UNIVERSITY OF Southampton

Southampton Statistical Sciences Research Institute

Working Paper A13/02

Applications and Policy

Expenditure as proxy for UK household emissions? Comparing three estimation methods

Milena Büchs and Sylke V. Schnepf

Abstract

Due to a lack of emissions data at the household level, studies examining the relationship between UK household CO_2 emissions and household characteristics currently rely on expenditure surveys to estimate emissions. There are several possible methods available for doing so but so far there is no discussion in the literature about the advantages and disadvantages related to these options. Such a comparison is relevant because studies in this area often draw policy-relevant conclusions.

To address this gap, this paper compares three different methods of estimation to discuss two questions: first, is it at all necessary to convert household expenditure into emissions, given that household expenditure and emissions are strongly correlated, and does research that takes this approach add anything to the insights that already exist in the extensive literature on the determinants of household expenditure? Second, if we assume that it is necessary to convert household expenditure into emissions, are more detailed (and time-consuming) methods of doing so superior to less detailed approaches? The analysis is based on expenditure data from the UK Living Costs and Food Survey 2008-9 and its predecessor the Expenditure and Food Survey 2006-7.

Expendituare as proxy for UK household emissions? Comparing three estimation methods[1]

Milena Büchs and Sylke V. Schnepf

Dr Milena Büchs is a Senior Lecturer in Sociology and Social Policy at the University of Southampton, UK.
Dr Sylke V. Schnepf is a Lecturer in Social Statistics at the University of Southampton, UK.

Abstract

Due to a lack of emissions data at the household level, studies examining the relationship between UK household CO_2 emissions and household characteristics currently rely on expenditure surveys to estimate emissions. There are several possible methods available for doing so but so far there is no discussion in the literature about the advantages and disadvantages related to these options. Such a comparison is relevant because studies in this area often draw policy-relevant conclusions.

To address this gap, this paper compares three different methods of estimation to discuss two questions: first, is it at all necessary to convert household expenditure into emissions, given that household expenditure and emissions are strongly correlated, and does research that takes this approach add anything to the insights that already exist in the extensive literature on the determinants of household expenditure? Second, if we assume that it is necessary to convert household expenditure into emissions, are more detailed (and time-consuming) methods of doing so superior to less detailed approaches? The analysis is based on expenditure data from the UK Living Costs and Food Survey 2008-9 and its predecessor the Expenditure and Food Survey 2006-7.

Corresponding author:
Dr Milena Buchs
Faculty of Social and Human Sciences
University of Southampton
Southampton
SO17 1BJ
UK
m.buechs@soton.ac.uk

[1] Acknowledgements: We would like to thank Ben Anderson and Nicholas Bardsley for comments on this paper and Frederico De Luca for assistance in preparing the dataset. This research was funded by ESRC grant RES-000-22-4083 “Who emits most - An analysis of UK households' CO2 emissions and their association with socio-economic factors”.

1. Introduction

From a social science and policy perspective it is important to understand how the distribution of CO2 emissions is related to household characteristics as it provides us with insights into potential distributional implications of climate change mitigation policies. However, research on household CO_2 emissions faces a range of challenges. To start with, there is currently no representative dataset available that combines emissions at the household level with household characteristics. Studies in this area have thus relied on expenditure surveys to estimate household emissions. All of the following studies on UK household emissions have utilised UK expenditure data: Baiocchi et al. (2010), DEFRA (2008), Dresner and Ekins (2006), Druckman and Jackon (2008), Druckman and Jackson (2009), Fahmy et al. (2011) and Gough et al. (2011). Exceptions are two studies on UK household transport emissions (Brand and Boardman, 2008; Brand and Preston, 2010) that are based on travel surveys in Oxfordshire. A range of studies on household emissions in other countries are also based on expenditure datasets, including Burney (1995), Cohen et al. (2005), Girod and De Haan (2010), Herendeen and Tanaka (1976), Herendeen et al. (1981), Kerkhof et al. (2009), Larivière and Lafrance (1999), Lenzen et al. (2006), O'Neill and Chen (2002), Pachauri (2004), Reinders et al. (2003), Weber and Perrels (2000), Weber and Matthews (2008), Wier et al. (2001) (also see Table A 1 in the appendix).

Using survey expenditure data for estimating household emissions is limited in several ways, including 1) issues around the relationship between expenditure, consumption and emissions (e.g. Baker et al., 1989; Girod and De Haan, 2010); 2) the quality of emission factors, particularly those derived from input-output analysis (e.g. Baiocchi et al., 2010; Druckman and Jackson, 2009); and 3) survey errors (e.g. Kerkhof et al., 2009). This implies that whilst expenditure based studies on household CO_2 emissions may be able to approximate total household emissions reasonably well, they are likely to be affected by these limitations when it comes to analysis at the household level, e.g. on the distribution of emissions across different types of households. Whilst these limitations cannot be easily overcome given the current lack of data on CO_2 emissions at the household level, it is important to identify these limitations and to explore methodological choices that can address some of these limitations to some extent.

