How the U.S. Census Bureau E-Commerce Figures Overestimate Output and Online Sales

Roger R. Betancourt

Department of Economics

University of Maryland

College Park, MD 20742

USA

Email: betancou@econ.umd.edu

August 14 2018

Abstract:

I identify biases in the definition of e-commerce employed by the Census Bureau. They arise from using a concept based on the role of ICT in a transaction that is measured empirically on the basis of sales/revenues. I quantify this bias as a measure of output growth in the retail sector for the most important e-commerce retail sector at the NAICS 4-digit level (4541). In addition, a range of estimates for the bias as a measure of the growth of actual online sales in this sector is derived. Potential solutions are provided in a concluding section.

Key Words: E-commerce definition, bias as output measure, bias as online sales measure, purchase process, omni behavior, channel customer/user distinction

JEL Classification Codes: L81; L8, M3

1. **Introduction**

For many years the U.S Census Bureau and other statistical agencies, e.g., BLS, have been trying to incorporate the measurement of the digital economy into their activities. Early on they settled on a conceptual definition of e-commerce (e.g., Fraumeni, Manser and Mesenbourg 2000) based on the nature of the transaction "Electronic agreement, not payment, is the key determinant of an e-commerce transaction." Empirical implementation of this conceptual definition, however, has been based on the use of sales/revenues. With the advantage of hindsight, this brief essay brings out a fundamental problem with this actual implementation by the U.S. Census Bureau and suggests potential solutions to the misperceptions generated by the actual implementation of the concept.

2. **Essence of the Problem.**

If one looks at the Annual Retail Trade Survey (ARTS) data for anyone year, e.g., 2015, one finds a table with E-commerce and Total of Retail Trade data provided by the U.S. Census Bureau at the 3- 4digit NAICS level titled, Estimated Annual Retail Sales – Total and E-Commerce: 1998-2015. It is one of 12 tables publically available at The U.S. Census Bureau site for the retail trade sector based on this survey. According to this table, in 1998 the value of e-commerce for electronic shopping and mail order houses (NAICS 4541) was $4.008 billion and in 2014 it was $256.259 billion. Thus, the annual compound growth rate of e-commerce sales in this category over this period that can be derived from the table figures is 29.68%. It seems too good to be true. Indeed, it is not true in two very precise senses. That is, this e-commerce growth rate is not a good indicator of the growth of online economy output or online revenues in the retail sector.

Mechanically, the reason for the seemingly too rapid growth rate can be deduced from a footnote defining e-commerce in any table where it appears. It is also available on the U.S. Census Bureau website as a response to frequently asked questions about e-commerce. "E-commerce sales/revenues are defined as the sales of goods and services where the buyer places an order, or the price and terms of sales are negotiated over an Internet, mobile device (m-commerce), extranet, EDI network, electronic mail, or other comparable online system. Payment may or may not be made online." The last sentence in the definition introduces the possibility of substantial double counting in the e-commerce measure.

In the retail sector the double counting problem can be illustrated with a simple example. A customer visits a brick and mortar (B&M) site and uses a mobile to 'negotiate' the following terms: current availability of the item at the store, price, and reservation of item for pick-up and payment after finishing shopping at store. When finished with shopping, the customer pays for all of the items, including the one 'negotiated' over the mobile, at the store. At least the 'negotiated' item, but perhaps all, would count as an e-commerce sale by the US Census definition even though it was picked-up and paid at the store and, thus, could or should also count as a B&M sale.

Furthermore, many other examples would entail double counting of a B&M sale in terms of where payment is made as an e-commerce sale due to the use of online tools as an input in the purchase process. All revenues from such purchases, however, are treated as an output of e-commerce when the e-commerce figures are used as an indicator of output. Moreover, the definition is also asymmetric. If a customer places an order and pays online but picks up the item at the store, it is classified as an e-commerce sale without any mention or record by the U.S. Census Bureau of the B&M input in completing the purchase.

