A Consumption-Based Hybrid Life Cycle Assessment of Carbon Footprints in California: High Footprints in Small Urban Households

Jukka Heinonen

Abstract—Higher density reduces distances, private car dependency and thus reduces greenhouse gas emissions (GHGs). As a result, increased density has been given a central role among urban development targets. However, it is not just travel behavior that changes along with density. Rather, the consumption patterns, or overall lifestyles, change along with changing urban structure, particularly with changing housing types and consumption opportunities. Furthermore, elevated consumption of services, more frequent flying and less intra-household sharing have been shown to potentially outweigh the gains from reduced driving in more dense urban settlements. In this study, the geography of carbon footprints (CFs) in California is analyzed paying close attention to the household size differences and the resulting economies-of-scale advantages and disadvantages. A hybrid life cycle assessment (LCA) framework is employed together with consumer expenditure data to assess the CFs. According to the study, small urban households have the highest CFs in California. Their transport related emissions are significantly lower than those of the residents of less urbanized areas, but higher emissions from other consumption categories, together with the low degree of sharing of goods, overweight the gains. Two functional units, per capita and per household, are used to analyze the CFs and to demonstrate the importance of household size. The lifestyle impacts visible through the consumption data are also discussed. The study suggests that there are still significant gaps in our understanding of the premises of low-carbon human settlements.

Keywords—Carbon footprint, life cycle assessment, consumption, lifestyle, household size, economies-of-scale.

I. INTRODUCTION

Climate is warming [1]. Significant reductions of anthropogenic GHG emissions are needed rapidly to mitigate the warming to an adaptable level [1]. At the same time, rapid urbanization is taking place around the globe with the share of urban residents already exceeding 50% of the global population. As a consequence, reducing the GHGs caused by urban settlements has quickly become one of the critical issues around the globe.

GHGs are often considered to negatively correlate with urban density [2]-[4], and thus increased density has been given a central role among urban development targets. However, two features in the dominant assessment approaches lead to this result: (1) Utilization of Kyoto Protocol type of production-based approaches, which allocate the emissions based on the location of the emissions source, and (2) Including only the emissions from transport and housing energy. Especially in developed countries the production-based approaches often lead to blaming the less urbanized areas for high GHG emissions, since the most urbanized areas tend to outsource the emissions by producing little and concentrating on consumption of imported goods [5]-[6]. Also, when looking at the emissions from very limited sector scope, e.g. just from transportation, an inherent assumption is made that nothing else but travel behavior changes in different urban structures. However, literature, e.g. [7]-[10], depicts that the lifestyles are likely to be quite different through all consumption patterns when housing types and the surrounding service structures change. Elevated consumption of services, more frequent flying and less intra-household sharing have been shown to potentially outweigh the gains from reduced driving in denser settlements [7], [9], [11], [12]. Limited scope and geographically restricted production-based assessments might lead to biased estimations.

Another type of assessment approach adopts a consumer responsibility perspective and allocates the emissions of all utilized goods to the consumer [13], [14]. The consumption-based approach allows for analyzing how the environmental impacts caused by the residents vary with lifestyles across space. Consumption-based assessments are not meant to replace, but complement the production-based approaches by offering a way to analyze the emissions based on the demand and use of goods and services of consumers. This perspective is especially important with regard to GHGs, since a key feature of them, distinguishing them from the majority of other pollutants, is that their global warming impact is independent of the location of the emission.

It is evident that higher density reduces private driving [3], [14] and might also reduce housing energy requirements [15], [16]. However, these two consumption categories are not the only ones the urban structure affects. The surrounding structure with varying opportunities for consumption affects the lifestyles of the residents and thus the emissions they cause through all consumption categories [7], [17]. As the result, it is largely the affluence of the residents and the consumption opportunities they have which predict the CFs, not the density of an area [9], [10], [18]. Heinonen and Junnila [19] also showed that the lifestyles may differ so strongly along with the housing mode that even the connection between high density and low housing energy consumption might not hold. Furthermore, Ottelin et al. [12] showed evidence on how flying

Jukka Heinonen is with the University of Iceland, Reykjavik, Iceland (phone: +354-823-0064; e-mail: heinonen@hi.is).
can act as a substitute for driving and how the reduced GHGs from private driving in denser structures might be fully compensated by increased emissions from flying. Thus, more research is needed on who or which types of households actually cause the emissions and how they could be efficiently reduced.

