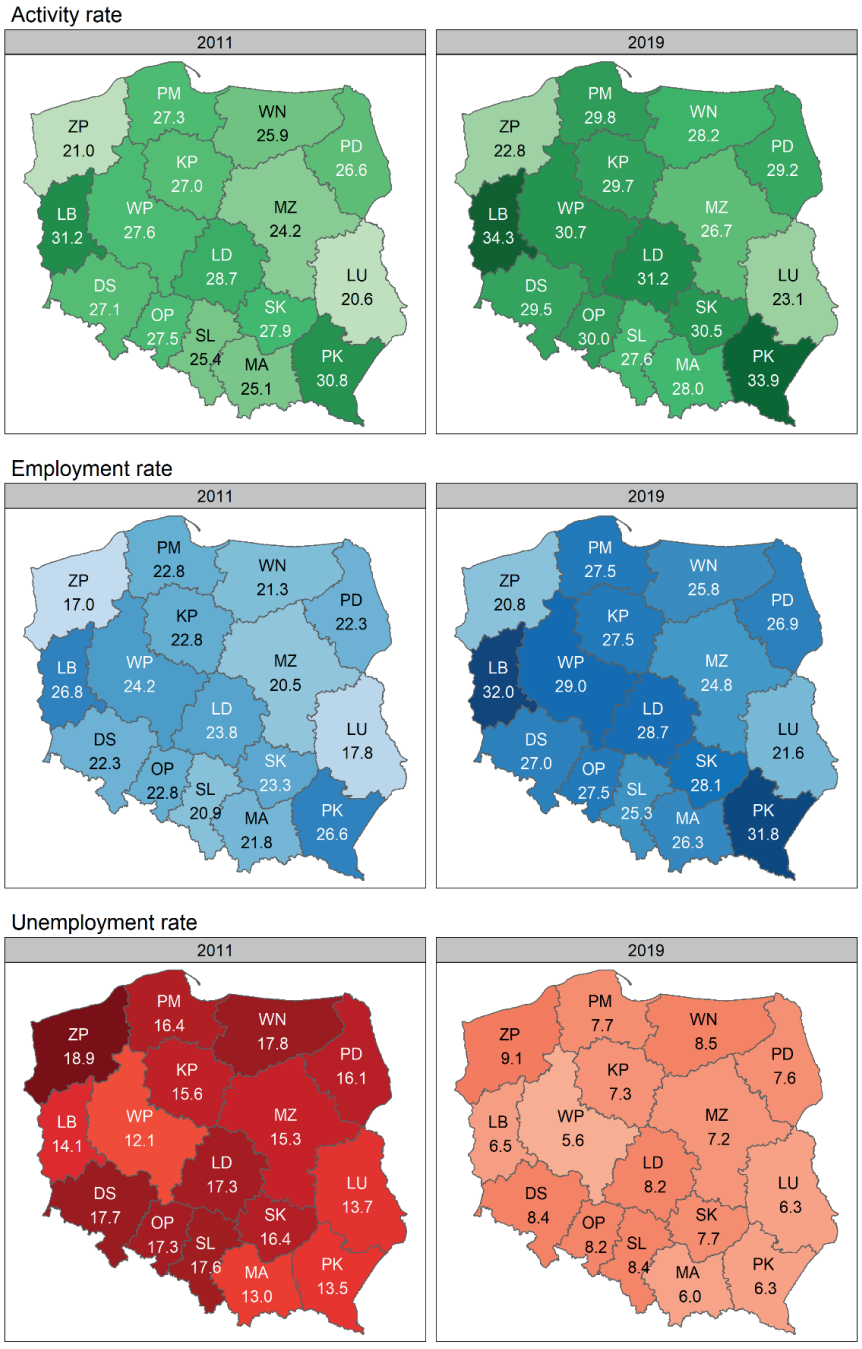
Direct (LFS-based) and SPREE estimates of labour market indicators for disabled people of working age in selected provinces between 2011 and 2019



Source: the authors' elaboration based on LFS and the 2011 Census.