3. **Quantitative Importance of Bias in use of E-commerce Figures as Output Measure**.

Does the use of this definition matter quantitatively for any concept we want to measure as economists? If we want to use e-commerce figures as an indicator of output of the retail sector and its evolution, it certainly does. It is conventional wisdom that the appropriate measure of gross output in retailing empirically (e.g., Yuskavage 2006) and conceptually (Triplett and Bosworth 2004: Ch. 8) is the retail gross margin. The digital economy does not change this feature. The same ARTS 2015 generates a table of retail gross margins over the 1993- 2015 period as one of the 12 available. Hence, we can calculate the annual compound growth of the gross margin for the retail category electronic shopping and mail order houses for the same 1998-2014 period. The yearly growth of this 'true' measure of output is 10%.

As a measure of the 'true' growth rate of output, the yearly percentage bias in e-commerce (B) would be B= (GE –T)/T x 100%, where GE is the e-commerce growth rate in the electronic shopping and mail order category (NAICS 4541), and T is the true growth rate of output in this category, i.e., B = 196.8 %. This measure also represents a substantial overestimate of the 'true' growth rate of e-commerce measured in terms of sales/revenues paid online, Te. In 2014 the same table shows that total retail sales in this category amounts to $386.065 billion. Thus, Census measured e-commerce represents 66.4% of total retail sales in this category. Hence, Te, can be found from solving 10 = Tne (.336) + Te (.664), given that the true growth rate of non-e-commerce, Tne, is measured without bias. The latter, estimated as the annual compound growth rate of the difference between the total revenues and the e-commerce revenues as calculated by the Census, equals 3.38%. Thus, Te using end of period weights above is 13.35% and using beginning period weights (.9501, 0499) is 10.35%. As a measure of true retail

growth of e-commerce in terms of online sales, the bias in e-commerce, as calculated by the Census, ranges from 122.32% to186.76%.

4. **Extent of the Problem**

The Census definition of e-commerce is employed in four sectors: retail trade, merchant wholesalers, selected service sectors and manufacturing. For manufacturing, it is defined in terms of the value of e-shipments but also regardless of whether or not payment is made online. E-commerce figures in these other three sectors would also overestimate output but by how much quantitatively would depend on what the appropriate reference true output is in these sectors. The latter is not as easily identified as in the retail sector (e.g., for a discussion of the difficulties in service sectors see Griliches1992).

More generally, besides e-commerce figures leading to a substantial overestimate of online revenues as an output measure, what is the damage done by this double counting of B&M revenues as online revenues? One answer is generalized confusion to such an extent that even the producers of the data seem unaware of their limitations. The latter is best illustrated by a recent report from the Office of the Chief Economist of the Economics and Statistics Administration of the US Department of Commerce where the following statement is made (Nicholson 2017, p.5) "The index for non-e-commerce retail sales rose from 100 in 1998 to 170 in 2015, an average of just 3 percent per year. The index for e-commerce sales increased from 100 in 1998 to a whopping 6830 in 2015—an annual growth rate of 28 percent, nine times the growth of all other retailers."

While the non-e-commerce figures are measured without double counting revenues, the e-commerce ones are. If you eliminate the double counting of e-commerce sales figures relying on B&M payments that are somewhat arbitrarily also classified as online revenues, because of

the usage of ICT tools somewhere in the purchase process, the growth rate of e-commerce retailers (in terms of their actual online revenues) would be considerably less than nine times that of the non-e-commerce ones. Indeed, it ranges from as low as 3.06 to 3.36 times based on section 3 estimates. Finally, the problem is not limited just to U.S. statistical agencies but it also affects other advanced nations, e.g., Eurostat uses a similar definition of e-commerce, i.e., ignoring whether or not payments are made online.