An additional important, but often not recognized perspective is the impact of household size and the economies-of-scale effect. Bradbury et al. [20] actually suggest the diminishing household sizes to pose one of the biggest threats to sustainable utilization of the globe. This phenomenon is especially strongly related to urbanization and affects the distribution of CFs significantly. E.g. Heinonen et al. [8] depict how the high-rise residents have equally high or higher CFs in Finland than the residents of low-rise areas at the same level of affluence strongly due to the economies-of-scale advantage of the low-rise households, and Ala-Mantila et al. [11] demonstrate how the intra-household sharing increases towards less dense areas and compensate for the increasing GHGs from housing and transport.

In this study, we assess and analyze the geography of CFs in California paying close attention to the household size differences and the resulting economies-of-scale advantages and disadvantages. We employ an input-output (IO) analysis based hybrid LCA framework to assess the consumption-based CFs in California. We utilize US Consumer Expenditure Survey data of the Bureau of Labor Statistics [21], the EIO-LCA tables of Carnegie-Mellon University [22] and enhance the assessment with several external data sources. We depict how it is actually the small urban households that have the highest CFs in California due to high consumption power and disadvantage in economies-of-scale in comparison to larger households. We utilize two functional units, per capita and per household, to analyze the CFs and to demonstrate how important the household size is. We also discuss the lifestyle impacts visible through the consumption data. From the method perspective, we employ newer and more disaggregated data than the authors of previous studies from the US [23]-[25]. We also propose an improved way to include the emissions from the manufacture of certain durable goods into the CFs.

The remainder of the paper proceeds as: In Section II a literature review on consumption-based CF assessments is presented, Section III describes the method and data, the results are presented in Section IV and in Section V the results are discussed and positioned with regard to the previous studies, and the main uncertainties of the study are presented.

II. BACKGROUND

Connecting lifestyles to the environmental burden is not a new idea. Already in 1989, Schipper et al. [26] studied the relationship between income and energy consumption through different lifestyles. Consumption-based methods have since then become an established approach to assess the environmental impacts of entities from neighborhoods to nations [13]. The methods are able to capture the so called "outsourced emissions", that is, polluting industries concentrating to different locations than consumption. Several studies have shown that a significant share of the impacts associated with consumption can occur outside the place of consumption [27]-[30]. Regarding the United States, this foreign impact has been reported to vary from 11% to 29% [25], [30], [31].

Sub-national carbon footprinting has recently become a widely studied topic within this field of research. One key direction of this research is the city level, for which different approaches have been developed [5], [18], [32]-[36]. The city level has been seen important since it is obvious that the smaller the unit under study, the higher the share of the outsourced emissions potentially is.

Consumption-based assessments allow also for analyses of the lifestyle impacts. One direction of this research has studied the claim that urban living is less carbon-intensive on a per-capita basis than suburban or rural living. It is this branch of research which has questioned the density-GHG-relationship. Affluence, rather than urban density, has actually been often shown to drive the carbon/energy footprints. This finding has been replicated in Canada [6], the United [25], the United Kingdom [18], Australia [10], the Netherlands [37], Finland [7] and nations overall [38].

A perspective still understudied inadequately is how the urban structure affects the lifestyles and the resulting GHG emissions [7]. Focusing only on transportation and/or housing energy, emissions strongly correlate with population density with the densest regions having the lowest emissions (e.g. [2]). There would thus seem to be a discrepancy between these studies and those having found the affluence to drive the emissions when the overall CFs are considered. It seems that while we might understand relatively well how transportation patterns change along with the urban structure, there is a significant gap in understanding how other consumption patterns are related to the surrounding structure and how the emissions change as the consumption patterns change.

III. DATA AND METHODOLOGY

A. Research Materials

The primary input data utilized in the study is the 2011 U.S. Consumer Expenditure Survey (CE), collected annually for the U.S. Bureau of Labor Statistics (BLS) by the U.S. Census Bureau [21]. The CE data are gathered through independent quarterly interview and weekly diary surveys of approximately 7,000 sample households across the U.S. Each household can participate the survey for one quarter or several quarters, and thus the yearly sample is higher than 7,000 sample households across the U.S. Each household can participate the survey for one quarter or several quarters, and thus the yearly sample is higher than 7,000. Regarding consumption categories, the survey covers over 700 categories of daily and durable goods and services. The interview part includes all other personal consumption except alimentation and certain other daily consumption goods, which are covered with the diary part. The classification follows the Universal Classification Code (UCC). Furthermore, a wide variety of background variables is provided to be used e.g. for sampling and descriptive purposes.

