Journal of Economic Psychology 30 (2009) 622–633

Contents lists available at ScienceDirect

Journal of Economic Psychology

journal homepage: www.elsevier.com/locate/joep

Analyzing consumers heterogeneity and self-reported tastes: An approach consistent with the consumer's decision making process ☆

Victor Fernandez-Blanco, Luis Orea, Juan Prieto-Rodriguez *

Department of Economics, University of Oviedo, Avda. del Cristo s/n, 33006 Oviedo, Spain

ARTICLE INFO

Article history:
Received 14 July 2006
Received in revised form 7 October 2008
Accepted 27 April 2009
Available online 18 May 2009

PsycINFO classification:
3920

JEL classification:
D12
Z10

Keywords:
Preferences
Consumer behaviour
Latent class models
Movie demand

ABSTRACT

This paper analyses whether self-reported valuation of goods is an adequate proxy for underlying tastes. In this case, different self-reported appraisals, which we can associate with different utility functions, would imply different demand curves. To estimate these kinds of relationships we have used data on reported tastes on new film releases and cinema attendance. We have used a latent class approach in order to imitate the data generating process underlying the demand functions, where consumer's preferences are determined before consumers purchase. With this procedure we can reject the hypothesis of a unique demand function for all consumers. As expected, moreover, prices have a different influence depending on individuals' self-reported tastes and specific market policies for each consumer group could therefore be designed.

1. Introduction

A critical problem in estimating demand or behaviour functions in social science is the presence of unobserved individual heterogeneity. Since the shape of demand functions relies on individual preferences, which are not observed, estimating a common demand function encompassing every sample observation may not be appropriate in the sense that the estimated function is not likely to represent the 'true' preferences or consumer's behaviour. Therefore, we need to control for unobserved individual heterogeneity in order to avoid biases in the estimate of the underlying preferences. A common approach has been to resort to some demographic and personal variables which are often reported in surveys such as sex, age or education. Although these variables can capture differences in individual's tastes and be used to identify different consumer segments, they might be weak proxies for the underlying preferences due to the fact that they are mainly collected for purposes

☆ Research for this paper was financed by the Spanish Ministry of Science and Technology Projects #ECO2008-04659 and SEJ2007-64700/ECON. The database was served by Foundación Autor and Sociedad General de Autores y Editores (SGAE). We would like to thank the annonymous referee and the Editor for their helpful suggestions. This research has also benefited from helpful comments by William Greene, Subal Kumbhakar and Mike Tsionas. The usual disclaimer applies.

* Corresponding author. Tel.: +34 985103768; fax: +34 985104871.
E-mail addresses: vfernan@uniovi.es (V. Fernandez-Blanco), lorea@uniovi.es (L. Orea), jprietor@uniovi.es (J. Prieto-Rodriguez).

0167-4870/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.joep.2009.04.005

other than approximating preferences. Moreover, introducing these elements into the demand functions as control variables only accounts for intercept heterogeneity, with the slope remaining unchanged.

To obtain more robust estimations than those using only observable characteristic variables, in this paper we advocate extending the set of separating variables by including self-reported valuations and the use of latent class models, which tried to imitate the data generating process underlying demand functions, i.e. it allows us to identify consumer groups with similar preferences and then to investigate differences in demand given consumer's preferences.

Demand functions are linked to utility functions, which are determined by different tastes. Since it is reasonable to think that self-reported valuations are highly correlated with unobservable tastes (otherwise it would imply that people were lying), they should be good proxies for the underlying utility preferences and supplement the information coming from observable characteristic variables[1]. Under the assumption that self-reported valuations are "close" enough to the underlying preferences, access to this subjective, though invaluable, information allows us to avoid the biases which are present in the estimates of underlying demand functions which use only observable characteristic variables, as is customary in empirical social science research[2]. If these biases are large enough, self-reported valuations (if available) should be included to control for individual heterogeneity. If this information is not available, consumer surveys should be extended in order to include such information.

