

Regional Update: Self-Reflections from a Project that Links Education Data from Various Sources in Ontario, Canada

Karen Robson^{a,*}

^a*McMaster University, Canada*

*Corresponding author: Email: klrobson@mcmaster.ca, Address: McMaster University, Ontario, Canada

I have had the privilege in my career of working with many longitudinal data sets that allowed me to follow the educational trajectories of young people from youth and beyond. That privilege is largely due to working and studying in the UK, where longitudinal and life course data have been collected on birth cohorts since 1946. It is certainly the case that researchers in the UK have a vast array of data from which to choose (at least 6 cohort studies, several other types of longitudinal household panels) which are quite reasonably housed and archived at the UK Data Service and are often downloadable upon agreeing to the terms of the archive online. Virtually any research question you have around youth, education and school to work transitions can be addressed using one or more of these data sources. This was the “normal” I was exposed to when working on my PhD in the early 2000s.

After repatriating to Canada to take up my first tenure track job at York University in 2004, I slowly began to realize that my “data situation” in the UK was special. As most newly minted PhDs do, I continued to work with my old UK data sets to publish out of my dissertation. But years passed and it only became logical to examine research questions in my own region. At the time, Canada had two longitudinal cohorts collected by Statistics Canada – the *National Longitudinal Study of Children (NLSCY)* and *Youth and the Youth in Transition Study (YITS)*. The former ran from 1994 to 2009 and the latter from 1998 to 2009. Yes, it has been nearly a decade since the federal government in Canada has collected longitudinal data on Canadian youth. The defunding of data was inextricably linked to the decade of cuts to various scientific agencies undertaken by the Conservative government that held office between 2006 and 2016.

Data protection laws in Canada require that researchers can only access Statistics Canada data in highly secure Statistics Canada Research Data Centers (RDCs), which are similar to other government data set-ups in the United States. A researcher must fill out a detailed application to access the data and undertake all analyses in the RDC. Results are vetted by the RDC staff. I attempted to work this way with the NLSCY data in the mid 2000s, as I was involved in a study examining the youth from military families compared to those from civilian families and we used many instruments from the NLSCY. Having to go through that process and work in the RDC was not ideal – I was still yearning for the data panacea of the UK.

Around 2010 I was asked to help with data analysis on a project with the local school district. It has been through my ongoing partnership with the Toronto District School Board that I have been involved in researching the determinants of academic success as well as the transitions to postsecondary education in Ontario. The partnership with the district has allowed me to undertake comparative analyses with partners in other cities like New York, Chicago, and London. What is visibly absent from this list of comparative cities is *Canadian cities*. While it is certainly a privilege to examine postsecondary trajectories of at-risk youth between major cities in the world, it seems only logical that comparisons should also take place closer to home.

Unlike the vast majority of countries, Canada does not have a national education system. Responsibility for education is delegated to provinces, and it has been this way since the beginning of our history as a nation. As such, although education is broadly comparable across the country, its administration and policy development has been cultivated regionally. Toronto schools are in no way “linked” to schools in Vancouver or Montreal.

Indeed, these major cities exist in different provinces with different approaches to data collection. Additionally, although all schools in the country collect administrative data on their students, with the exception of British Columbia, *there is no linkage of their school records between secondary and postsecondary*. This makes it infinitely frustrating for education researchers to be able to say anything about postsecondary transitions. In the case of Toronto, we are able to link district records to a central application data centre for colleges/universities in Ontario. We can know if student applied to postsecondary and whether or not they were given an offer of admission. What we cannot know is if they showed up and if they stayed at the institution in question. Obviously being offered a place and graduating a degree are very different things. And notably, we don't know anything about any other students living in Ontario outside of Toronto and who do not attend the public school board (there is a large publicly funded Catholic board in Ontario as well, servicing over 90,000 students a year in Toronto alone).

Trying to link the data from the public school system and postsecondary institutions in Ontario is a project that is currently being undertaken by myself and several other researchers. We don't have the National Student Clearinghouse Research data linkages that my American colleagues enjoy. In fact, data infrastructure is only being given attention lately because of the push that many researchers are giving to evidence-based policy around postsecondary access. Increasingly, the script for adulthood requires postsecondary education, even for the most basic entry-level jobs. As such, the issue of access: - i.e. who gets in, who does not – it being regarded as an important policy topic, and rightly so, as postsecondary education is increasingly being seen as a required ticket to gain entry into the labour market. In Ontario, the Toronto district has published numerous reports on racial inequities in special education, academic success, and in the streaming process in the secondary education system. Such relationships undoubtedly are exacerbated in the postsecondary sector; however the absence of data makes the research question impossible to study.

Partnerships between the Toronto district and individual postsecondary institutions are slowly developing. I am personally involved in an initiative that seeks to link data from my local district (Hamilton-Wentworth Public and Catholic) to data from my university (McMaster) and a large community college (Mohawk). This initiative is being driven by the Higher Educational Quality Council of Ontario, an agency of the government of Ontario responsible for evidence-based research on the postsecondary education in Ontario. This is, however, very new territory for Ontario bureaucrats, so it is not a quick process. Safeguards must be in place to protect data confidentiality at all partner institutions and trust relationships must be developed between the parties involved. In our Hamilton-based pilot, we are hoping to have some very preliminary results by the end of 2018. It is these kinds of district-based pilots that we hope we eventually be “scaled up” to include all districts and postsecondary institutions in Ontario. In the absence of federally-based data collection on students, this is the best we can hope for.