Journal of Retailing and Consumer Services 39 (2017) 35–42

Contents lists available at ScienceDirect

Journal of Retailing and Consumer Services

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

Willingness to Pay-inference in the absence of rejected propositions

Sander van Cranenburgh, Caspar G. Chorus*

Transport and Logistics Group, Delft University of Technology, The Netherlands

ARTICLE INFO

Keywords:
Willingness to Pay
Discrete choice
Shopping location

ABSTRACT

This paper presents a method to infer a Willingness to Pay (WtP) distribution based on a sample of observations of individuals who pay a particular price for a particular quality increase. Crucially, no observations are available of individuals who reject a higher price/higher quality proposition, and choose a lower price/lower quality alternative instead. While at first sight it may seem impossible to infer a WtP distribution in the absence of such rejected propositions, we show that this is possible under the assumption that there is a certain degree of alignment of supply (of propositions) and demand (WtP) in the market. The method is Maximum Likelihood-based, and easy to implement. The method is shown to have a promising empirical performance on a synthetic dataset and a dataset of revealed shopping destination choices.

1. Introduction

Inference of Willingness to Pay (WtP) is a key research aim for applied econometricians in a variety of research fields such as marketing (Wertenbroch and Skiera, 2002; Li et al., 2012a), transport (Molin and Timmermans, 2006; Abrantes and Wardman, 2011), health (de Bekker-Grob et al., 2012; Boeri et al., 2013) and environmental economics (Thiene and Scarpa, 2009; Perez-Pineda and Quintanilla-Armijo, 2013). Classical applications include WtP-inference among consumers (e.g., for quality improvements of consumer goods); travelers (e.g., for travel time reductions); patients (e.g., for more effective treatments); or citizens (e.g., for reductions in flood risk), to name just a few. Although many studies have used Stated Preference (SP) surveys to extract WtP information, it is widely acknowledged that – if available – Revealed Preference (RP) data provide the preferred empirical context for such inference, given its high level of external validity.[1] However, a well-known drawback of RP data is that it is often more 'noisy' and less complete or detailed, than data collected in SP-settings. This paper focuses on a particular type of data-imperfection which may occur in RP-settings, and it provides a new method for WtP-inference in the context of such imperfect data.

Before we discuss the type of RP data-imperfection that is the focus of this paper, it should be noted that many types of data that are being used for WtP-inference can be conceptualized or recast as a series of propositions (binary choice sets) to 'buy' an increased level of 'quality' (e.g., a larger smartphone-screen, a lower travel time, a better medicine, or a smaller flood risk) for a given price.[2] When there is sufficient variation in propositions (i.e., in levels of quality increases as well as in levels of price), decision-makers' WtP can be inferred from observed patterns of acceptance and rejection of different propositions. Acceptance of a proposition helps to determine a lower bound of WtP, while a rejected proposition helps determine an upper bound. Jointly, they enable the analyst to pinpoint mean WtP, or more generally the WtP distribution, for a given sample.

In some cases, however, only 'accepted propositions' are observed. Take for example the situation where only choices of individuals who are willing to pay a premium for a larger smartphone screen are observed. Or, the situation where only travelers on a faster but more expensive toll road are observed. Or, the situation where only visitors of a larger, but further away shopping center are observed (in this latter case, WtP takes the form of willingness to travel further). Such a failure to observe rejected propositions may be due to particularities – or even flaws – of a dataset or data collection process; or they may be the consequence of intrinsic factors. An example of the latter may be that the data are supplied by an entity (e.g., firm, road operator, or retailer) which only supplies the 'premium product' (i.e., smartphone with larger

* Correspondence to: Jaffalaan 5, 2628 BX Delft, The Netherlands.
E-mail address: c.g.chorus@tudelft.nl (C.G. Chorus).

