

The Sixth IMF Statistical Forum

SESSION III: WHAT IS THE VALUE OF UNPRICED SOFTWARE AND DATA? PLATFORM AND MARKETS

November 19, 2018

CHAIR: Kristina Kostial (IMF Strategy, Policy, and Review Department)

PRESENTERS: Carol A. Robbins (National Science Foundation, U.S.);

Wendy Li (U.S. Bureau of Economic Analysis);

Rana Foroohar (Financial Times).

SESSION TOPIC

Free software and free services supplied by platforms that collect users' data are prominent in the digital economy. Valuing open source software and the data of digital platforms at zero seems unsatisfactory even though in national accounts value is usually inferred from prices. On the downside, do platforms put a risk to growth and stability?

SUMMARY OF PRESENTATIONS

1. ***Open Source Software as Intangible Capital: Measuring the Cost and Impact of Free Digital Tools*** (Carol A. Robbins, Gizem Korkmaz, José Bayoán Santiago Calderón, Claire Kelling, Stephanie Shipp, and Sallie Keller)

The authors develop an approach to document the scope and impact of open source software (OSS) packages created by all sectors of the U.S. economy. They use a bottom-up approach to measure a subset of OSS projects and languages, collecting data on open source software R, Python, Julia and JavaScript, as well as from the Federal Government's code.gov website. The authors use lines of code and a standard model to estimate a package developer's time, converting lines of code to resource cost. They find that the resource cost for developing the aforementioned software packages exceeds US\$ 3 billion, based on 2017 costs. They also find that software programs available on code.gov can be valued (resource cost) at more than US\$ 1 billion, also based on 2017 costs. Finally, the authors explore network relationships between OSS packages (that is, their re-use statistics) to measure their impact; they observe that the R dependency network has the highest number of connected components.

2. ***Value of Data: There's no Such Thing as a Free Lunch in the Digital Economy*** (Wendy Li, Makoto Nirei, and Kazufumi Yamana)

Online platform companies exchange “free” goods and services for consumer data. Welfare analysis on digital goods or services without considering the value of data can mislead policy analysis. The authors classify online platforms based on underlying business models and conduct case studies to analyze the related data activities. The authors show how online platform companies create the value of data and present a data value chain to show how the value of data varies by step. Unlike R&D that may depreciate due to obsolescence, data can produce new values through data fusion, creating unprecedented measurements challenges. Initial estimation shows that the value of data can be tremendous. Moreover, online platform companies can capture most benefits of the data because they create the value of data and consumers lack knowledge to value their own data. Lastly, the Internet of Things, the trend of 5G, and the emerging online-to-offline transition are accelerating the speed of data accumulation. The valuation of data will have important policy implications for investment, trade, and growth.

3. *Platforms, Markets and Risks to Growth and Stability* (Rana Foroohar)

The speaker presents two main reasons for the rise of large platform tech firms: (i) the network effect, which is the ability to harness users and increase value of offerings by leveraging greater and greater numbers of those users; and (ii) the power of data, which has major policy ramifications. She notes how these topics are playing out in the business ecosystem and the policy field, the new huge risks posed by the use of consumer data, and the need for regulation. There is a move towards “industrial internet”, as all kinds of companies are using data as a valuable asset, changing their historical business models. On the risks, some of these new business models can be detrimental to some consumers (e.g. less risk pooling in the insurance sector, information asymmetry and reduced competitiveness and fairness in the market in some sectors) and reduce transparency (with companies’ practices appearing more and more as algorithmic black boxes). The speaker closes the presentation by advocating for a rethink of regulation, particularly antitrust action.

QUESTIONS AND ANSWERS:

Questions focus on: (i) the role regulation can play to facilitate data access by statisticians for economic and social analysis, while preserving anonymity of firms and individuals; (ii) the appropriate method for valuing data (given the high variability of values obtained depending on the method used) and how policy makers should approach the large number of measures of the value of data; and (iii) whether the so-called “value of data” may just be capturing some sort of “winner-take-all” dynamic created by network effects. Some speakers note legal barriers and not obtaining explicit consent from survey respondents as challenges to data sharing, while others note data could be shared without disclosing consumer identity, and that more transparency is needed for consumers to understand potential data use and to facilitate access to own data. Regarding different valuation methods, measuring resource costs may be a lower bound estimate but it is the approach most easy to implement, as well as a standard approach for measuring own-account software and R&D investment; in addition, it is easy to explain to data users. Whether data should be capitalized (possibly by using the cost-based approach) should be investigated in

a future update of the System of National Accounts (SNA). Finally, companies, especially small and medium sized companies (SMEs) use other means besides patents to protect their intellectual property, such as through copyrights and trademarks.