

## Appendix

### **Variable Definitions**

**Population Density:** Natural logarithm of district population per square kilometer. It is computed using Ottoman and Turkish Censuses and the surface area of each district. The 1893 Ottoman Census is the first census to record men and women according to their millet (religious and ethnic affiliation). Figures are reported for each Ottoman kaza (district). Ottoman kaza (district) borders were manually drawn using ArcGIS. Surface area is calculated using these geocoded polygons. Missing information is filled using provincial yearbooks from the same period. The 1893 Ottoman Census has been transcribed into Latin characters by Karpat (1985). Modern Turkish Census results can be accessed through TurkStat's web application.

**Number of Missionary Stations:** Total number of Protestant missionary stations in a given kaza. The station data is sourced from primary source material from the Archives of the American Board of Commissioners for Foreign Missions (ABCFM) at Houghton Library, Harvard University. Secondary sources, annual station reports, journals, and letters are consulted to find the exact location of missionary stations.

**Number of Missionary Schools:** Total number of male and female Protestant missionary schools in a given kaza. The schooling data is sourced from Alan (2015). Figures for number of schools per kaza are calculated through spatial computations in ArcGIS.

**Average Elevation:** The average elevation of a given district. The primary source is the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Digital Elevation Model Version 3 (GDEM 003). The dataset is jointly created by the Ministry of Economy, Trade, and Industry (METI) of Japan and the United States National Aeronautics and Space Administration (NASA). The ASTER GDEM's coverage spans from 83 degrees north latitude to 83 degrees south, encompassing 99 percent of Earth's landmass, with 30-meter posting and 1x1 degree tiles. The raster image for elevation is used to calculate average figures for each kaza using ArcGIS software.

**Annual Mean Temperature:** The average annual temperature of a given district. The primary source is the CHELSA (Climatologies at High Resolution for the Earth's Surface Areas) dataset. The CHELSA dataset includes monthly mean temperature and precipitation patterns for the time period 1979-2013. The raster image for regional temperature is used to calculate average figures for each kaza using ArcGIS software.

**Annual Mean Precipitation:** The average annual precipitation of a given district. The primary source is the CHELSA (Climatologies at High Resolution for the Earth's Surface Areas) dataset. The CHELSA dataset includes monthly mean temperature and precipitation patterns for the time period 1979-2013. The raster image for regional precipitation is used to calculate average figures for each kaza using ArcGIS software.

**Distance to Nearest Lake:** The logarithm of the minimum distance from the centroid of a district (in kilometres) to the nearest major lake. The shape file is downloaded from HydroLakes database (<http://wp.geog.mcgill.ca/hydrolab/hydrolakes/>) in vector format. The primary source is the HydroLakes database. The spatial computations are made using ArcGIS software.

**Distance to Nearest River:** The logarithm of the minimum distance from the centroid of a district (in kilometres) to the nearest major river. The shape file is downloaded from HydroLakes database (<https://hydrosheds.org/page/hydrorivers>) in vector format. The spatial computations are made using ArcGIS software.

**Distance to Nearest Sea:** The logarithm of the minimum distance from the centroid of a district (in kilometres) to the nearest major sea. The shape file is downloaded from World Ocean Basemap ([https://services.arcgisonline.com/ArcGIS/rest/services/Ocean/World\\_Ocean\\_Base/MapServer](https://services.arcgisonline.com/ArcGIS/rest/services/Ocean/World_Ocean_Base/MapServer)). The spatial computations are made using ArcGIS software.

**Distance to Nearest Major Port:** The logarithm of distance of a district (in meters) to the nearest major port. The spatial computations are made using ArcGIS software. The source of this data comes from the 1899 R. Huber Map. The major ports are Constantinople (Istanbul), Dardanelles (Çanakkale), Gallipoli (Gelibolu), Smyrna (Izmir), Silifke, Mersin, Alexandrette (Iskenderun), Erekli (Karadeniz Ereğli), Inebolu, Sinop, Samsun, Ordu, Dunie, Kiressoun (Giresun), Tireboli, Trebizonde (Trabzon), and Rize.

**Distance to Ottoman Trade Routes:** The logarithm of distance of a district (in kilometers) to the nearest Ottoman Trade route. The spatial computations are made using ArcGIS software. The source of this data is the Old World Trade Routes (OWTRAD) Project (<http://www.ciolek.com/owtrad.html>).

**Distance to Carriageway:** The logarithm of distance of a district (in kilometres) to the nearest 19<sup>th</sup> century carriageway (chaussées). The spatial computations are made using ArcGIS software. The source of the data is Huber's Ottoman Map of 1899. Roads are manually traced using ArcGIS software.

**Carriageway Density:** The magnitude-per-unit area of 19<sup>th</sup> century carriageways that fall within given kaza. The spatial computations are made using the line density tool in ArcGIS software. The source of the data is Huber's Ottoman Map of 1899. Roads are manually traced using ArcGIS software.

**Distance to the Capital Constantinople (Istanbul):** The logarithm of distance of a district (in kilometers) to Constantinople (Istanbul) which was the capital of the Ottoman Empire (1299-1922). The spatial computations are made using ArcGIS software.

**Central Kaza Dummy:** Indicates whether the kaza in question was the administrative capital of a sanjak (sub-division) according to the 1893 Ottoman Census. The primary source I consulted for this information was the historical gazetteer of the Ottoman Empire compiled by Tahir Sezen (2017).

**Ibtidâî Enrolment Figures (1914):** The primary source for these figures are the Statistical Review of the Ministry of Education (Maarif-i Umumiyye Ihsaiyat Mecmuasi) for the Muslim calendar year of 1336. This report enumerates the number of schools, teachers, students, and those who have reached a compulsory schooling age. Using this information, it is possible to calculate relatively accurate enrolment rates for the period in question.

**Ibtidâî Gender Parity Index (1914):** The ratio of the absolute number female students to male students for the Ibtidâî schools for the school year starting in 1914. The primary source for these figures are the Statistical Review of the Ministry of Education (Maarif-i Umumiyye Ihsaiyat Mecmuasi) for the Muslim calendar year of 1336.

**Educational Attainment Figures (2010):** The share of the population above the age of 15 that has received an elementary or high school degree. The education data is obtained from the Turkish Statistical Institute (TÜİK) for the year 2010.