Challenges in Coupling Digital Payments Data and Input-output Data to Change Consumption Patterns

Pizzol, Massimo; Vighi, Elena; Sacchi, Romain

Published in:
Procedia CIRP

DOI (link to publication from Publisher):
10.1016/j.procir.2017.11.004

Publication date:
2018

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Pizzol, M., Vighi, E., & Sacchi, R. (2018). Challenges in Coupling Digital Payments Data and Input-output Data to Change Consumption Patterns. *Procedia CIRP*, *69*, 633-637. https://doi.org/10.1016/j.procir.2017.11.004

Available online at www.sciencedirect.com
ScienceDirect
Procedia CIRP 69 (2018) 633 – 637
www.elsevier.com/locate/procedia

25th CIRP Life Cycle Engineering (LCE) Conference, 30 April – 2 May 2018, Copenhagen, Denmark

Challenges in coupling digital payments data and input-output data to change consumption patterns

Massimo Pizzol[a]*, Elena Vighi[a], Romain Sacchi[a]

[a]*Department of Planning, Aalborg University, Rendsburggade 14, 9000 Aalborg, Denmark*

* Corresponding author. Tel.: +45-9940-8369. *E-mail address:* massimo@plan.aau.dk

Abstract

As society become more digitalized, new possibilities of combining data arise. This study describes the attempt to couple information on digital expenditures obtainable from bank accounts with LCA-like data to calculate personalized and consumer-specific environmental footprints of consumption patterns. In Denmark, most purchases are performed using credit cards, and this information is available from consumers' bank accounts. These data were matched with the categories of the Multi-Regional Input-Output database EXIOBASE via different conversion tables for Danish and European industrial classifications (DB07, NACE1-2) with an automated procedure. This matching allows then to calculate the carbon footprint of the specific consumer. In principle, this information could be used to raise awareness on consumption-related impacts and trigger behavioral change. However, several challenges exist to reach this goal, such as data aggregation, privacy, and communication issues. This article describes a first experimental attempt to match digital expenditures with Input-Output data and touches on the challenges ahead.

Peer-review under responsibility of the scientific committee of the 25th CIRP Life Cycle Engineering (LCE) Conference

Keywords: Digitalization; Life Cycle Assessment; Environmentally Extended Input-Output Analysis; Electronic payment

1. Introduction

Together with technological advancements, a deeper environmental awareness and behavioural change are urgently needed to disrupt the world's current unsustainable production and consumption patterns [1]. While consumers are generally aware that low-tech behavioural changes like shifting to car-free commuting or reducing meat consumption can bring substantial environmental benefits, especially when adopted by large groups [2], they are mostly unaware of what and how large the environmental effects of their personal consumption pattern are. In principle, empowering consumers with this kind of *information feedback*, i.e. detailed and personalized knowledge about their environmental impact, should allow them to revise their everyday consumption by changing the type and reducing the amount of products purchased [3].

The information feedback has been mostly studied in the context of household energy consumption where its effects are well-known [4]. Several studies observed behavioral change [5] in terms of energy savings achievable with smart-meters [4]. Since households are unaware of their domestic energy use, providing them a feedback that makes energy consumption visible allows reducing it. This feedback is achievable *directly* via raw information provided by an energy meter with display or *indirectly* via pre-processed information e.g. a monetary energy bill. The frequency, duration, and level of understanding of the feedback all contribute to its effectiveness, which is up to 20% energy saved [4]. This information feedback can be provided for other types of consumption than energy, for example several tools are available to calculate the environmental impact of an individual's *consumption pattern* (type of diet, travel habits, house, and products purchased) measured as carbon footprint (c.f. WWF [6], US EPA [7], and

Peer-review under responsibility of the scientific committee of the 25th CIRP Life Cycle Engineering (LCE) Conference
doi:10.1016/j.procir.2017.11.004

others [6–9]). These calculators are usually accessible via websites where by prompting personal details about his/her consumption pattern the user obtains a carbon footprint estimated via fixed coefficients. The process is time consuming and could be automatized by using information on the consumer's digital payments, i.e. payments made using electronic devices and channels.