Figure 2

Labour market indicators for disabled people of working age between 2011 and 2019 based on SPREE estimates



Source: the authors' elaboration based on LFS and the 2011 Census.

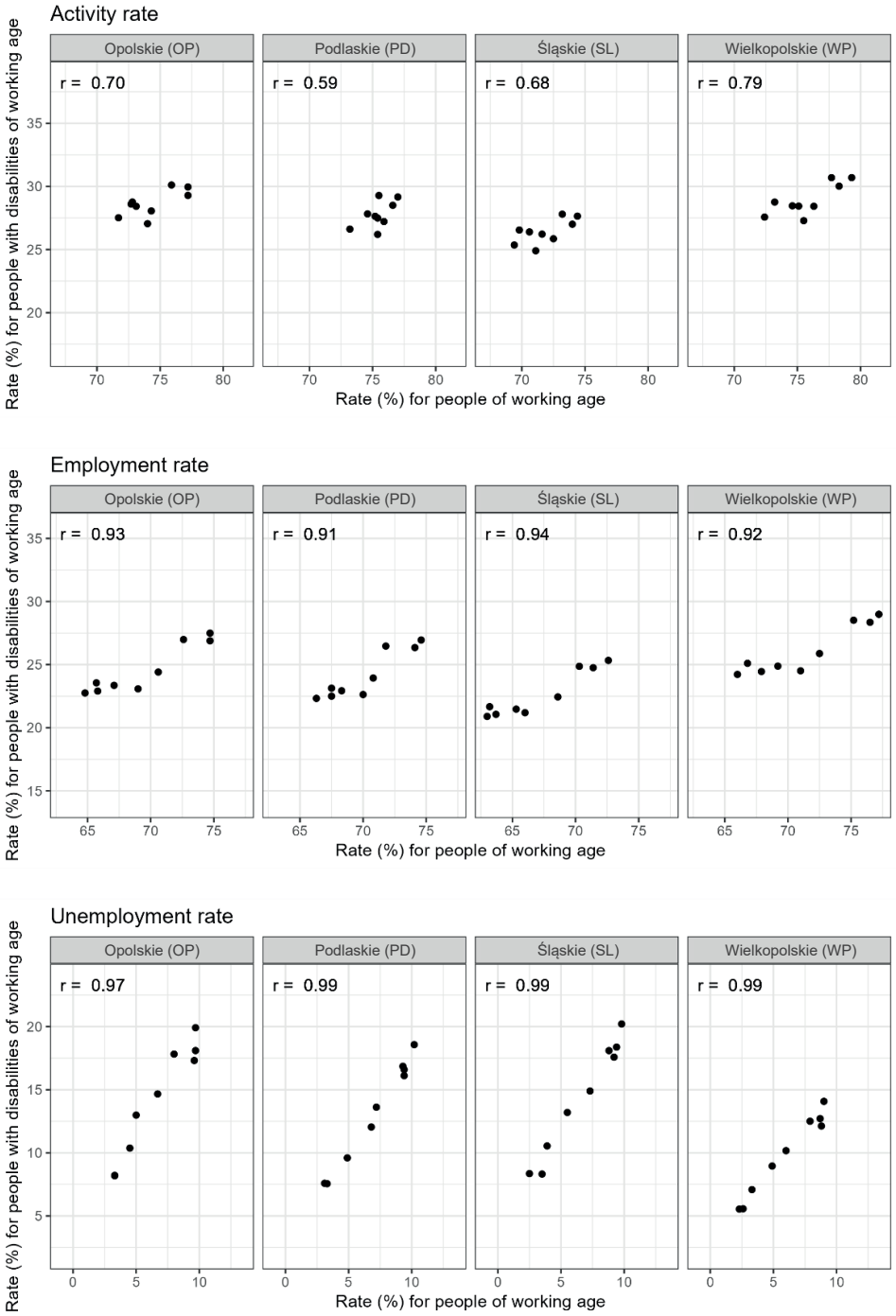
This is consistent with other labour market data for this province, which has had one of the lowest unemployment rates in Poland for many years. The next three provinces with the lowest unemployment rate, namely Małopolskie (MA), Podkarpackie (PK) and Lubelskie (LU), are located in the south-eastern part of the country, where the labour market situation is known to be different.

