

[doi:http://dx.doi.org/10.15414/afz.2015.18.si.148-150](http://dx.doi.org/10.15414/afz.2015.18.si.148-150)

Organic market data: Bridging the gap between data collectors and end users

Robert Home*, Catherine Gerrard, Corinna Feldmann

Research Institute of Organic Agriculture FiBL, Frick, Switzerland

Successful decision-making in the organic sector is dependent on the decision makers having high quality market data. However the quality of organic market data in Europe varies greatly and only basic statistics exist in many European countries. Improvement of data quality first requires a usable definition of what constitutes data quality so that data suppliers can make any necessary changes. In this study, we aim to evaluate the quality criteria that are used in Europe, and we use the case of retail sales data to identify problems in the quality of that data type. We surveyed data collectors and data end users in two separate surveys. The results show that the commonly used data quality indicators represent at least two, or possibly even three dimensions, which has implications for their use for data collectors who wish to improve the quality of the data they supply. We conclude with some first steps that collectors may take to improve data quality.

Keywords: organic market data collection, challenges, needs of end users, Europe

1 Introduction

Despite the growth of the organic market in Europe, only basic statistics exist in most countries. Data and market information are needed by members of the organic supply-chain to make investment decisions, and by policy makers to calibrate measures targeted to the sector. The assurance of data quality has become a central issue for national and international statistical institutions that deal with agricultural commodity data (Recke, 2004). There is a need for quality in organic market data and statistics, or the users, and therefore by extension, the whole organic industry, will be at a competitive disadvantage. Hamm and Zanolli (2006) gave a number of reasons for the urgent need for reliable data on European organic production, retail sales and trade. These include a lack of market transparency, which meant that producers and policy-makers were unaware of products for which demand was high and also unaware of other areas of the market where demand was already being met by production. This resulted in a situation where incentives to convert to organic production were sometimes targeted at areas where production then exceeded supply, resulting in reduced premiums; some producers re-converting away from organic production; and effectively, a failure of the incentives to produce a long-term shift to organic production.

Assessment of the quality of the data used by market actors has proven to be difficult and often relies on the intuition of data collectors. While the opinions of collectors are useful in indicating whether quality is acceptable, it is without theoretical basis and inadequate to provide diagnostic information (Michelsen et al., 1999). What this means is that a general rating of data quality gives little clue as to what is good and what is not so good. The European Foundation for Quality Management (EFQM) Excellence model takes the approach of dividing quality into identifiable dimensions (or criteria) (Recke, 2004) that can be analysed individually. This approach of dividing quality into identifiable criteria appears promising and is explored further in this study.

The aim of this study is to identify whether the quality criteria, which have been developed from the perspective of data collectors, are also appropriate from the perspective of end users of organic market data. A further aim is to assess whether the data quality needs,

* **Correspondence:** Robert Home, Research Institute of Organic Agriculture FiBL, Ackerstrasse 113, Postfach 219, 5070 Frick, Switzerland. E-mail: robert.home@fibl.org

expressed in terms of measurable quality dimensions, are being met by the existing methods of data collection in Europe. There are different types of organic market data including production, retail sale, price, and import /export data. In this paper we will concentrate on the case study of retail sales data, because of the relevance and collection frequency of this data type (Gerrard et al., 2012). We present an overview of the situation in Europe to gain insights into how data collection methods impact perceived data quality so that prescriptive conclusions can be drawn.

2 Material and Methods

To understand the quality of data, we take the approach of surveying the organisations that collect, analyse and/or disseminate them and identifying the methods that they use. Furthermore, we survey end users of organic market data to identify their needs and demands, and to find areas of information asymmetry. To organise and compare the results of both surveys, we will use the commonly used data quality dimensions. An online survey with additional structured interviews was carried out with data collectors in 2012 (Gerrard et al. 2012; Feldmann and Hamm 2013), which was complemented by an online survey of data users, also in 2012 (Home et al. 2013).

3 Results

The notable finding of the study was that remarkably few collection organisations appear to systematically control the quality of their data. The comparison of the data types, when listed in descending order according to the proportion of collectors who conduct data quality checks was reasonably similar to the descending order of the mean quality rating of the data types. Exceptions to this trend were, for example, export volume data, for which only 3.2 % of collecting agencies conduct quality checks, but which received a high mean quality rating. This result suggests the advantage of conducting such quality checks, but the exceptions further suggest that quality checks cannot guarantee good quality data. Users of retail sales data reported a mean agreement that the data is of overall good quality with a mean rating of 3.30 (on a scale of 1 to 5), which may be related to the reasonable number of collection organisations that report conducting quality checks on retail sales data. Although relatively few collection organisations conduct quality checks on their data, the accuracy rating, with mean 3.45, was quite high for this data type.