[1] In addition to external validity issues, the *internal* validity of SP surveys is increasingly being questioned by scholars. For example, in the travel behavior literature, where SP-surveys are routinely used to infer WtP for travel time reductions (leading to so called Value of Time-estimates), recent papers have shown that the design of such surveys influences WtP-estimates in unwanted ways (Fosgerau and Börjesson, 2015; Ojeda-Cabral et al., In Press).

[2] With this statement, we do not wish to suggest that this binary choice set-approach is the dominant empirical approach to infer WtP; in fact, most empirical studies multinomial choice sets and more than two attributes. We merely wish to point out that conceptually speaking, a binary choice between a low-quality/low-price alternative and a high-quality/high-price alternative is the most simple and archetypical choice situation from which WtP for the quality attribute can be inferred.

http://dx.doi.org/10.1016/j.jretconser.2017.07.001
Received 24 August 2016; Received in revised form 21 April 2017; Accepted 3 July 2017
0969-6989/

screen, faster toll road, larger retail center) and hence has no data available concerning individuals who chose not to buy the premium product.

Conceptually as well as econometrically speaking, at first it sight appears impossible to infer WtP from these observations of accepted propositions – or: choices for the premium product – alone. One might expect to be able to derive only lower bounds of (mean) WtP from such observations. However, in this paper, we propose and test a method that can be used to infer WtP in the absence of observations concerning rejected propositions.[3]

The method is fairly straightforward conceptually and easy to implement econometrically. It is based on two assumptions, which we will argue are likely to be reasonably accurate in practical applications. Given these assumptions, the likelihood of an observed accepted proposition can be conceived as being the product of the probability that a particular proposition is drawn from a particular distribution of propositions, and the probability that an individual is drawn from the distribution of individuals, whose WtP is at least equal to the price embedded in the proposition. As such, standard Maximum Likelihood Estimation routines can be applied to infer the distribution of WtP. Empirical analyses based on RP shopping destination choice data provide a first sign of the method's potential, although we stress that replication on other datasets is needed to get insight into the validity of the approach.

The remainder of this paper is structured as follows: after presenting the proposed method and its economic and econometric rationale in more depth in the next section, Section 3 provides the empirical applications, one based on synthetic data (giving a proof of concept and a test of robustness) and one based on real data. Conclusions and directions for further research are presented in Section 4.

2. Theoretical considerations and the proposed method

2.1. Setting the stage and problem formulation

WtP is assumed to vary across the population of interest (i.e., the population of individuals targeted by suppliers of propositions); it is conceived as a random variable denoted *WtP* which follows a probability density function $f_{WtP}(WtP)$. Each individual n in the population of interest is assumed to receive a proposition to pay a price p_n for a quality increase of size q_n. Propositions, too, are assumed to be stochastic in the eyes of the analyst; that is, the probability that a randomly sampled individual from the population of interest receives a particular proposition is a priori unknown to the analyst. Variable *proposition* is described by probability density function $f_{proposition}(proposition)$.

The assumption that propositions vary across the population of consumers is non-trivial, but can be justified building on key results in academic literature on price dispersion. Starting with seminal work by Stigler (1961), price dispersion has been explained by economists as following from the fact that in many markets, the search for price and quality information is costly for consumers. The core of the argument put forward by these and other 'information economists' is that when search for price and quality information is costly, a firm may benefit from charging a higher price per unity of quality than competing firms; it will attract under-informed consumers who would have been better off buying from another firm. During the 1970s and 1980s, a plethora of studies has explained how different market structures (e.g. monopoly vs. competitive), different levels of search costs and preference heterogeneity among consumers, lead to different levels of price dispersion (e.g., Reinganum, 1979; Salop and Stiglitz, 1982; Dahlby and West, 1986). A key finding from these studies is that price dispersion is (or: should be theoretically expected to be) the rule, not the exception, in most markets. During the early days of the internet economy, the expectation gained traction that by substantially lowering consumers' search costs, the internet (and comparison-websites in particular) would reduce or even eliminate price dispersion in many markets – especially those where easily comparable goods are sold, such as books. However, more recent studies have shown that price dispersion persists even in markets in which internet is a main channel of information provision and sales (e.g., Biswas, 2004; Baye et al., 2004; Clemons et al., 2002; Roma et al., 2015; Zhao et al., 2015; Moreno-Izquierdo, 2015; Meseguer-Artola and Rodríguez-Ardura, 2015). A recent study has shown that even in a market where price dispersion is prohibited by law, subtle mechanisms arise which cause implicit prices to vary across consumers (Li et al., 2012a, 2012b). In sum, the assumption used in the remainder of this paper – i.e., that propositions (price charged per unit of quality) vary across potential buyers – can be regarded realistic in many markets.

