

*17th Ottawa group meeting  
Rome, 7 - 10 June 2022*

## **Title: Applying advanced techniques to the estimation of Multipurpose Digital Device Price Indices**

**Authors:** *Kevin Roche, Lance Taylor, Roobina Keshishbanoosy*

As part of the efforts to continuously improve the quality of the Consumer Price Index (CPI) at Statistics Canada, new methodologies and data sources are being researched for measuring price change in multipurpose digital devices. The research demonstrated here integrates web scraped data with modern supervised modelling approaches to produce quality adjusted price indices. Furthermore, this experimental set of approaches also explores different aggregation structures and sets of weights for the indices, as well as increasing the monthly sample size and reducing manual price collection.

We compare the results of different algorithms in modeling prices for quality adjustment, as well as different methods for constructing the corresponding quality adjusted price indices. Different specifications of ordinary least squares, random forest, and extreme gradient boosting (xgboost) modelling algorithms are tested, with the xgboost models (widely considered the gold standard in predictive machine learning) delivering the best out-of-sample fits. Comparisons of hedonic imputation and hedonic time dummy approaches are made, including the novel use of double machine learning – a recently developed technique that allows for causal inference when using non-parametric methods (i.e. with the use of xgboost). Additionally, we compare the effects of using different weights and structures on the indices.

To our knowledge, no other national statistical office has experimented with double machine learning for the purpose of producing quality adjusted price indices, making our research a significant contribution to the field. By allowing for causal inference in an otherwise non-parametric framework, we demonstrate the increased flexibility of the application of machine learning in quality adjustment processes. The final results of this novel research will be applied in the price estimation for multipurpose digital devices in the Canadian CPI.

**JEL Classification System:** E31, C59, D12, L63

**Keywords:** Consumer Price Index, hedonics, web-scraped data, consumer electronics