2017 Journal of Industrial Ecology – www.wileyonlinelibrary.com/journal/jie

SUPPORTING INFORMATION FOR:

Min, J. and N.D. Rao. 2017. Estimating uncertainty in household energy footprints. *Journal of Industrial Ecology.*

Summary

This supporting information S1 describes each step of the process adopted for the methodology. At the end, it includes further explanation of the Randomized Iterative Proportional Fitting Procedure (rIPFP).

Details on how we bridge household consumption to MRIO sectors

1) **Map household survey consumption categories to sectors in the MRIO.** Figure 1 schematically shows the various intermediate third-party classification schemes we use to map consumption categories to EXIO sectors. This involves two steps.

   a. **Standardize CES consumption classification:** We adopt mappings between country-specific classifications for expenditure survey items and a standardized household consumption classification created by the World Bank for its Global Consumption Database (the first arrow in Figure 1).[i] The standard classification use here is the International Comparison Program (ICP) classification, equivalent to the International Classification of Individual Consumption According to Purpose (COICOP). However, this ICP classification only have a few aggregate fuel categories, for which we have higher resolutions in both surveys. Hence, we compose our own fuel classification with 13 fuel types, so as not to sacrifice the resolution in the surveys.[ii]

   b. **Use UN concordance matrices to map CES consumption categories and EXIO commodity sectors:** We develop a qualitative mapping matrix (with elements of either 0 or 1, referred to as $\boldsymbol{Q_{EXIO}}$ hereafter) which relates each of these standard ICP items to one or more ISIC sectors of EXIOBASE (the second arrow in Figure 1). The dimension of $\boldsymbol{Q_{EXIO}}$ is 164×200 with rows for ICP items and columns for EXIO sectors.[iii] To develop this, we use two concordance matrices between COICOP-CPC (Central Product Classification) and between CPC-ISIC from the UN classification registry and link them serially to get a COICOP-ISIC mapping.[iv] Note that for cases of one-to-many or many-to-one mappings, these public sources provide no information on how to allocate expenditure in the one to the corresponding many sectors.

2) **Translate purchaser prices to basic prices**. This tasks involves knowledge about subsidies and taxes, and trade and transport margins in various retail sectors, so that one can accurately characterize the flow of money from households to industry. To our knowledge, however, there are no publicly available data for India or Brazil, except those developed and provided by EXIO upon request. In the case of Brazil, we have a second source from Brazilian national IO tables, which allows us to conduct a simple sensitivity of our energy intensities to these two sources. We refer to this translation as a valuation matrix $\boldsymbol{V}$, which relates a final household demand vector in the IO table in purchaser prices $\boldsymbol{y}_{\boldsymbol{0},\boldsymbol{pp}}$ to one in basic prices, $\boldsymbol{y}_{\boldsymbol{0}}$ (the third arrow in Figure 1).

3) **Assign shares among the mapped EXIO sectors:** The last step is to assign expenditure shares among the IO commodity sectors mapped to each ICP item in the qualitative mapping $\boldsymbol{Q}_{\boldsymbol{EXIO}}$ using the algorithm described below (See Figure S1-1b). This allocation process needs to respect two constraints: 1) the total sectoral demand for each IO sector; and 2) the total consumption expenditure by survey item.

However, there is also uncertainty in defining these constraints, because total household consumption expenditure in surveys typically falls well short of that reported in national accounts, particularly in developing countries (Deaton, 2005). For example, we find for India that NSS 2011-2012 covers about 40 percent of the final household demand reported in its national account. This gap needs to be reconciled for a consistent estimation of consumption-based footprints. However, we find that similar studies stay silent or are ambiguous about how, or even whether, they reconcile total household expenditure from the two sources.[v] We adjust the total consumer expenditure using the two-step approach described below before commencing the random allocation process.

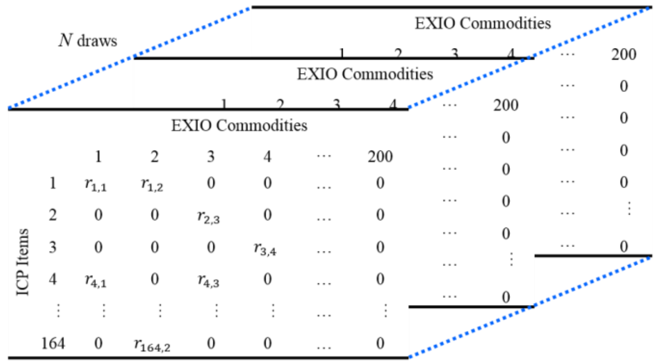
a. We start by proportionately scaling up the household expenditure from the survey to equal the sum of final demand $\boldsymbol{y}_{\mathbf{0},\boldsymbol{pp}}$, called $\boldsymbol{z}_{\mathbf{0}}$, a similar approach taken by Lenzen et al. (2004). We then identify obvious inconsistencies between the two constraints $\boldsymbol{z}_{\mathbf{0}}$ and $\boldsymbol{y}_{\mathbf{0},\boldsymbol{pp}}$ given the mapping $\boldsymbol{Q}_{\boldsymbol{EXIO}}$ between them. These inconsistencies occur when there are no possible allocation schemes that meet both constraints. In such cases, we adjust the sectoral scaling of $\boldsymbol{z}_{\mathbf{0}}$ from the surveys to match the other from the IO model (based on national accounts) called $\boldsymbol{z}_{\boldsymbol{adj}}$.

