GABRIELLA DONATIELLO

ITALIAN NATIONAL INSTITUTE OF STATISTICS - ISTAT. ITALY

MARCELLO D'ORAZIO

ISTAT, ITALY

DORIANA FRATTAROLA

ISTAT, ITALY

ANTONY RIZZI

ISTAT, ITALY

MAURO SCANU

ISTAT, ITALY

MATTIA SPAZIANI

ISTAT, ITALY

STATISTICAL MATCHING OF INCOME AND CONSUMPTION EXPENDITURES

Abstract:

In recent years, there has been increasing interest in using appropriate instruments to measure household living conditions. Actually defining material living condition needs to consider the level of consumption as well as the economic resources in terms of income and wealth that enable household consumption of goods and services. Collecting information on the joint distribution of income, consumption and wealth at the micro level poses several difficulties for National Statistical Institutes. In particular, setting up a new survey is unfeasible because of budget constraints as well as a significant reporting burden on respondents given the high amount of data to be collected in a single survey. As a result a better exploitation of existing data sources becomes of vital importance and statistical matching techniques could represent a valid alternative for producing statistics on the distribution of variables not jointly collected in a single survey. However several critical issues need to be taken into account for assessing the quality of the results and of the whole matching process.

The purpose of this paper is to evaluate the possibility of applying statistical matching on two different data sources to create an integrated database with detailed information on households income and consumption expenditures in Italy. The data to integrate are those of EU-SILC (European Union Statistics on Income and Living Condition) 2012, with income reference year 2011, and the HBS (Household Budget Survey) 2011. Both surveys are conducted by ISTAT. This paper explores which are the matching approaches more suitable with the final objective and provides insights concerning some important steps of the integration process. It is worth noting that in our case it is not possible to perform statistical matching under the conditional independence assumption (CIA, independence between income and consumption given some common information in both the data sources). To avoid the CIA it is evaluated the usage of the available auxiliary information (e.g. household monthly income, housing costs). In alternative, the statistical matching approach based on the exploration of the uncertainty due to the absence of joint information on households expenditures, income and wealth is considered. In order to improve the quality of the matching procedure the advantage in having a more efficient ex-ante data

collection system as well as a better harmonization of common variables of SILC and HBS and other important social surveys is discussed. The main results related to the integrated data set are finally presented.

Keywords:

Statistical matching, Survey data integration, Income, Consumption

JEL Classification: C15, C14