5. **Potential Solutions**

A simple short-run solution would be to expand the footnote defining e-commerce to include a warning that the figures imply double counting as a measure of output or as an indicator of strictly online revenues. Nonetheless if a problem is so persistent and widespread, and incorporated in the work of many well-trained professionals, it is unlikely to have persisted by accident. It elicits the question-- What underlies its persistence? Online technology is ubiquitous and it can affect all aspects of the purchase process. As economists we want a measure of its importance. The Census Bureau provides one measure by relating it to the generation of revenues both directly and indirectly. An index based on revenues is unsatisfactory for two reasons: conceptually, it is not a good measure of online output; practically, it leads to comparisons of one series based on non-e-commerce measured by B&M payments and a series based on e-commerce defined by including some B&M payments by one criterion as online revenues for e-commerce measurement.

More constructive solutions would entail developing an index of ICT usage in the purchase process that can be related to variables other than revenues, e.g., time on a task in the purchase process. This would require gathering of new data in the case of the U.S. Census Bureau, but perhaps not in the case of Eurostat. The latter agency has a survey on ICT usage by

households and by individuals (http://ec.europa.eu/eurostat/cache/metadata/en/isoc_i_esms.htm) in addition to the one on enterprises (http://ec.europa.eu/eurostat/statistics-explained/index.php/E-commerce_statistics). An alternative index of usage might be designed from a distinction in the literature on omni behavior (e.g., Betancourt 2018). The latter concept differentiates between customers of a channel, who as a result of purchasing in a channel can be users of all of the channel services, and users of a channel services, who may or may not be purchasers in a channel.

The distinction generates five mutually exclusive categories that can be applied to purchase processes: 1) purchases by B&M customers who are not users of online channels; 2) purchases by online channel customers who are not users of B&M channels; 3) purchases by customers of B&M channels that are only users of online channels 4) purchases by online channel customers who are only users of B&M channels and 5) customers and, thus, users of both channels. The current Census Bureau definition of e-commerce treats the sum of categories 1) part of 3) and part of 5) purchases as non-e-commerce purchases, and the sum of categories 2), 4) and the remaining parts of 3), and 5 purchases as e-commerce purchases. Perhaps the five basic categories could be the basis for an e-commerce **usage** index that is not based directly on asymmetric, ambiguously defined sales/revenues. Sales/revenues should be measured wherever payment is made without exception.

References

Betancourt, Roger R.2018. "Understanding the Digital Economy in Retailing and Its Economic Impacts," (last accessed August 9 2018) http://econweb.umd.edu/~betancourt/distribution/Understanding%20the%20Digital%20Economy%20in%20Retailing%20and%20Its%20Economic%20Impact.FinalRev2.pdf

Fraumeni, Barbara, Marilyn E. Manser and Thomas L. Mesenbourg. 2000. "Government Statistics: E-Commerce and the Electronic Economy," paper prepared for presentation to the Federal Economic Statistics Advisory Committee (FESAC). (last accessed August 9 2018). https://www.census.gov/library/working-papers/2000/econ/fraumeni-01.html

Griliches, Zvi (editor). 1992. *Output Measurement in the Service Sectors.* Chicago: University of Chicago Press.

Nicholson, Jessica R. 2017. "New Insights on Retail E-Commerce," ESA Issue Brief #04-17(last accessed August 9 2018) http://www.esa.doc.gov/sites/default/files/new-insights-retail-e-commerce.pdf

Triplett, Jack E. and Barry P. Bosworth 2004. *Productivity in the U.S. Services Sector*. Washington DC: The Brookings Institution Press.

U.S. Bureau of the Census. 2015. Annual Retail Trade Survey. (last accessed August 9 2018) https://www.census.gov/data/tables/2015/econ/arts/annual-report.html

Yuskavage, Robert E. 2006. "Distributive Services in the U.S. Economic Accounts," Paper Prepared for the NBER Conference on Research Income and Wealth Summer Institute. July 17. (last accessed August 12 2018) https://www.bea.gov/papers/pdf/distributive_services_in_the_us_economic_accounts.pdf