An interesting result with regard to the accuracy rating was the strong correlation (Pearson Correlation = 0.844) with the overall quality of the data. This result underlines the importance of accuracy to perceived data quality and shows that it should remain a focus for data collection agencies. Comparability also correlated reasonably strongly with overall data quality (Pearson Correlation = 0.676), although it can be argued that this attribute relates more to the usability of data than to its quality. The conclusion then is that usability is a valid contributor to the overall construct of quality. Users of retail sales data reported a mean agreement that the data is timely and punctual of 3.03, which is not particularly high. Most data are usually published annually but most users wish for it to be published monthly and 78% expressed the wish that it be published more often than annually. The correlation analysis revealed that general data quality correlates reasonably with both timeliness (Pearson Correlation = 0.572) and punctuality (Pearson Correlation = 0.785), and thereby underlines their validity as indicators of perceived data quality.

We also conducted a principal components analysis using principal axis factoring, which revealed two factors which combined to explain 72.5% of the variance. This suggests that the selected attributes are very well suited for describing organic market data. The first factor explained 36.5 % of the variance and loaded against everything except affordability. The first factor loaded most strongly against accuracy, punctuality, and overall quality. We can interpret this factor as an indicator of the perceived quality of the data. This suggests that affordability is not considered to be a quality criterion for organic market data. The second

factor explained a further 36 % of the variance and loaded against everything except relevance. The second factor loaded most strongly against affordability, timeliness, and accessibility. We can interpret this factor as revealing indicators of the convenience to use the data. The finding that relevance did not load against this factor can be explained in that relevance does not contribute to convenience, but is rather a prerequisite for data use. End users will not use data that is irrelevant to them and will use relevant data; even when that data is not convenient to use.

4 Conclusions

The results of this study suggest that some of the quality characteristics are useful and others less so. Timeliness, punctuality, accessibility, comparability and especially accuracy are perceived to be indicators of data quality. The strongest correlation between data quality and any of the indicators was with accuracy. The attributes that appear to be less useful as indicators of quality are relevance and affordability. Relevance is all well and good, but all data will be relevant to some people and not to others, which makes it a prerequisite for data use rather than an indicator attribute of quality. The results show that people will use data that is relevant: even if the other quality criteria are judged quite badly. Affordability appears not to be perceived as an indicator of data quality in the minds of end users but can rather be seen as an indicator of convenience. A clear and simple undertaking that data collectors might make to improve the quality of their data is to perform quality checks. A further step that can be taken to enhance uptake and use of their data is to concentrate on dissemination. Many respondents report that existing data does not exist, which suggests that the users cannot find the data and therefore points to dissemination problems.

5 Acknowledgements

Work for this publication was undertaken as part of the research project "Data network for better European organic market information" (OrganicDataNetwork). This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no. 289376. The opinions expressed in this contribution are those of the author and do not necessarily represent the views of the European Commission

References

- FELDMANN, C. and HAMM, U., (2013) *Report on collection methods: Classification of data collection methods*, Unpublished report, OrganicDataNetwork.
- GERRARD, C.L., VIEWEGER, A. and PADEL, S. (2012) *Report on data collectors: Inventory of data collecting and publishing institutions*, Unpublished report, OrganicDataNetwork.
- HAMM, U. and ZANOLI, R. (2006) *The need for a long-term strategy for data collection on organic markets*, [Online]. Joint Organic Congress, Denmark, May 30-31, 2006. Available at: <http://orgprints.org/7140/>
- HOME, R., LOŠTÁK, M. and STOLZE, M. (2013) *Survey of the organic market data needs of end users*. Unpublished report, OrganicDataNetwork.
- MICHELSSEN, J. et al. (1999) *The European market for organic products: Growth and development, Organic farming in Europe: Economics and policy*. Stuttgart: University of Hohenheim.
- RECKE, G. et al. (Eds). (2004) *Development of a European Information System for Organic Markets – Improving the Scope and Quality of Statistical Data, Proceedings of the 1st EISfOM European Seminar, held in Berlin, Germany, 26-27 April, 2004*.