

Big Data and Audience Measurement: A Marriage of Convenience?*

Lorie Dudoignon, Fabienne Le Sager and Aurélie Vanheuverzwyn

Key Question

Since it appeared in 1995, digital convergence has gradually altered both the data and media worlds, highlighting the power of Big Data. Should Big Data be seen as risking a return to the paradigm of census that prevailed until the end of the 19th century – whereas the 20th century belonged to sampling and surveys? This is the question that Médiamétrie is facing in its audience measurement systems.

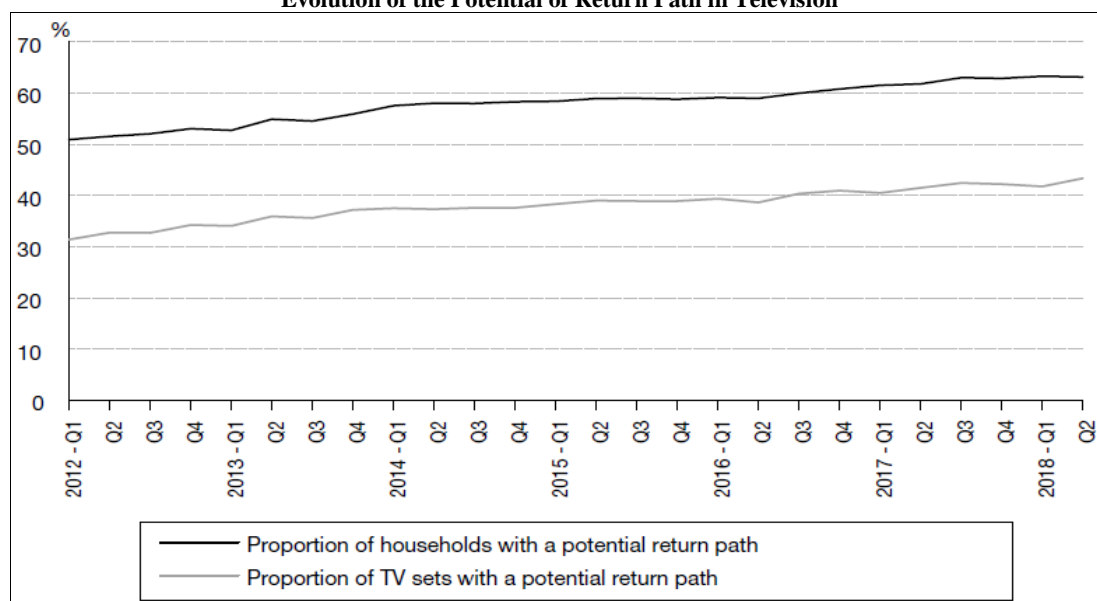
Methodology

The analysis focuses first on the relative advantages and limits of survey data and big data in terms of quality, which allow to explain the choice of so-called hybrid approaches, which consists in mixing two data sources to create a third, richer or more detailed one.

Main Results

- Quality issues concern big data as well as survey data.
- Internet audience measurement is based on two sources: the first-one is user-centric and the second one is site-centric. The proposed hybrid approach for Internet audience measurement consists in introducing site-centric results as auxiliary variables in the survey calibration.
- In television, the return path is technically possible for all set-top boxes that are connected to the internet. Panel audience measurement does not always allow the most detailed measurement of very fragmented usages. The use of return path data is therefore particularly interesting for special interest channels. The proposed hybrid approach in this case relies on hidden Markov chains.

Evolution of the Potential of Return Path in Television



Coverage: Metropolitan France
Sources: Médiamétrie, Home Devices.

Message

Big Data should not be seen as an alternative to traditional audience measurement systems, based on sample surveys, but as an opportunity to improve them.