Figure 3 shows correlations between the SPREE estimates of the three indicators (economic activity rate, employment rate and unemployment rate) for disabled people of working age and direct estimates published by Statistics Poland for all people of working age (with and without disability) for four selected provinces. Each of the nine points in the charts represents one reference year (2011–2019). In addition, each chart shows the value of Pearson's correlation coefficient, which measures the linear correlation between the two sets of estimates. As can be seen, regardless of which indicator or province is considered, there is a strong positive correlation between the SPREE estimates for people with disability and the direct estimates published by Statistics Poland for all people of working age. This means that the labour market situation of people with disability is strongly correlated with the general labour market situation. In other words, changes in the labour market situation of people with disability are consistent with the trends observed in the whole working age population. It is worth noting that this type of relationship would not be observed if the direct estimation approach was applied to people with disability (owing to the low precision of estimates at province level). As shown in Figure 1, direct estimation is characterized by high instability over time (high fluctuations), which is inconsistent with trends observed in the labour market as a whole.

Figure 4 shows the correlations between the SPREE estimates of the three labour market indicators for disabled people of working age and the official direct estimates published by Statistics Poland for all people of working age in 2011 and 2019. This time, each point in the charts represents one province. As we can see, no correlation can be seen between the two sets of estimates regardless of which year is considered. For example, the chart for 2019 shows six provinces – Lubuskie (LB), Łódzkie (LD), Małopolskie (MP), Śląskie (SL), Świętokrzyskie (SK), Zachodniopomorskie (ZP) – for which direct estimates of the economic activity rate range from 74.4 to 75.6%, while the corresponding SPREE estimates for disabled people display much more variation: from 22.8% for Zachodniopomorskie (ZP) to 34% for Lubuskie (LB). In other words, no spatial patterns can be identified for either of the reference years involving the official estimates of key labour market indicators for all people of working age published by Statistics Poland and the SPREE estimates for people with disability.

