The European Community Household Panel: A review

Franco Peracchi*
Tor Vergata University, Faculty of Economics 00133 Rome, Italy
(e-mail: franco.peracchi@uniroma2.it)

First version received: July 2000/Final version received: October 2000

Abstract. The European Community Household Panel (ECHP) is a standardized multi-purpose annual longitudinal survey carried out at the level of the European Union. It is centrally designed and coordinated by the Statistical Office of the European Communities (Eurostat), and covers demographics, labor force behavior, income, health, education and training, housing, migration, etc.

After reviewing the organization of the survey, this paper discusses some of the issues that an applied economist potentially interested in the ECHP may face when trying to deal with these data. In particular, the paper focuses on survey attrition and nonresponse, the weighting procedures with which Eurostat tries to compensate for differential nonresponse and attrition rates, and the imputation methods adopted by Eurostat in order to deal with item nonresponse.

Key words: Panel data, nonresponse, attrition, Europanel

JEL classification: C4, C8

1. Introduction

The European Community Household Panel (ECHP) is a standardized multi-purpose annual longitudinal survey carried out at the level of the European Union (EU). It is centrally designed and coordinated by the Statistical Office of the European Communities (Eurostat), and covers demographics, labor force behavior, income, health, education and training, housing, migration, etc.

* I thank an Associated Editor and seminar participants at CEMFI, ISTAT, New York University, Texas A&M University and the University of Turin for helpful comments.
The project was launched in 1992 “in response to the increasing demand in the European Union for comparable information across the Member States on income, work and employment, poverty and social exclusion, housing, health, and many other diverse social indicators concerning living conditions of private households and persons” (Eurostat 1996b).

The objective of the ECHP is to represent the population of the EU at the level of households and individuals. It aims at being both cross-sectionally and longitudinally representative, with changes in the population over time reflected by the continuous evolution of the sample – through births to sample households and the formation of new households from the split off of existing ones. This aspect differentiates the ECHP from other panels, such as the National Longitudinal Surveys in the USA, that only follow a few cohorts of individuals as they age. In this, and many other dimensions, the ECHP is patterned after the Panel Study of Income Dynamics (PSID) conducted in the USA by the Survey Research Center at the University of Michigan.1

The ECHP, initially conceived as a 3-wave panel, has now reached its 7th wave. The first wave, conducted in 1994, included all current members of the EU except Austria, Finland and Sweden. Austria was added in the second wave conducted in 1995, Finland in the third wave conducted in 1996, and Sweden in the fourth wave conducted in 1997. Funding for 3 additional waves was approved by the European Commission in 1996 and then again in 1999. In its current form, the survey is expected to end in year 2002. At that point, it will be redesigned thoroughly or simply discontinued.

The attractive features of the ECHP for socio-economic research are its comparability across countries and over time, and the range of economic and socio-demographic information that it collects. The purpose of this paper is to discuss some of the issues that an applied economist interested in exploiting the opportunities offered by the ECHP may face when trying to deal with these data. Our discussion is based on the documents issued by Eurostat and the evidence from the publicly released data, which currently contain the first three waves of the panel.

The remainder of this paper is organized as follows. Sections 2 and 3 describe the organization of the survey (coverage, sampling design, following rules, questionnaire, fieldwork, and data editing), and data access and organization. The substantive problems of attrition and nonresponse are addressed in Sections 4–6. Specifically, Section 4 discusses the issue of nonresponse, distinguishing between unit nonresponse, item nonresponse, sample attrition, and new entry. Section 5 discusses the weighting procedures with which Eurostat tries to compensate for differential nonresponse and attrition rates. Section 6 discusses the imputation methods adopted by Eurostat in order to deal with item nonresponse. Finally, Section 7 offers some conclusions.

---

1 The PSID is a widely used longitudinal survey of the U.S. population. The study, designed to provide a continuous representation of individuals and the families they live in, contains detailed information on socio-economic and demographic variables. It has been ongoing since 1968. Its initial sample comprised two different components: an approximately random sample of families (the “SRC sample”) and a sample of low-income families (the “SEO sample”). Data are collected annually and the publicly released micro-data files, now downloadable free of charge from the Internet (http://www.isr.umich.edu/src/psid/), contain the full span of the information collected over the course of the study. For an introduction to the PSID, see Hill (1992).