Digital payments are extremely common in Scandinavia and in the Nordic countries where consumers are acquainted with cash-free purchases in both physical shops and web-shops, as confirmed by a recent study of the Danish National Bank [10]. According to another Danish study 40% of payments under 100 Danish Krone (approx. 13.44 euro) are made with credit card, and almost 70% of larger payments [11]. The Central European Bank reports that in Denmark 81,3 % of all digital payments were card payments in 2016, the highest value in Europe [12], and that this value has been increasing over time. Today, new online banking features allow Danish consumers to generate reports where their digital payments are classified into different categories such as "energy", "groceries", "transport", etc. This feature builds upon a pre-existing national registry of all companies operating in e.g. Denmark, that are categorized into so-called branches. The Danish Branch Code 2007 (DB07) is in fact a statistical classification of all economic activities, and is the Danish version of the European Classification NACE Rev. 2. Interestingly, the NACE classification is also what new Multi-Regional Input-Output databases (MRIO) are built on, such as EXIOBASE [13]. MRIO databases allow quantifying the impact of nearly all economic activities in a life cycle perspective, i.e. taking into account the contributions from all other activities, with high completeness [14–16]. Previous research shows that MRIO databases can be effectively integrated into carbon calculators [17,18].

Summing up, the digitalization of consumer expenditures via electronic payment (e.g. credit cards) allows for new possibilities of delivering an information feedback to consumers. In principle, it is now possible to match digital information on which goods are purchased with information on the environmental impact of the respective production activities. This information could be used to raise awareness on consumption-related impacts and trigger behavioral change. However, several challenges exist to reach this goal: it is unclear how well the data matching can be achieved, how and in which format the information should be delivered to consumers, and how large the effect of the feedback would be. Besides calculation and communication issues, data aggregation and privacy issues may constitute a challenge.

In this context, the objective of this study is to present the results of an experimental attempt to match digital expenditures and input-output data, and to address the challenges in upscaling this process.

2. Methods

In this study, electronic payment data were matched to EXIOBASE via NACE using an automatized procedure implemented in the Python programming language. By matching payments with the input-output database, it is in principle possible to calculate what the environmental impact of these expenditures is, e.g. in terms of carbon footprint, and to approximate the environmental burden of the consumer. This is possible because EXIOBASE is a so-called Environmentally-Extended database, where the e.g. carbon emissions associated with each industrial sector are reported.

The study is a proof of concept to illustrate the suggested approach. One year of expenses drawn from a Danish bank account statement have been used as test sample. Entries in the list are in the format *date,text*. For example: *2016-07-28,DK-NOTA30373 FØTEX FREDERIKS AL*. In this specific case the text *DK-NOTA* indicates that the payment was done via credit card. The first step was to format the entries in the list of expenditures. Special characters were removed and the text was converted to lower case. Extra white spaces were stripped and recurrent non-significant characters removed. This allowed to isolate the names of the creditors from any undesired "noise". Fuzzy string comparison was then used to match the entries in the list of expenditures with the more than 500 thousand Danish companies in the register, to find the industrial code (DB07) of the creditor. Fuzzy string comparison is a technique that allows to match, with a returned level of confidence, two strings of characters that are not strictly similar, based on the calculation of the editing distance between the two strings.

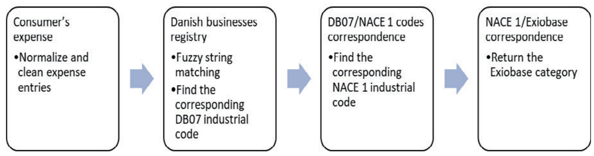

Figure 1 Connecting the consumer's expenses to EXIOBASE categories.

Two Python's fuzzy string comparison libraries are tested: *fuzzywuzzy*[*] and *difflib*[†], both with a cut-off ratio of 80% (a match between two strings is here deemed valid when less than 20% of the characters of one string need to be edited to equal the other). The expenses and their Danish industrial classification codes are then linked to their NACE 1 code counterpart. Finally, a correspondence map between NACE 1 codes and EXIOBASE categories[‡] is used to connect the expenses and their amounts to the EXIOBASE categories. All of the matched expenses are linked to at least an EXIOBASE category. It should be noted that while DB07 refers to NACE Rev. 2., EXIOBASE refers to NACE Rev. 1. In other words, the two databases refer to two different versions of the same classification. Since there isn't a direct connection between

[*] https://github.com/seatgeek/fuzzywuzzy

[†] https://docs.python.org/2/library/difflib.html

[‡] The correspondence map links the yet unpublished EXIOBASE v3 categories with NACE Rev 1 classification and is an internal working version kindly provided by Jannick Schmidt [24]. EXIOBASE v3 categories are almost identical to v2 categories.

DB07 and EXIOBASE, these two databases were linked using the conversion key between NACE Rev. 1. and Rev- 2.. The sequence is summarized in Figure 1.

3. Results

The selected set of expenses encompasses over 500 transactions with 195 different recipients, distributed as per Figure 2 . The high number of payments to domestic creditors and the limited use of cash withdrawals leads to believe that we capture most of the expenditures.

