Hungarian Historical Social Mobility file

Description of the sampling method

Population: marriages in present-day Hungary (were established in 1920) between 1850 and 1950

Sampling method: two-stage sampling, first stage stratified cluster sample of municipalities, second stage random sample from marriage records

Sampling frame: first stage: list of municipalities from the 1930 Hungarian Census; second stage: municipal parish marriage registers

Detailed description of sampling method:

Our predominantly concern by designing the sampling method was that the legal statuses of settlements (e.g., has town rights) in this historical period of Hungary do not necessarily reflect the level of their development or extent of urbanization. In a study on Hungarian settlement structures in 1910, Beluszky (2001) showed that approximately 300 settlements had some urban functions (more than twice the number of the officially acknowledged towns), but some official municipalities lacked any urban character (Beluszky, 2001). We therefore stratified Hungarian municipalities both by the legal status of the settlement (villages, towns, and regional centers) and by the level of its development. To obtain the latter, we used demographic and development indicators from the 1930 Hungarian Census and performed cluster analyses. The 1930 Hungarian Census contains information on demographics, labor, and housing conditions for the Hungarian population. The figures were aggregated to the municipal level and supplemented with information on economic establishments from the Hungarian Central Statistical Office. Full information for all relevant statistics was available for 3,417 municipalities. Published volumes of the Census, containing data aggregated to the municipal level, were used (Census 1930, 1935). The following settlement clusters were identified: *rural villages, developing rural villages, urban-type villages, agrarian towns, industrializing towns, developed urban towns, and regional centers with municipal rights*.

The second concern with the design of the sample was that Hungarian settlements were overwhelmingly agrarian; rural villages made up more than two-thirds of all municipalities. Non-

agrarian municipalities, where less than half of the population works in agriculture, composed only 5.4 percent of all settlements. The distribution of the population was, however, more even across agrarian and non-agrarian settlements: 37 percent of the Hungarian population lived in non-agrarian settlements, whereas 63 percent lived in agrarian-type municipalities. To represent all types of municipalities, we used a two-stage stratified cluster sample design. We used the 1930 Census as a sampling frame. Sampling within the randomly selected, larger regions of Hungary was performed by first randomly selecting a maximum of two towns from each developmental cluster with municipalities present in the region. For three of the randomly selected regions, we also sampled one regional center with municipal rights and included two districts from the capital city of Budapest. Next, for each town or regional center, we randomly selected one or two villages in the micro-region of the town, again, one or two from each developmental cluster. Although the sample must be weighted to represent the country, this method allowed us to include each region and type of municipal development in the sample.

For each municipality, we proceeded by digitizing the marriage acts from the church books of all local religious congregations, including Roman Catholic, Hungarian Reformed, Lutheran, and Jewish. For the South Hungarian town of Kalocsa and two surrounding villages, church books had already been digitized and put at our disposal. Our estimates for the data collection, e.g., the expected number of marriages per period, were based upon the inspection of these data. The marriage records were registered by the registrar or priest who, in some cases, did not document the occupation of the father, groom, or both. Church books did not always document occupational information because of different customs. It was necessary to pre-select sampled towns to circumvent towns with very few marriages or little occupational information documented in their church books. Towns were pre-selected by first counting the number of church-marriage records every five years and the number of marriages that did not contain occupational information for the father, the groom, or both. Based on these counts, we decided to either proceed with data collection for the town or select another town. The decision rule was that if valid observations for the most popular denomination were absent over a period of 30 years or the number of valid observations made up less than 30 percent of all marriages within that denomination, we dropped the town from the sample and sampled another town from the same region and developmental type. If a town was sampled, we repeated the same procedure as below for each of the sampled villages in the micro-region, dropping those with sparse marriage

records and randomly selecting a replacement village from the micro-region with a similar developmental profile.

For larger cities, systematic random sampling of marriages was necessary because there were too many marriages. Random sampling was achieved by assigning a sampling interval for each year, denomination, and municipality. The starting point for a given page of a church book was selected randomly.

References:

Beluszky, P. (2001). The spatial differences of modernization in Hungary at the beginning of the 20th century. *Center for Regional Studies of Hungarian Academy of Sciences Discussion Papers, No. 37.*

Census 1930. (1935). *Az 1930. évi népszámlálás, 1. és 2. kötet* [The 1930 Hungarian census, 1st and 2nd volume]. Budapest: Magyar Kir. Központi Statisztikai Hivatal.