While consumption surveys describe very well the
consumption patterns of certain households, some weaknesses should be corrected for carbon footprinting purposes. (1) The rental payments often cover utility usage, without adjustments thus potentially leading to underestimation of the CFs of those paying these, especially of those living in apartment buildings (e.g. [33]). The CE data include information about the embedded payments, but not the amounts. We extracted these payments for electricity, heat, gas, water and trash payments using a rather robust assumption that the embedded payment is as large as the particular utility payment on average when paid separately. (2) The GHGs embedded in the durable goods paid through loans over long time periods are not reflected well in consumption surveys, especially those associated with building construction and car manufacturing. Their impacts on CFs are thus easily underestimated (e.g. [39]). We corrected for these biases by omitting the principal payments and down payments of housing and vehicle loans from the CE data, and assessed the emissions separately. Regarding construction, we assessed the emissions for a certain area using construction permit data for the reference year from the U.S. Census and California Construction Industry Research Board and the EIO-LCA residential construction sector from the 2002 producer price model adjusted to 2011-2012 dollars with the construction cost index. For vehicles, the approximated values for new vehicles purchases in the reference years were taken from the CE data to assess the GHGs from manufacture. For used vehicles only the retail emissions related to their sales were included.

B. Methodology

LCAs

The basic categorization of LCAs divides them into bottom-up and top-down approaches. In bottom-up approaches the emissions are assessed and allocated to the entity analyzed process by process. The method is in general considered accurate, but it is laborious and inescapably suffers from truncation error due to drawing of the system boundary [40]-[41]. The share of the emissions left outside of the boundary can be significant (e.g. [40], [42]).

In top-down approaches, called input-output (IO) LCAs, the emissions are assessed with environmentally extended tables describing the output in emissions related to a monetary transaction in one sector of an economy. In IO LCAs the number of processes included in the assessment is infinite (see e.g. [40], [43]) and thus it does not suffer from truncation error. IO LCAs do have inherent problems, however. The most important are aggregation error and homogeneity and linearity assumptions (e.g. [44]). Even in the most disaggregated IO models each sector actually comprises multiple sectors of an economy, and the aggregation arises since the GHG intensities between the comprised sectors vary and as the result the IO tables can describe wrongly the emissions associated with the assessed entity. The homogeneity and linearity assumptions mean the naive assumptions that all the goods within a sector would embody the same emissions per monetary unit and that expenditure would linearly correlate with emissions. Furthermore, IO tables are predominantly based on national economies, giving arise to an additional error source, an assumption that imports would be produced with domestic technology and carry the same emissions per monetary unit as domestic production. The tables are seldom updated continuously and thus the technological improvements may not be well reflected.

Hybrid approaches intend to combine the better qualities of the two. They can include varying shares of process data and IO tables usage, and based on the way the two approaches are combined, three approaches can be distinguished: tiered hybrid LCA, IO based hybrid LCA, and integrated hybrid LCA (e.g. [41]).

Employed Tiered Hybrid Model

We employ here a streamlined hybrid LCA (Crawford, 2011) taking only the GHGs into account. Primarily the tiered hybrid method is used, replacing the first tier emissions of the IO table with local and more recent data. The EE IO model utilized is the Carnegie Mellon University Green Design Institute’s 2002 U.S. Benchmark Purchaser Price Economic Input-Output Life-Cycle Assessment (EIO-LCA) model [22]. The EIO-LCA is among the most disaggregated models available with 428 economic sectors. The classification follows the North American Industry Classification System (NAICS). Being a purchaser price model, the emission outputs are adjusted to the final end-user market prices.

The IO sectors amended with process data are housing energy and private driving fuel combustion. Regarding electricity production, the emission factors for the full life cycle were taken from Horvath and Stokes [45] and the price data from U.S. Energy Information Administration (EIA) [46], leading to ~560 g/kWh for California. For natural gas, the emissions from combustion were added to the production and supply emissions of EIO-LCA according to the factors provided by the U.S. Environmental Protection Agency (EPA) [47] and the prices provided by EIA [48]. For gasoline and diesel combustion in private driving the combustion emissions were added similarly according to EPA [49] and EIA [50] on top of the rest-of-the-life cycle emissions of EIO-LCA.