The *fuzzywuzzy* library performs better with almost 32% of the domestic expenses matched in the company register (against 22% for *difflib*). However, we were not able to match 68% of the domestic expenses. These include 1) expenses that can actually be matched with a Danish company, but the company name differs in the register and in the expense list, e.g. for local shop part of a chain of stores; 2) expenses that can't be matched with any Danish company, e.g. expenses done abroad, transfers between bank accounts, or positive intake like e.g. salary or refunds.

The correspondence between the Danish company register and the EXIOBASE categories via the Danish/European industrial classification codes was rather good as 91% of the entries connect to a category. However, a challenging 41% of the companies in the register connect to more than one category in EXIOBASE as some companies inherit from more than one industrial code. On average, the companies that connect to more than one EXIOBASE category can be linked to 2,65 categories. Moreover, 132 of the 148 (89%) EXIOBASE categories link to at least one company in the Danish register.

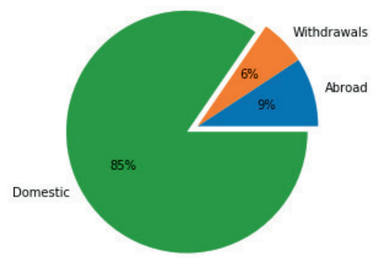

Figure 2 Distribution of bank transactions between withdrawals, payment to creditors based abroad and domestic creditors.

The distribution between the EXIOBASE categories that match with the expenses is reported in Figure 3. The categories i55 ("hotels and restaurants") and i52 ("retails and shops") represent more than half of the matched expenses. They are followed by i51 ("wholesale trade"), i80 ("education") and i70 ("lodging", "real estate activities").

4. Discussion

The study is a proof of concept, and for this reason only data from a single bank account were used in the experimental analysis, even though this clearly limits the general validity of the results. Nevertheless, this experiment allows for several learning opportunities. It also highlights some of the challenges in upscaling this process to generate an information feedback about the environmental impact of a consumer's expenditures.

The main problem is that not all expenses were matched with a registered company. As mentioned earlier, web banking services are currently able to categorize almost all expenses and create reports for the bank account owner. This means that the providers of this service have access to better data than we have, and that in theory the matching can be substantially improved by collaborating with banking institution and the Danish Central Business Register.

A practical problem of the input-output data is that we fail to obtain a complete match between the DB07 categories and the EXIOBASE categories. This means that a digital expense can be classified according to several different EXIOBASE categories. In order to calculate the carbon footprint of these expenditures, either a selection of the most relevant category or a sort of allocation between the EXIOBASE categories would be necessary, if the procedure has to be automatized. In both cases, it is unclear whether this would result in an increased accuracy or increased uncertainty.

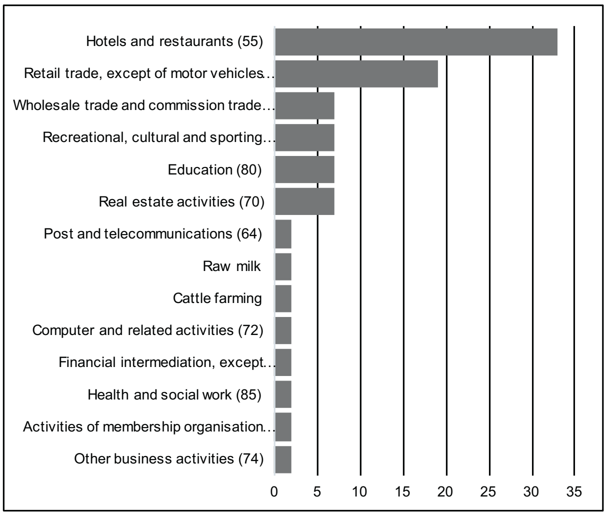

Figure 3 Frequency distribution of EXIOBASE categories in regards to matched consumer's expenses

It is worth reflecting on the general applicability of this approach. Together with Denmark, Sweden, Norway, Finland, the Netherlands and the United Kingdom are countries where cashless payments are very common[10]. It was beyond the scope of this study to investigate whether these countries possess a national register of companies linked to NACE, like the DB07 in Denmark, but is clearly a limiting factor for implementing the approach. Beyond Europe, Canada and Australia are also countries with substantial use of digital payments, but it is unlikely that their national register of companies, if available, would be linked to NACE, and other conversion keys may be necessary to match the digital expenditures with EXIOBASE. As a rule of thumb, and considering economies of scale, we estimate that this approach is worth implementing if the percentage of cashless payments on total payments exceeds 50%.

