Abstract for a proposed contribution to the 7th European User Conference for EU-Microdata, Mannheim, March 25-26, 2021

Cross-national income differences in sectors and occupations producing “societally necessary services”

The research project GenDis\(^1\) analyses working conditions in sectors and occupations which are deemed essential to individual welfare in modern societies. Against the backdrop of long-standing trends like demographic ageing, the project seeks to extend the existing knowledge on what can be done to secure a sufficient labour supply in these parts of the economy. Recently, the Coronavirus pandemic has drawn increased public attention to the contradiction between the “systemic importance” of certain sectors and occupations and the poor rewards granted to their protagonists.

Part of the research programme of the GenDis project is a comparative analysis of income from employment of the concerned groups of workers in Germany and other European countries. The contribution proposed for the 7th European User Conference for EU-Microdata presents work in progress and has three parts:

1. It explains how workers producing “societally necessary services” can be identified in European microdata using the NACE and ISCO variables.
2. It discusses which of the existing harmonised datasets are suited for comparative analysis of employment income for specific sectors and occupations, given their respective strengths and weaknesses.
3. It presents first descriptive results on employment incomes of selected groups of workers in selected countries, based on recent EU-LFS data.

Ad 1: Our identification strategy of workers that produce what we call “societally necessary services” draws on a broad literature on that concept and on related concepts. It is described in Lehweß-Litzmann et al. (2020). We choose sectors that seem to form a consensus in that literature and identify the relevant NACE codes. Knowing that any such list may vary between different countries and historical periods, ours may be valid for present-day Germany and other Western countries. As our project is mainly interested in the question of potential shortages in parts of the labour market, we focus on services that can hardly be automated and will (probably) keep on being produced by

\(^1\) http://www.sofi-goettingen.de/projekte/gesellschaftlich-notwendige-dienstleistungen-sicherstellen-ist-arbeit-am-gemeinwohl-attractiv-gendid/projektinhalt/
human workers. Within the chosen sectors, we thus pick occupations delivering personal services, because they often require an interaction between provider and client in the process of service delivery (which sets limits to automatisation). A third criterion we use is that the occupations chosen should be intrinsically related to the core product of their sector. We identify the relevant categories in the ISCO classification system. We thus pick the cases to be analysed from an imagined grid structure of sectors and occupations. Both the NACE and the ISCO variable are contained in datasets relevant for income analysis.

Ad 2: If we abstract from purely national data sources for comparability reasons, a comparative analysis of employment income in Europe can be based on one of the data products offered by Eurostat: the Structure of Earnings Survey (SES), the Labour Force Survey (LFS), the Statistics on Income and Living Conditions (SILC), and potentially also the Labour Cost Survey (LCS). In addition, there is the Luxembourg Income Study Database (LIS), provided by an independent civil society organisation. These sources differ in terms of the income concepts used (Schröder and Schwarzhappel 2020) and have their respective strengths and weaknesses (coverage, sample size, share of non-response, degree of detail in relevant classification variables, data quality, degree of harmonisation, documentation, etc.). For an analysis by sector and occupation, we absolutely need a data source with a sufficient number of cases within the relevant combinations. Also, the data has got to be sufficiently harmonised with regard to industry, occupation and the income variable, and ideally, we should be able to verify this way of looking at available pieces of documentation.

Ad 3: Based on the EU-LFS, first descriptive results will be presented for a subset of countries with sufficiently reliable income information. Like for other European datasets, countries participating in the LFS differ sharply in their way of collecting information. Unlike other datasets, income information of the LFS is distributed to researchers only in the form of deciles. This poses additional obstacles to data reliability due to potential effects of non-random non-response (in case of interviews). Ideally, the income data thus derives either from register data, which holds for a minority of participating countries, or if from interviews, deciles should be defined on the basis of external sources (which do not suffer from non-response, like administrative data). In countries that form deciles based on the LFS data itself, it is necessary to ensure that respondents did not tend to round numbers, as this poses problems with cut-off values (Eurostat 2012, 15). In case data is collected by income bands, it needs to be ensured that the number of categories offered to respondents are sufficiently fine-grained, thus significantly exceed the number of deciles (i.e. 10). Countries that satisfy these criteria will be used for a descriptive comparison of earnings of sectoral-occupational groups producing “societally necessary services”. Some adjustment, e.g. for qualification and hours worked, will be done. Potentially, first multivariate analyses will also be ready to present by the time of the conference.

References