Consider now individual n who is observed to have accepted a proposition to pay a price p_n for a quality increase of size q_n. Note that this conceptual representation of a proposition encompasses a variety of sorts. For example, take an individual that faces a choice task that involves a product characterized by price p_n^* and quality level q_n^*, and a product with price $p_n^* + p_n$ and quality $q_n^* + q_n$ (e.g. a choice between a fast but expensive toll road and a cheap but slow alternative). Here, the proposition to pay a price of p_n for a quality increase of size q_n is implicit in the choice task. Or, take a situation where the individual is explicitly confronted with a proposition to pay a price p_n for a quality increase of size q_n (e.g., an offer to upgrade an airline ticket from Economy to Business class). That is, the sample of observations available to the analyst is fully characterized by a set of accepted propositions, which is assumed to be randomly drawn from the distribution of accepted propositions in the population. This sample of accepted propositions can be denoted as a set of N price-quality increase ratios $\left\{\frac{p_1}{q_1}, \dots, \frac{p_n}{q_n}, \dots, \frac{p_N}{q_N}\right\}$, which normalizes propositions in terms of a price per unit of quality increase. Crucially, the analyst does *not* observe rejected propositions. That is, she does not observe choices for lower-quality/lower-price alternatives such as choices for the non-toll road in the first example or rejections of the Economy-Business class upgrade-proposition in the second example.

The aim is to infer, from the set of observed accepted propositions, the distribution of WtP in the population, measured in terms of the (maximum) price different individuals are willing to pay per unit of quality increase.[4] It is instructive first, to discuss the reasons why at first one would be inclined to believe that it is impossible to infer WtP from such a set of accepted propositions alone. Conceptually speaking, an observation of an accepted proposition p_n/q_n implies that the individual is willing to pay at least as much as p_n for a quality increase of size q_n. In other words, p_n/q_n only gives a lower bound estimate for the individual's WtP. No information at all regarding an upper bound for her WtP can be inferred from this data point. In the process of Maximum

[3] At first sight, it may seem that the type of choice situation we consider, may be analyzed using choice based sampling techniques (e.g., Manski and Lerman, 1977): these techniques allow the researcher to obtain unbiased parameters estimates in situations where the probability that a choice made by a particular individual is included as an observation in the dataset depends on the chosen alternative. For example, in a country where the market share of public transport is very low compared to that of the private car, a travel behavior researcher may oversample individuals that have chosen to travel by train (e.g. by distributing surveys at a railway station) in order to obtain a sufficiently diverse set data to estimate a choice model. Our considered type of data is different from such a choice based sample in one crucial aspect: we consider a binary choice set, where the probability that a choice for the low-price/low-quality alternative ends up in the data set is zero. This precludes the use of choice-based sampling techniques, which is designed for the situation where there is a non-zero probability of inclusion in the data set for each type of considered alternative (e.g Imbens, 1992; page 1189).

[4] Throughout the paper, the assumption of linearity is maintained, in the sense that if an individual is willing to pay a price p for a quality increase of size q, then she is assumed also to be willing to pay $K * p$ for a quality increase of size $K * q$. Relaxation of this assumption, which is used in the majority of WtP-studies, is left for further research.