Figure 3

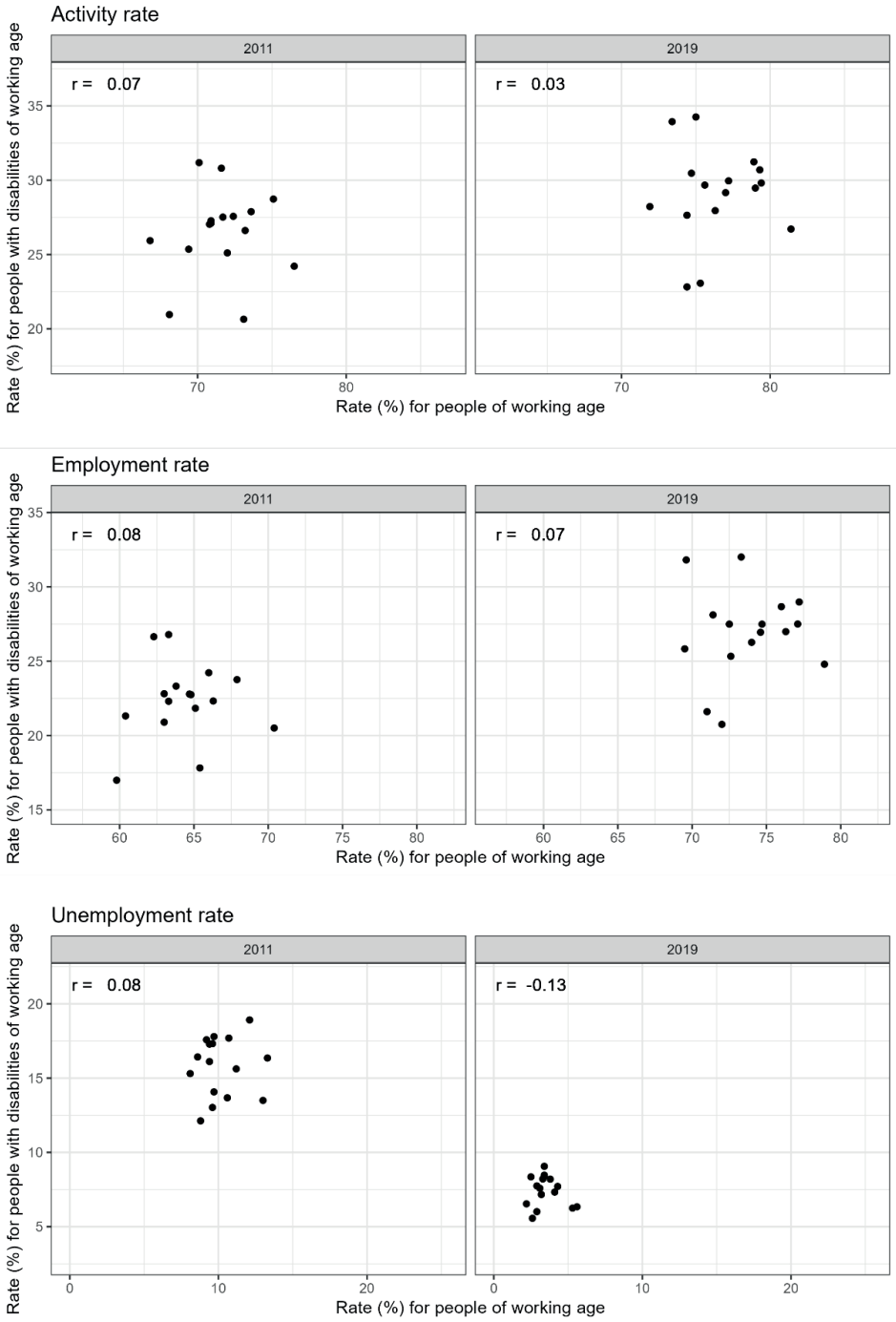
A comparison between SPREE estimates of labour market indicators for disabled people of working age and official direct estimates for all people of working age in selected provinces for all years



Source: the authors' elaboration based on LFS and the 2011 Census.

Figure 4

A comparison of SPREE estimates of labour market indicators for disabled people of working age and official direct estimates for all people of working age in selected years



Source: the authors' elaboration based on LFS and the 2011 Census.

VII. CONCLUSIONS

Disability is a major economic, social and health problem, which is increasingly affecting the functioning of modern societies. As various studies have shown, around 15% of the world's population live with disabilities, and this percentage is growing, mainly as a result of civilizational changes and a continuous rise in life expectancy, which is connected with population ageing. Other factors that contribute to this trend include cultural changes that have an impact on people's physical and intellectual fitness and patterns of social behaviour. The same process are underway in Poland, where the results of recent censuses indicate that the share of people with disability has increased from 12% in 2011 to around 14% in 2021.

The labour force participation of people with disability is also changing. Despite limitations caused by their impairments and architectural barriers, people with disability are increasingly eager to take part in cultural, social and sports activities. They also want to be economically active, as evidenced by the results presented in this article. Between 2011 and 2019, the economic activity rate and the employment rate for people with disability increased across all provinces of Poland, while the unemployment rate declined.

