

Webscraping prices to estimate hedonic models and extensions to other predictive methods

Mathilde Poulhes, Jean-Denis Zafar

Abstract: When an item is missing and has to be replaced, the difference of quality between the disappearing product and the new one must be taken into account in the price index, in order to measure comparable prices. Hedonic regressions can be used to estimate this difference, using product characteristics as explanatory variables for the price. However, the quality of the models can be insufficient due to the small size of samples. This paper explores the use of webscraping in order to gather bigger volumes of information on prices and characteristics, in particular for electronic goods. Traditional hedonic regressions will be compared with other predictive methods, including machine learning algorithms in terms of predictive power.

Keywords: consumer price index; hedonic regression; quality adjustment