# APPENDIX
## ADMIN3 – PANEL OF LINKED ADMINISTRATIVE DATA
**ANNA SEBŐK**

In the summer of 2019, The Databank of the Centre for Economic and Regional Studies (KRTK) established Admin3, the third round of the Panel of Administrative Data. Admin3 was made by linking individual-level as well as firm-level data provided by the National Insurance Fund Administration, the Hungarian State Treasury, the Educational Authority, the Ministry of Finance, and the National Tax and Customs Administration. The dataset is a 50% random sample of the Hungarian population, containing health, education, labour market and unemployment data as well as numerous characteristics of Hungarian firms. The Admin3 is an anonymized dataset covering multiple years.

Admin3 (covering data from 2003 to 2017) has been created using a data-integrational method. First, a 50% sample was drawn from the people who held a Social Security Number in 2003 using the National Insurance Fund Administration's registry which virtually covers the entire Hungarian population. The result of this process is a list consisting of the Social Security Numbers and the Employment Tax Numbers. Then these two identifiers were hashed using an algorithm created specifically for the purpose of the actual data linking process by the National Infocommunications Services Company (NISC Ltd.). In the next step, the other data providers queried the data they possess about those in the sample, including the two identifiers hashed with the same algorithm. They then handed these over to the NISC Ltd. which later merged and anonymized them. Thus the dataset does not contain any natural identifier. After merging the individuals' data and dropping the original identifiers, the linked dataset was sent to the KRTK Databank in a raw format yet not suitable for research work. The data cleaning then began as a joint effort of the fellow workers of the KRTK Databank and those researchers who have been engaged with the different registries for a long time. The anonymized dataset is solely available for scientific research via safe server connection and under controlled conditions.

The following healthcare-related information is available on an individual level for the period of 2009–2017: Social Security Number registry (gender, birth year and month, as well as information on Social Security Number validity), district code of residence, data about the term of social insurance, public health care, general practitioner care, inpatient and outpatient care, mortality, prescription take-outs, monetary provisions (sick benefit, baby-care allowance, childcare benefit, sick-leave). Out of these, the following data have been used

in this volume: district code of residence, number of visits to the GP, number of cases and expenses[1] of outpatients' care, number of days and expenses[2] of hospitalization, number of prescriptions taken out and expenses[3] – in total and by the main Anatomical Therapeutical Category (ATC) groups – as well as mortality. It is important to bear in mind that the aggregation of data does not allow accurate recovery of aggregated healthcare statistics due to a slightly different dataset and the nature of sampling.

The section on labour market consists of data about the employee, referral of public employment and labour force on an individual level. The Hungarian Labour Market utilizes data about labour market status and wage income.

Social transfer includes data about pension payments, monetary provisions, unemployment, and labour force related programs, on an individual level.

The section on education contains data about the participation in higher educational training, higher educational relationship, public educational relationship, maturity exam and the National Assessment of Basic Competencies, on an individual level.

Information about firms comes from the corporation tax declaration and NES Wage Survey records, at a firm-level although linked to employees/individuals.

More details of the Admin3 are available in *Sebők* (2019).

1 Including laboratory care, except for CT and MRI diagnostics.
2 Including active and chronic care.
3 Expenses covered by social security and by the patient are both available.

## Reference

sеbők, а. (2019): The Panel of Linked Administrative Data of KRTK Databank. Working Papers on the Labour Market (BWP) 2019/2. Institute of Economics, Centre for Economic and Regional Studies, Budapest.