A drawback of digital payments is that some information may be too aggregated to be used as feedback. For example, only the total amount of money spent at the supermarket is reported, not the amount spent on single groceries. This may impede consumers to observe the effect of very specific changes, e.g. buying less meat. This means that the information feedback may be too coarse to be insightful. The information on expenditures can be disaggregated either by consumers directly via questionnaires, or via information from statistical agencies on the average share of budget spent by consumers on different goods.

Previous research shows that the efficacy of the feedback is reduced if the information provided to consumers is inaccessible, complex, or ill-timed [19]. The feedback may be more informative for some categories of products than others, e.g. consumables rather than durable goods [20]. It is also unclear what privacy issues may arise when making use of digital payment data from banking institutions. Moreover, the level of understanding of the different indicators need to be assessed and the mechanism beyond the behavioral change unveiled [21]. This should be investigated ideally using qualitative and semi-quantitative research methods of the social sciences, such as the involvement of consumers via focus groups, interviews, and questionnaires, and building on the theoretical framework and methods of behavioral science and environmental psychology.

Differently from energy meters, consumers may have difficulties in understanding and trusting the indirect feedback provided by aggregated indicators of environmental impact (e.g. carbon footprint measured in Carbon Dioxide equivalents), as this information is highly pre-processed. There is the need to investigate, by asking consumers directly, to what extent different indicators are understandable, meaningful, and informative in order to revise their consumption pattern. In particular, the focus should be on the assessment of the communication efficiency of different indicators (e.g. physical versus monetary), levels of aggregation of the indicators (single score versus multiple dimensions), and of the data.

The topic of how information on environmental impacts affects consumer behavior has been much debated in the vast literature on environmental labelling: studies on carbon labelling show that it may [21,22] or not [20,23] have an effect on consumption. However, research is needed to investigate and unveil the drivers beyond consumers' behavioral change for the specific case of the information feedback from digital payments. What triggers behavioral change in this context? Is information sufficient or other incentive-based mechanisms (e.g. gamification, social feedback between similar users, self-commitment) are needed? What role plays timing in the provision of feedback? Do privacy issues play any role? Answering this questions is a necessary step before large-scale implementation of any feedback mechanism based on complex impact indicators.

5. Conclusion

This study described the first empirical steps in the development of a specific application of information feedback, i.e. an experimental attempt to match digital payments and input-output data. Moreover, the study addressed the challenges in upscaling this process.

The study contributes in the direction of development of technical solutions to use Environmentally Extended Input-Output data in context and of effective disaggregation strategies. These solutions are of great interest for industrial ecologists and LCA practitioners on a broader perspective, as they can find general application in product environmental footprint assessment. Beyond these communities, the information of perception of complex indicators and behavioural change have scientific and policy relevance for e.g. the development of effective environmental labelling schemes.

The experimental work described in this study has the potential to be further developed into large scale applications of information feedback, e.g. in collaboration with developers of web and mobile apps, and banks in the Nordic countries or the EU, and has therefore high upscaling potential. At a societal level, this could bring increased levels of environmental knowledge and awareness, and lead to environmental impacts reduction.

References

[1] J.C.J.M. Van den Bergh, Environment versus growth - A criticism of “degrowth” and a plea for “a-growth,” Ecol. Econ. 70 (2011) 881–890. doi:10.1016/j.ecolecon.2010.09.035.

[2] B. Goldstein, S.F. Hansen, M. Gjerris, A. Laurent, M. Birkved, Ethical aspects of life cycle assessments of diets, Food Policy. 59 (2016) 139–151. doi:10.1016/j.foodpol.2016.01.006.

[3] United Nations Environmental Program (UNEP), Towards a Life Cycle Sustainability A ssessment: Making informed choices on products, 2011. doi:DTI/1412/PA.

[4] A. Faruqui, S. Sergici, A. Sharif, The impact of informational feedback on energy consumption—A survey of the experimental evidence, Energy. 35 (2010) 1598–1608. doi:http://dx.doi.org/10.1016/j.energy.2009.07.042.

[5] S. Seebauer, A. Wolf, Disentangling household and individual actors in explaining private electricity consumption, Energy Effic. 10 (2017) 1–20. doi:10.1007/s12053-016-9435-x.

[6] GFN, Footprint calculator, Glob. Footpr. Netw. (2016). http://www.footprintnetwork.org/en/index.php/GFN/page/calculators/ (accessed August 1, 2016).

[7] American Forest, Carbon footprint calculator, Direct. (2008). http://www.americanforests.org/discover-forests/carbon-calculator/ (accessed August 1, 2016).