From a methodological point of view, we advocate using latent class models to control for intercept and slope individual heterogeneity using observable characteristic variables and self-reported data.[3] Unlike other techniques such as cluster analysis which permit the identification of different groups in two stages (i.e. observations are first divided in several classes and then separate analyses are carried out for each subsample), latent class analysis is a one-stage technique. In this model we do not need prior knowledge as to which group an observation belongs since both the consumer's demand and the probability of her being a member of a particular group are estimated *simultaneously*. That is, in a mixture model the sample is segmented *endogenously* into different classes or groups, which differ in terms of the parameters of the estimated demand functions. The information contained in the data is exploited more efficiently in a latent class model and the estimated demand functions are therefore more robust because the model exploits not only sample data (which may include self-reported valuations) but also the goodness of fit of each estimated demand to provide additional information with which to identify groups of individuals.

We use data from a Spanish survey on culture conducted during 1998 to illustrate how self-reported valuations might play an important role in controlling for individual heterogeneity when demand functions are estimated. We have chosen this data for several reasons. First, this survey contains information about subjective valuations of both American and Spanish films. Second, the motion picture industry is receiving increasing attention in the economics literature, especially with regard to its industrial organization (Canterbery & Marvasti, 2001; Chisholm, 1997; De Vany, 2004; Moul, 2005) and the determinants of the economic success of a movie (De Vany & Walls, 1996; De Vany & Walls, 1997; De Vany & Walls, 1999; Prag and Cassavant, 1994; Ravid, 1999; Walls, 2005). Most of these papers are Hollywood-oriented, while in Europe the relatively little research that has been carried out has been devoted mainly to the estimation of demand functions (Cameron, 1990; Cameron, 1999; Dewenter & Westermann, 2005; Fernández-Blanco & Baños-Pino, 1997). Third, the film industry is one of the largest cultural sectors in economic terms[4]. Fourth and finally, in the US core copyright industries are in first place in terms of exports ranking and, after computer services, movies and video exports are the main categories of copyright exports. It is understandable therefore why Hollywood is increasing its interest in foreign markets (Marvasti & Canterbery, 2005). The European Union (EU) is the number one export market for American films and Spain is a relevant market inside EU.

The paper has the following structure. Section 1 describes the variables contained in the Spanish survey, paying special attention to the definition and structure of the self-reported valuation variables. Section 2 presents the empirical model to be estimated. The results are displayed in Section 3. Finally, Section 4 offers the main conclusions of our research.

2. Preferences and personal valuations: the case of Spanish and American films

In this paper, we use a data set coming from the Spanish cultural consumption habits survey (*Encuesta sobre Hábitos de Consumo Cultural – CCHS*) established by the Spanish Authors' Copyright Foundation (see Fundación Autor, 2000) and which was conducted in Spain during 1998. In each quarter of that year a new random sample of 3018 people over fourteen years of age were interviewed. The total sample size was thus 12,072 people[5]. The CCHS is an opinion survey covering the most important fields of cultural consumption: performing arts, cultural industries – including movie theatre attendance- and other leisure

[1] Clark, Etilé, Postel-Vinay, Senik, and Van der Straeten (2005) establish that the connection between self-reported valuations and tastes "... requires (i) associating utility to observable characteristics (gender, age, marital status, etc.) and (ii) relate discrete reported valuations to latent utility functions". Thus, we can face two types of individual heterogeneity: "heterogeneity in the utility function" and "heterogeneity in the expression function".

[2] This is so-called omitted variable bias in the sense that the estimated parameters not only capture the effect of its own variable but also the overall effect of those variables that have been excluded from the model and are correlated with the included variables.

[3] This approach is also known as the semiparametric heterogeneity model (Heckman & Singer, 1984) and also finite mixture models. See Greene (2008) for a survey of discrete choice models including latent class models.

[4] In 1997, the audiovisual sector, including TV and radio, represented 21.4% of added value and 17% of employment in the cultural sector in Spain (García Gracia & Zofío Prieto, 2003).

[5] The EHCC was also conducted quarterly during 1997 but during that year the survey was in the development stage and the interview questionnaire content was significantly different and is not suited to the type of work we carry out in this paper. Accordingly, it was decided to use only the 1998 data.