Finally, we use two functional units. The primary functional unit is per capita, supported by per household analyses to depict the impact of household size on all shareable goods.

C. Analyzed Samples

The CE data include approximately 725 households from California in each quarterly interview sample of 2011, and 200 households in diary data. To analyze the aimed perspectives on the CFs in California, we conducted the assessments on two levels. In the first stage, the households were divided according to the population size of their home settlement into Town, City and Metropolitan samples. In the second phase the Metropolitan sample was disaggregated according to the household type with a two-category breakdown of ‘Adult Households’ (AH) and ‘Families with Children’ (FC). Table I presents the samples and certain qualities related to each sample.
TABLE I

CHARACTERISTICS OF THE STUDIED AVERAGE HOUSEHOLDS

<table>
<thead>
<tr>
<th>Settlement size</th>
<th>Town</th>
<th>City</th>
<th>Metropolitan</th>
</tr>
</thead>
<tbody>
<tr>
<td>125,000-329,900</td>
<td>1.20-4 million</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.4</td>
<td>2.7</td>
<td>2.8</td>
<td></td>
</tr>
<tr>
<td>37,000</td>
<td>49,000</td>
<td>54,000</td>
<td></td>
</tr>
<tr>
<td>1.9</td>
<td>1.7</td>
<td>1.7</td>
<td></td>
</tr>
</tbody>
</table>

D. Research Process

The research proceeded in the following five steps:
1. Extracting the consumption profiles for the analyzed samples.
2. Selection of proper EIO-LCA sector to describe each consumption category.
3. Inflation adjustment of the EIO-LCA sectors.
4. Conducting the assessments.
5. Analysis and interpretation of the results.

(1) The CE microdata including all the responses is given in quarterly samples. Regarding interview data, each quarterly sample contains the interviews conducted during that quarter but gathering data from the previous three months. We employed 2011 Q2-2102 Q1 samples to generate a 12-month dataset. The 2012 Q1 sample includes data from 2012, but the overall result is a 12-month period predominantly describing the year 2011. The diary data are given in four quarterly samples for the year 2011. (2) 138 EIO-LCA sectors were found to match with the CE consumption categories, meaning thus that same sectors were used for several consumption categories. Still, 138 sectors is very high among the previous CF studies. (3) The 2002 EIO-LCA GHG intensities were adjusted to 2011-2012 prices according to the BLS Consumer Price Indices (CPI) for All Urban Consumers (CPI-U). The CPI follows a very similar sectoral breakdown to CE, and thus a sectoral instead of more aggregated adjusting was possible, unlike done in the majority of previous CF studies. (4-5) The results were calculated and the 138 sectors were aggregated into the following eight consumption categories for presentation and interpretation purposes:
1. Housing energy
2. Other housing
3. Private driving
4. Vacations
5. Food
6. Leisure goods and services
7. Education
8. Health care.

IV. RESULTS

A. CFs According to the Size of the Settlement

The functional unit selection makes a significant difference to the results. On a per capita basis the CFs in California were found to substantially increase along with the settlement size. Town residents have lowest average per capita CFs at 8.7 tons of CO2e/a (tCO2e/a), followed by Cities with 10.7 tCO2e/a and Metropolitans with 11.9 tCO2e/a. However, the household sizes decrease towards the larger settlements, small adult households concentrating towards the bigger centers. Thus, using per household as the functional unit leads to quite a different outcome; city households have the lowest CFs at 28.9 tCO2e/a, followed by Towns with 29.9 tCO2e/a and metropolitans with 33.5 tCO2e/a. The FCs according to the two functional units are depicted in Fig. 1.

Fig. 1 The annual average FCs in California in the Town, City and Metropolitan samples on per household and per capita.
The data provide several explanations for the higher per capita CFs in the larger settlements. Firstly, the largest emissions category for all settlement types is private driving. Cars are shareable goods and even though Town households cause the highest emissions per household with their driving, on per capita basis the emissions are the smallest. Secondly, even though housing energy requirements are affected by the geographic location quite heavily, on overall the housing-related emissions are another category where the economies-of-scale impact from household size is strong. As the result, the housing-related emissions are the smallest in Towns and the highest in Metropolitans on per capita basis. Thirdly, the consumption tends to increase towards the larger settlements due to increasing disposable income and the concurrent improvement in the diversity of consumption opportunities (see further discussion from e.g. [7]). In the case of California, Metropolitans end up causing the highest per capita emissions in all the eight categories, leading to significantly higher overall CFs.