The need to estimate key characteristics of the labour market for people with disability is increasingly relevant given the scale of this phenomenon. A comprehensive, horizontal, cross-sectoral public policy to support people with disability, which takes into account their needs for independent living and social inclusion, requires up-to-date, reliable and detailed statistical data. Although Statistics Poland is taking steps to better diagnose the scope of disability and create tools for collecting statistical information in this area, work is still needed to produce reliable estimates at lower level of spatial aggregation and for more detailed domains. As we have shown in this article, this demand can be met by applying SPREE estimators and using data from different sources (2011 Census and LFS 2011–2019). The proposed approach was successfully applied to estimate the number of working-age people with legal disability by the labour market category and key labour market indicators for this group at the province level. As already mentioned, our estimates represent the level of aggregation for which no official statistics have been available so far, since direct estimates based on LFS data have relatively high estimation errors for such detailed domains. It is also worth noting that the method of estimation described in this article, compared to that presented in the literature (see Section II), is characterized by less computational complexity and is generally more intuitive.

More reliable estimates obtained by employing the SPREE approach are therefore likely to increase the effectiveness of social policy aimed at improving the quality of life for people with disability and can be used to evaluate and monitor measures taken by various institutions, including those tasked with facilitating disabled people's participation in the labour market.

The approach proposed in the article can be used to produce estimates for intercensal years when LFS data do not provide reliable estimates of the economic activity of people with disability. However, when relevant data from the latest census are used in the SPREE method, they may produce discontinuities in the estimates. This problem will be the subject of future studies.

Other directions for future research include the application of various extensions of the SPREE approach developed in the literature, which are based on generalized linear structural models (GLSM). Thanks to these extended approaches, interactions can be more flexibly transferred from a contingency table based on census data to estimates in intercensal years.

References

Calegari, E., Fabrizi, E., & Mussida, C. (2022). Disability and work intensity in Italian households. *Review of Economics of the Household*, *20*(2), 533–552. https://doi.org/10.1007/s11150-021-09559-6

Cocks, E., Thoresen, S., & Lee, E. A. L. (2015). Pathways to employment and quality of life for apprenticeship and traineeship graduates with disabilities. *International Journal of Disability, Development and Education, 62(*4), 422–437. https://doi.org/10.1080/1034912X.2015.1025714

European Commission. (2010). *The EU disability strategy 2010–2020: A renewed commitment to a barrier-free Europe*. Brussels, 15.11.2010. COM(2010) 636 final.

European Commission. (2016). *Disability and labour market integration*. Publications Office of the European Union.

European Commission. (2021). *Union of equality: Strategy for the rights of persons with disabilities 2021–2030*. Publications Office of the European Union.

Fabrizi, E., Montanari, G. E., & Ranalli, G. M. (2015). A hierarchical latent class model for predicting disability small area counts from survey data. *Journal of the Royal Statistical Society Series A: Statistics in Society, 179*(1), 103–131. https://doi.org/10.1111/rssa.12112

Gołata, E., & Dehnel, G. (2021). Credibility of disability estimates from the 2011 population census in Poland. *Statistics in Transition new series, 22*(2), 41–65. https://doi.org/10.21307/stattrans-2021-016

Hanga, K., DiNitto, D. M., & Wilken, J. P. (2015). Promoting employment among people with disability: Challenges and solutions. *Social Work and Social Sciences Review, 18*(1), 31–51. https://doi.org/10.1921/swssr.v18i1.847

Jiang, J., & Lahiri, P. (2001). Empirical best prediction for small area inference with binary data. *Annals of the Institute of Statistical Mathematics, 53*, 217–243. https://doi.org/10.1023/A:1012410420337

Jiang, J., & Lahiri, P. (2006). Estimation of finite population domain means: A model-assisted empirical best prediction. *Journal of the American Statistical Association, 101*, 301–311. http://www.jstor.org/stable/30047458

Jones, M. K. (2008). Disability and the labour market: A review of the empirical evidence. *Journal of Economic Studies*, *35*(5), 405–424. https://doi.org/10.1108/01443580810903554

Józefowski, T., & Szymkowiak, M. (2013). Zastosowanie estymatora typu SPREE w szacowaniu liczby osób bezrobotnych w przekroju podregionów [Using a spree estimator to estimate the number of unemployed people across subregions]. *Studia Oeconomica Posnaniensia*, *1*(10), 120–136.