[8] American Forest, Carbon footprint calculator, Direct. (2008). http://www.americanforests.org/discover-forests/carbon-calculator/ (accessed June 20, 2008).

[9] MyClimate, CO2 Emissions Calculator, Orient Overseas Contain. Line Ltd. (2015) Carbon Calculator. http://www.oocl.com/eng/aboutoocl/Environmentalcare/ooclcarboncalculator/Pages/default.aspx (accessed August 1, 2016).

[10] Danmarks Nationalbank, Analyse: Danskerne er mestre i at betale elektronisk, (2017) 8. http://www.nationalbanken.dk/da/publikationer/Sider/2017/03/Danskerne-er-mestre-i-at-betale-elektronisk.aspx.

[11] J. Gustav, K. Jacobsen, Betalingsvaner i Danmark, Copenhagen, 2011.

[12] European Central Bank, Table 7.2 Relative importance of payment

services, ECB Payments Stat. Sept. 2017. (2017) 17. http://sdw.ecb.europa.eu/reports.do?node=1000001388 (accessed November 1, 2017).

[13] R. Wood, K. Stadler, T. Bulavskaya, S. Lutter, S. Giljum, A. de Koning, J. Kuenen, H. Schütz, J. Acosta-Fernández, A. Usubiaga, M. Simas, O. Ivanova, J. Weinzettel, J. Schmidt, S. Merciai, A. Tukker, Global Sustainability Accounting—Developing EXIOBASE for Multi-Regional Footprint Analysis, Sustainability. 7 (2015) 138. http://www.mdpi.com/2071-1050/7/1/138.

[14] A. Tukker, A. de Koning, R. Wood, T. Hawkins, S. Lutter, J. Acosta, J.M. Rueda Cantuche, M. Bouwmeester, J. Oosterhaven, T. Drosdowski, J. Kuenen, EXIOPOL - Development and illustrative analyses of a detailed global MR EE SUT/IOT, Econ. Syst. Res. 25 (2013) 50–70. doi:10.1080/09535314.2012.761952.

[15] S. Nakamura, K. Nansai, Special Types of Life Cycle Assessment, in: M. Finkbeiner (Ed.), Spec. Types Life Cycle Assess., 1st ed., Springer Netherlands, 2016: pp. 219–291. doi:10.1007/978-94-017-7610-3.

[16] G. Finnveden, M.Z. Hauschild, T. Ekvall, J. Guinee, R. Heijungs, S. Hellweg, A. Koehler, D. Pennington, S. Suh, Recent developments in Life Cycle Assessment, J. Environ. Manage. 91 (2009) 1–21.

[17] K. Roelich, A. Owen, D. Thompson, E. Dawkins, C. West, Improving the policy application of footprint indicators to support Europe's transition to a one planet economy: The development of the EUREAPA tool, Sci. Total Environ. 481 (2014) 662–667. doi:10.1016/j.scitotenv.2013.10.084.

[18] S. Huysman, T. Schaubroeck, M. Goralczyk, J. Schmidt, J. Dewulf, Quantifying the environmental impacts of a European citizen through a macro-economic approach, a focus on climate change and resource consumption, J. Clean. Prod. 124 (2016) 217–225. doi:10.1016/j.jclepro.2016.02.098.

[19] S.E. West, A. Owen, K. Axelsson, C.D. West, Evaluating the Use of a Carbon Footprint Calculator: Communicating Impacts of Consumption at Household Level and Exploring Mitigation Options, J. Ind. Ecol. 20 (2016) 396–409. doi:10.1111/jiec.12372.

[20] N. Borin, D.C. Cerf, R. Krishnan, Consumer effects of environmental impact in product labeling, J. Consum. Mark. 28 (2011) 76–86. doi:10.1108/07363761111101976.

[21] M. Guenther, C.M. Saunders, P.R. Tait, Carbon labeling and consumer attitudes, Carbon Manag. 3 (2012) 445–455. doi:10.4155/cmt.12.50.

[22] C. Groening, J.J. Inman, W.T.R. <suffix>Jr.</suffix>, The role of carbon emissions in consumer purchase decisions, Int. J. Environ. Policy Decis. Mak. 1 (2015) 261. doi:10.1504/IJEPDM.2015.074719.

[23] M. Kortelainen, J. Raychaudhuri, B. Roussillon, Effects of carbon reduction labels: Evidence from scanner data, Econ. Inq. 54 (2016) 1167–1187. doi:10.1111/ecin.12278.

[24] S. Merciai, J.H. Schmidt, Methodology for the construction of global multi-regional hybrid supply and use tables for the EXIOBASE v3 database, J. Ind. Ecol. (n.d.).