B. Metropolitan Adult Households and Families

The highest CFs in California were found from the largest metropolitan areas. Within this group, a comparison between the adult households (Adults) and families with children (Families) reveals how important the economies-of-scale effect can be and how difficult low carbon living is for the small adult households with high consumption power. Fig. 2 depicts how dramatically the functional unit affects the results. On per household level Adults have CFs of approximately 27.0 tCO2e/a and Families 43.6 tCO2e/a, but per capita Adults’ CFs of 19.5 tCO2e/a exceed significantly the CFs of 11.4 tCO2e/a of Families. Furthermore, the GHG emissions caused by Adults exceed those caused by Families in every category on per capita basis. Housing, travelling in general and leisure goods and services are the categories where the differences are the highest. Thus the high CFs of Adults are a product of both, weak economies-of-scale with shareable goods, and consumption-intensive lifestyles.

Housing and private driving give good examples of the impact of the economies-of-scale effect. Housing energy and other housing account together for roughly 15 tCO2e/a for an average Family and only 10 tCO2e/a for an average Adult household. However, on per capita the average Adult still causes emissions of 7.3 tCO2e/a whereas in Families the caused emissions per capita drop to 3.9 tCO2e/a. The situation is very similar in Private driving category. On per household the emissions of Families are double to those of Adults at 14.6 tCO2e/a and 7.0 tCO2e/a, but per capita Families cause 3.8 tCO2e/a whereas Adults 5.1 tCO2e/a.

The GHGs from vacations and from the consumption of leisure goods and services, give some indication about different lifestyles of the two household types on average. The GHGs from these two add nearly 4 tCO2e/a/capita into the CF of an average Adult, whereas only 1.8 tCO2e/a/capita in Families group. Again an important explanatory factor is the consumption power, which is significantly higher per capita in Adults group even though as a household the average Family spends 50% more than an Adult household.

V. DISCUSSION

A. Interpretation of Results

This paper was set to study the geography of urban CFs in the context of California. In the study we brought up an often overlooked question of how the household sizes have a very important role in determining the CFs, namely in the sense that the small affluent adult households tend to have significantly higher CFs per capita than families with children due to the economies-of-scale effect. The economies-of-scale effect affects strongly all the shareable goods like housing and private driving, which form a large share of the CFs. In addition, an important share of all consumption, especially durable goods like many appliances, housewares and furnishings, could be categorized as household goods with the number of households affecting their demand more strongly than the number of people per household. It is thus extremely difficult for a small
household to live low-carbon life due to the emissions of all shareable goods allocating a small number of people. The issue is also very complex, since the rapid population growth is the primary reason for the intolerably high environmental burdens caused by humans. On the other hand, the trend in several developed countries is that the household sizes decrease and the number of adult households increases. This phenomenon is also tied to urbanization and if continued, it may cause a significant environmental threat far into the future even if the population growth would settle down [20].

In the paper we showed first how the per capita CFs are the highest in the largest metropolitan areas in California, 11.9 tCO2e/a/capita compared to 10.7 in Cities and 8.7 in Towns. In the second stage, we looked further into the Metropolitan CFs and found Adults to have significantly higher CFs in comparison to Families, 19.5 tCO2e/a/capita vs. 11.4 tCO2e/a/capita. The differences arise primarily from the most important shareable goods, housing and private transportation. On per capita basis the emissions from private driving as well as from housing are the highest for Metropolitan and the lowest for Towns in the settlement size samples, and higher for Adults than for Families within the Metropolitan sample. This even though Families possess more cars (see Table I) and drive significant mileages, and they cause the highest GHGs from private driving on per household. The situation is very similar with housing; Families have much larger apartments and much higher energy consumption per household, but per capita the economies-of-scale advantage reduces the emission well below those caused by Adults.

The data indicated certain lifestyle differences as well. Especially the emissions caused by vacations, particularly from aviation, and from the consumption of leisure goods and services first in Metropolitan sample and then Adults sample significantly exceeded those in the comparison samples. Heinonen et al. [7] depicted similar pattern to exist in Finland, and call this “parallel consumption” meaning the increased use of service spaces to compensate for the loss in possessed living space in denser settlements. Especially regarding vacations this might be a sign of the so called compensation effect, the residents of the more urbanized areas searching for takeaways from the everyday life more eagerly than the residents of less urbanized areas, as found to be the pattern in Finland by Strandell & Hall [51].

