Kyle L Marquardt
NRU HSE
14.11.20

Identity as a latent variable

The mismatch between the methods by which scholars measure identity and contemporary constructivist theory is well-documented (Abdelal et al, 2009; Lee 2009). Individuals can have multiple identities even within the same domain (e.g. individuals can identify with both the Russian and Moldovan ethnic groups), and the salience of these identities can vary across contexts and time (e.g. an individual's civic identity may become more salient than their ethnic identity when they are in a foreign country). However, it remains all-to-common to measure an individual's identity using dichotomous indicators that reify these dynamic and continuous processes.

Latent variable models provide a potential way to bridge the divide between constructivist theory and measurement. In addition to allowing researchers to aggregate over multiple measures of identity to provide fine-grained data on this concept, more complicated modeling strategies allow scholars explicitly model the relationship between different forms of identity. However, the degree to which the increased theoretical rigor of these models is worth the trade-off in increased complexity is unclear.

In this paper, I provide insight into the advantages and disadvantages of modeling identity as a latent variable. Specifically, I develop different latent variable models of identity that build on insights from both cutting-edge political science latent variable research (Pemstein et al, 2019; Luhrmann et al 2020) and the more traditional application of item response theory models in individual-level educational testing. I apply these models to original survey data from the Moldovan separatist region of Pridnestrovie, estimating individual-level identification with different ethnic groups, languages and regions. I pursue a variety of different model specifications, illustrating the advantages and disadvantages of building more complex models that closely match theory. I conclude by comparing the estimates from these different latent variable models to their traditional dichotomous counterparts, examining the construct validity of both in empirical tests of their relationship to support for Pridnestrovien separatism.