

The hidden treasure: using admin data to improve policy making in the CEE

Budapest, 10 October 2017

As a legacy of centralised planned economies, post-socialist European countries share a strong tradition of extensive data collection by the state. These administrative data collected by the tax authority, pension funds and other government bodies are often quite high quality. Administrative data on individuals or firms has many advantages over survey data. It is often more accurate, especially compared to surveys asking people to remember things in the past, and it can provide large samples at a reasonable price. Also, by linking two or more databases, we can produce such amount of information on the (anonymised) individual that would be infeasible to collect in a survey.

Most countries also have highly skilled researchers and statisticians who could use these large datasets following the practice of Western Europe, where these data are regularly used to assess the impact of government interventions and also serve as a resource for the research community.

So why don't we use these data? One of the obstacles, ironically, lies in the laws on personal data protection endorsed when these countries reestablished democratic institutions in the early 1990s. Though researchers work with anonymised data, even that can be effectively blocked by strict data protection rules. Another common barrier is the lack of trust between academic and government organisations as well as within the government. Lastly, the relatively low efficiency of governance can play a role, as it implies that governments rarely commission proper (ie. scientifically reliable) impact evaluations on their policies.

There are some promising new developments however. Some countries are beginning to respond to internal challenges or the EU directives promoting the secondary use of public sector information ([Directive 2003/98/EC](#) and its revision in [Directive 2013/37/EU](#)). A new source of motivation has come from DG Employment of the EU Commission, who have recently introduced a rule that bidders to European Structural Funds should commit to producing at least one evaluation based on microdata and using counterfactual methods. The [Competence Centre on Microeconomic Evaluation](#) of the European Commission also actively promotes the use of admin data by governments of EU countries.

Open data initiatives have accelerated in Croatia, the Czech Republic, Slovakia and Slovenia. Poland has established a Ministry of Digitisation that promotes the secondary use of admin data and they also regularly use social insurance data to follow the labour market career of university graduates. Hungary has had an edge so far by adopting Act 101 in 2007, which eliminated the main barrier in the preceding legislation by establishing a legal basis for data owners to process personal data for the purposes of anonymisation.

Researchers can contribute to these positive developments by showing the value of research based on administrative data and sharing it with policy makers and other stake-holders, presented in an eye-catching and easy to understand way. The research community should also monitor the adoption of the new EU personal data protection regulation in the Member States to ensure it does not lead to further limitations in accessing data for the legitimate purposes of research and policy evaluation.

These lessons emerged from discussions at a workshop held in Budapest on 21-22 September. Participants came from the Visegrad countries, Albania, Bosnia-Herzegovina, Croatia, Romania, Serbia and Slovenia. The workshop was hosted by the Hungarian Society of Economists, supported by the Visegrad Fund and organised in cooperation with the Budapest Institute, CELSI, CERGE-EI and IBS. Detailed results will be published in a policy brief and follow-up events are planned in Warsaw on 29 November and in Prague on 13th December.