From the method and assessment techniques perspectives, our study contains some advanced features, especially related to the emissions associated to the production of the longest lasting durable goods of buildings and vehicles. Regarding these, the expenditure data from consumption surveys does not provide sufficient basis to assess the emissions due to the values being hidden in loans. Our method of assessing the construction phase emissions (including the materials) separately and including emissions only from new construction allows assessing these emissions similarly to all other emissions. In this study these emissions have a rather minor role, but the role of construction could change dramatically when smaller actively developed areas are analyzed (see e.g. [39]). The case is somewhat similar with private vehicles. In this study we omitted the down payments and the principal payments and assessed the emissions from manufacturing only for new vehicle acquisitions based on their values. Our approach is advanced also in the level of detail, since we use as many as 138 EIO-LCA sectors which allows more detailed input data as well in comparison to earlier U.S. based assessments [23]-[25]. The 2011-2012 CE data are also more recent and can be claimed to reflect better the actual current lifestyles. Finally, our approach accommodates the substantial sectoral inflation variation between the year of the EIO-LCA model, 2002, and the CE data, 2011, since we converted the expenditures into 2002 dollars sector by sector.

Our findings follow quite closely those of several earlier studies with somewhat similar settings. Heinonen et al. [7], [8] present very similar findings from Finland in that they report the highest CFs being found from the largest settlements, and present affluence and small household sizes as the explaining factors. Furthermore, Wiedenhofer et al. [10] depict how the energy footprints in Australia are the highest for the urban residents, and Minx et al. [18] find that the geography of affluence predicts the best the CFs, not for example density.

B. Limitations and Uncertainties

LCAs always contain limitations and uncertainties. In a CF study like this, these can be put into two main categories: the method related and the data related. The method related are the general weaknesses of IO LCAs. While EE IO LCAs remain the method most commonly used for carbon footprinting [13], the results should be interpreted understanding the method’s limitations. First, it is possible that the products purchased in different categories vary between the samples, making the homogeneity and linearity assumptions (see Section 3) not to hold. In general the larger the samples, the higher the probability the average purchased products to be similar, since the average consumer purchases exactly the average products. However, in our case it is possible that this method limitation amplifies the differences between the compared samples. E.g. Girod & de Haan [52] found in their comparative study that in Switzerland the quantity increases along with income, but not linearly, IO approach thus somewhat overestimating the impact of higher monetary consumption on higher income levels (although it cannot be easily estimated if the more expensive products also cause higher emissions than the market average products or not, or if the connection is the opposite). Second, we utilize the U.S. economy based IO model, but in California the average products may not comply with country averages. Hybridization of the model reduces this problem, but only regarding the hybridized sectors. Our approach also includes an inherent assumption of domestic production of imports, which is a weakness and could be corrected for in the future by utilizing a multi-region model. Higher housing prices in larger settlements, on their part, should not cause a bias in our study due to the omitting of the principal payments and disaggregation of the rental payments (see Section III). Third, the aggregation error is inevitably present in our assessment. While the amount of sectors, 428, in the EIO-LCA 2002 is high for an IO model, each sector still comprises many actual
industry sectors. Aggregation error is also a true error in its nature, meaning that the EIO-LCA sectors may randomly under- or overestimate the emissions related to any single product. Finally, CF studies commonly use attributional LCA which only depicts the situation at the time of the assessment. When interpreting the results and drawing policy implications this should be noticed. An important development step would be to amend the analyses with consequential LCA which would enable assessments of the impacts of system development.

The employed data give arise to another types of uncertainties. Especially the reliability of the data regarding the less frequently purchased goods is questionable. To reduce this uncertainty, we predominantly avoided using the highest level of disaggregation of the expenditure data. Second, the sample sizes decrease rapidly when sampling variables are added, which reduces the representatibility of the samples. To check the robustness of our samples we used two simple checkups: (1) that the income level in CE data corresponds with the American Community Survey (ACS), and (2) that the average household size corresponds with ACS. Regarding (1) we found that the CE sample for California might be slightly downwards biased in consumption power, but not significantly.

ACKNOWLEDGMENT
We thank The Academy of Finland, Grant 286747, for enabling the study.

REFERENCES