Kim, J., & Rao, J. N. K. (2012). Combining data from two independent surveys: A model-assisted approach. *Biometrika*, *99*, 85–100. https://doi.org/10.1093/biomet/asr063

Klimanek, T., Szymkowiak, M., & Józefowski, T. (2018). Badanie zjawiska niepełnosprawności w przekroju powiatów województwa wielkopolskiego z wykorzystaniem metod statystyki małych obszarów [A study of disability across districts of the Province of Wielkopolska

using small area estimation methods]. *Przegląd Statystyczny, 65*(4), 449–472. https://doi.org/10.5604/01.3001.0014.0604

Lo, V. I. (2012). Promotion of the employment of persons with disabilities in Japan, the United States, and China: Carrot, stick, or both. *Arizona Journal of International & Comparative Law*, *29*(3), 557–602. http://arizonajournal.org/wp-content/uploads/2015/10/Lo_Final_3-20-13.pdf

Maples, J. J. (2017). Improving small area estimates of disability: Combining the American Community survey with the survey of income and program participation. *Journal of the Royal Statistical Society Series A: Statistics in Society, 180*(4), 1211–1227. https://doi.org/10.1111/rssa.12310

Mussida, C., & Sciulli, D. (2015). Direct and indirect effects of disability on employment probabilities: A comparative analysis. DISCE – Quaderni del Dipartimento di Scienze Economiche e Sociali dises1507. Università Cattolica del Sacro Cuore, Dipartimenti e Istituti di Scienze Economiche (DISCE).

National Center for Health Statistics. (1968). Synthetic state estimates of disability. PHS Publication, no. 1759. U.S. Government Printing Office.

Peijen, R., & Wilthagen, T. (2022). Labour market reintegration of individuals with a physical and cognitive disability by a company-based work-experience program. *International Journal of Disability, Development and Education, 69*(5), 1–18. https://doi.org/10.1080/103491 2X.2020.1802648

Rao, J. N. K., & Molina, M. (2015). *Small area estimation*. Wiley.

Statistics Poland. (2017). *Labour force survey in Poland: IV quarter 2016*. https://stat.gov.pl/en/topics/labour-market/working-unemployed-economically-inactive-by-lfs/labour-force-survey-in-poland-in-4th-quarter-2016,2,23.html

Statistics Poland. (2022). *Labour force survey in Poland: II quarter 2021*. https://stat.gov.pl/en/topics/labour-market/working-unemployed-economically-inactive-by-lfs/labour-force-survey-in-poland-quarter-22022,2,46.html

Szymkowiak, M. (2019). *Podejście kalibracyjne w badaniach społeczno-ekonomicznych* [The calabration approach in socio-economic surveys]. Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu.

Turcotte, M. (2014). *Persons with disabilities and employment*. Statistics Canada.

United Nations. (2015a). *The Millennium Development Goals Report 2015*.

United Nations. (2015b). *Transforming our world: The 2030 Agenda for Sustainable Development*.

Wittenburg, D., Mann, D. R., & Thompkins, A. (2013). The disability system and programs to promote employment for people with disability. *IZA Journal of Labor Policy, 2*(1), 1–25. https://doi.org/10.1186/2193-9004-2-4

World Health Organization. (2016). *World Report on Disability*.

Xie, H., Barker, L. E., & Rolka, D. B. (2020). Incorporating design weights and historical data into model-based small-area estimation. *Journal of Data Science, 18*(1), 115–131. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7182002/

You, J., Datta, G. S., & Maples, J. J. (2014). *Modeling disability in small areas: An area-level approach of combining two surveys*. Research Report Series, Statistics. Center for Statistical Research & Methodology. https://www.census.gov/library/working-papers/2014/adrm/rrs2014-11.html