b. There are a number of expenditure allocation schemes that can satisfy these constraints for a given mapping $\boldsymbol{Q}_{\boldsymbol{EXIO}}$. There are no known data sources that shed light on this. Thus, we generate distributions of allocations that meet these constraints and calculate energy intensities for all these solutions.

For this last step, we develop a procedure that we call "randomized Iterative Proportional Fitting Procedure (rIPFP)", which is a Monte Carlo simulation scheme applied to the well-known IPFP (i.e. RAS process) used to balance IO tables (Ch7.4, Miller and Blair, 2009). We generate N random initial matrices ($\boldsymbol{B}_{\boldsymbol{i}}$, where $i = 1,2, \ldots, N$, same dimension as $\boldsymbol{Q}_{\boldsymbol{EXIO}}$), from which, by iteration, we select those that satisfy the two marginal constraint vectors ($\boldsymbol{z}_{\boldsymbol{adj}}$ and $\boldsymbol{y}_{\mathbf{0},\mathbf{pp}}$). We refer to this resulting matrix as $\widehat{\boldsymbol{B}}_{\boldsymbol{i}}$. Then, we divide each row of $\widehat{\boldsymbol{B}}_{\boldsymbol{i}}$ by the corresponding element of $\boldsymbol{z}_{\boldsymbol{adj}}$ to derive $\widehat{\boldsymbol{R}}_{\boldsymbol{i}}$, each row of which now consists of ratio value $\boldsymbol{r}$'s summing up to 1 (Figure S1-1b).

		EXIO Commodities					
		1	2	3	4	…	200
ICP Items	1	1	1	0	0	…	0
	2	0	0	1	0	…	0
	3	0	0	0	1	…	0
	4	1	0	1	0	…	0
	⋮	⋮	⋮	⋮	⋮	…	⋮
	164	0	1	0	0	…	0

(a)



(b)

Figure S1-1 Structure of a) $\boldsymbol{Q_{EXIO}}$ and b) $\boldsymbol{\widehat{R}_i}$ matrices. For the cells with 1's in the fixed $\boldsymbol{Q_{EXIO}}$, we first assign random initial shares $\boldsymbol{r_{j,k}}$ adding up to 1 in each row. Then we run the rIPFP process to find $\boldsymbol{\widehat{R}_i}$ with the cell values satisfying both marginal constraints. This is done for *N* draws.

The two marginal constraints can be expressed for each draw $i \in \{1,2,\cdots,N\}$ as:

$$\boldsymbol{\widehat{B}_i}^{\boldsymbol{T}} \cdot \mathbf{1} = \boldsymbol{\widehat{R}_i}^{\boldsymbol{T}} \cdot \boldsymbol{diag}(\boldsymbol{z_{adj}}) \cdot \mathbf{1} = \boldsymbol{y_{0,pp}} = \boldsymbol{V^{-1}} \cdot \boldsymbol{y_0} \tag{2}$$

$$\boldsymbol{\widehat{R}_i} \cdot \mathbf{1} = \mathbf{1}, \tag{3}$$

where **1** is a column matrix with all elements as 1.

References

Deaton, A. 2005. Measuring poverty in a growing world (or measuring growth in a poor world). *Review of Economics and Statistics* 87(1): 1–19.

Lenzen, M., C. Dey, and B. Foran. 2004. Energy requirements of Sydney households. Ecological Economics 49(3): 375–399.

Miller, R.E., and P.D. Blair. 2009. *Input-Output Analysis: Foundations and Extensions*. Cambridge University Press, Cambridge England ; New York, 784 p.

[i] See The World Bank Global Consumption Database, accessed 4/05/2016 at http://datatopics.worldbank.org/consumption/. Minimal changes are manually made on this mapping when more evident matches are found. These changes are in supporting information S2.

[ii] There are 13 fuel items in this classification: (Biogas, Charcoal/coal/briquette/coke, Diesel, Electricity, Ethanol, Firewood and other fuels, Other biomass, Fuel oil, , Gasoline, Kerosene, LPG, Natural gas, Other household fuel)

[iii] The entire list of these 164 consumption categories are provided in supporting information S2.

[iv] See United Nations Statistics Division Classifications Registry, accessed 9/08/2016 at http://unstats.un.org/unsd/cr/registry/default.asp?Lg=1.

[v] Steen-Olsen et al. (2016) say they scaled up the survey expenditure to match the totals based on “a static assumption”.