APPENDIX B. THE 1998 ITALIAN HOUSEHOLD INCOME SURVEY

Data set description and use

The civilian earnings tables presented in Chapter 2 are based on data collected by the Bank of Italy’s survey of household income and wealth in 1998, which was conducted between February and July 1999. The stratified sample included 7,147 households, and was drawn in two stages (municipalities and households).\textsuperscript{150} The summary data presented in this document appears in terms of percentiles of the salary distribution, as opposed to averages. Percentiles are more robust summary statistics, and given the limited sample size and high standard errors associated with the sub-samples used in the analysis, they are likely to be more reliable. A look at percentiles of the income distribution also allows the reader to get a sense for where on such a distribution military salaries are located. The percentile statistics were weighted using a set of analytical weights provided in the data set in order to adjust for the different sampling probabilities of households across Italy. Given the sample frame, the weight coefficient varies according to city size as follows:

\[
W_{hi} = \begin{cases} 
\frac{P_h P_{nh}}{P_{nh} n_{hi}} & \text{for cities above 40,000 inhabitants,} \\
\frac{1}{m_{h} n_{hi}} & \text{for cities with up to 40,000 inhabitants.}
\end{cases}
\]  
\text{(B.1)}

where the first expression applies to cities above 40,000 inhabitants, while the second is used for cities with up to 40,000 inhabitants. \(P_h\) is the resident population, \(P_{nh}\) is the population of the cities included in the survey, and \(m_{h}\) is that of the cities in the \(h^{th}\)

\textsuperscript{150} The primary sampling units (municipalities) were stratified by region and size. Within each stratum, the municipalities in which interviews would be conducted was selected by including all municipalities with a population of more than 40,000 and randomly selecting smaller towns. Households were then selected randomly.
stratum. \( P_{hi} \) represents the population of interviewees and \( n_{hi} \) is the number of interviewees in the \( i^{th} \) city of the \( h^{th} \) stratum (Bank of Italy, 2000).

**Use**

The survey differentiates between several sources of income—including employment income, self-employment income, capital gains, and government transfers. Only employment income was used for comparing civilian to military salaries. To exclude part-time workers, observations with employment income of less than Lit.5m were taken out of the sample. And since salary data dated back to 1998, an average nominal wage estimate was used to make the data current to the beginning of 2000.\(^{151}\)

**Limitations**

The data set and approach outlined above suffer from a series of limitations, the most important of which relates to sample size for some of the subgroups considered in the analysis. For instance, the data on the salary distribution of the 17-22-year-old cohort is composed of a mere 180 observations. The standard errors are high, and the results should be interpreted with caution. At the same time, however, the data is used for rough comparison purposes only, and as previously mentioned including statistics on median and other deciles adds robustness to the findings.

The fact that data from 1998 is compared with military salaries that were current in 2000 adds a further complication. Ideally, since real and nominal wage growth varies significantly between sectors, 1998 salaries could be converted in 2000 values using the appropriate “inflator.” In the absence of such statistics, the average nominal wage growth was used as a rough approximation. Again, given the rather modest data analysis objectives anticipated for this data set, this “short-cut” was deemed acceptable. Lastly, the Lit.5m cutoff to differentiate between full- and part-time workers is arbitrary, but

\(^{151}\) The “inflator” was obtained from the Economist Intelligence Unit’s Italy Country Profile for 2000. See EIU (2000).
reasonable. Moreover, including the relatively few observations below the Lit.5m floor does not substantially alter the income distribution by deciles, as this summary statistic is robust to outliers.