

Combining expenditure surveys and tax registers: the case of Milan municipality.

Lisa Crosato, Paolo Mariani, Mauro Mussini and Biancamaria Zavanella¹

Preliminary version, please do not cite.

Abstract

In this paper we combine two data sources: the sample survey on family expenditure conducted by the Milan Municipality and the Chamber of Commerce of Milan (wave 2007-2008) and the tax register matched to the local population and family register in the data-warehouse AMeRiCA, regarding residents in Milan in 2007.² We mainly focus on data combination and run a few matching exercises following Jenkins et al (2008) approach. Using both information on individuals and households we select four variables (Census Section, Type of Family and two compounded variables for age and sex) and combine them into four matching criteria applied in a first stage independently and in a second one hierarchically. Results are encouraging since we match more than 70% of the survey households through the automated matching exercise.

Keywords: Exact matching, administrative data, income, tax records.

¹ Corresponding author: Lisa Crosato, Statistics Department Università di Milano-Bicocca, Via Bicocca degli Arcimboldi, 8, 20126 Milano. email: lisa.crosato@unimib.it.

² We thank Milan Municipality (Statistics Office) and the Chamber of Commerce of Milan for providing us with the data. We are particularly grateful to Lorena Scarcello and Flavio Necchi for their support to our project.