This paper focusses on the first of the limitations mentioned above – related to the potential mis-matches between expenditure and consumption/emissions – and examines whether and if so how different, expenditure-based methods of estimating household emissions influence results. First, we ask whether a conversion of household expenditure into emissions adds anything to existing analysis of associations between household expenditure and socio-economic background, given that emissions and expenditure are correlated (e.g. Reinders et al., 2003; Weber and Matthews, 2008) [2]?

[2] Also see Figure A 1 in the appendix which shows a non-parametric regression of household emissions on household expenditure, indicating that both almost linearly related to each other at the household level.

This question is explored by comparing a model of emissions that applies a single conversion factor to all expenditure categories with two other models that apply more detailed emission factors. Second, if a conversion of expenditure to emissions is deemed necessary, do methods that exploit more information from the survey and match them to more detailed external data generate significantly different results compared to those employing simpler methods of converting expenditure to emissions? Specifically, we will focus on how the three estimation methods influence various estimates, such as total, mean and median emissions, measures of variance and inequality, and results from multivariate analysis such as OLS regression (beta coefficients and their significance, R2). Since many of these measures, in particular measures of inequality such as the gini coefficient and results from OLS regression are used to draw conclusions regarding potential distributional implications of emission reduction policies, comparing different methods of estimation provides important insights.

The following section discusses further details regarding the limitations of estimating household CO_2 emissions based on expenditure data and explains the three options of estimating household emissions that this paper aims to compare. Section three describes the data, estimation methods, and methods of analysis. Section four reports the results, section five discusses them and section six concludes. The appendix provides further details on the methods applied to estimate emissions.

2. Background

Expenditure-based studies on household CO_2 emissions are limited in several ways. Whilst a large number of such studies exist, several do not discuss these limitations or only mention them briefly without examining their implications. A brief review of these limitations is required to identify some of the general bounds within which this study operates – whilst this study aims to address the first limitation to some extent and to examine implications of doing so, many of the limitations discussed remain.

1) *Issues around the relationship between expenditure, consumption and emissions*: Since CO_2 emissions arise from the fossil fuel based carbon content of consumption, we would need a precise account of a households' consumption to estimate their emissions. However, expenditure datasets only provide us with a "best estimate" of actual consumption. For example, products and services are differently priced. Therefore, the same expenditure by two households on one type of product, for example electricity or clothing, can relate to different levels of consumption if one household subscribes to a cheaper electricity tariff (Baker and Micklewright, 1987) or bought less expensive clothes than the other. For instance, it is assumed that low income households often pay more per unit of domestic energy than high income households, because they are more likely to be on pre-payment meters and less likely to pay by direct debit which is the cheapest way of paying for domestic energy. This may lead to an overestimation of low in-

come households' domestic energy or emissions if this is based on expenditure and not corrected for the price of energy (Kerkhof et al., 2009: 1516).

Conversely, high income households might tend to purchase more expensive goods and services than poorer households. Several authors have examined how this might affect the estimation of high income households' emissions. Whilst Girod and De Haan's study on Swiss households (2010) concludes that whilst emissions were still rising with income, about half of the income elasticity was attributable to the purchase of more *expensive*, rather than *more*, products. However, Vringer and Blok (1997) only found a much small reduction of income elasticity of emissions of between 3-7 per cent for a study on Dutch households. Due to a lack of more precise product-related emission data (see below), neither of these studies could examine in detail the impact of potentially different energy intensities of more expensive or cheaper products.

Furthermore, households may not consume everything that they bought during the survey period or they might consume from stocks during this period and thus not record an expenditure. This problem is also known as the infrequency of purchase problem (e.g. Baker et al., 1989; Deaton and Irish, 1984; Tiffin and Arnoult, 2010). It particularly affects expenditures that are collected through diaries held over short periods. Does the infrequency of purchase problem influence analysis of household emissions? All previous studies using expenditure data for estimating CO_2 emissions implicitly or explicitly (DEFRA, 2008: 13) assume that *mean* CO_2 estimates derived from random sample expenditure surveys will be unbiased since zero expenditures for infrequently purchased items should be compensated by recorded purchases as those who do buy these items during the diary period may not fully consume them within this period. If large sample sizes are used there is no obvious reason to believe that this assumption does not hold, including for sub-groups within the sample. However, *measures of dispersion* such as standard deviation and variance are likely to be overestimated. OLS regression results can also be affected: given that the measurement error affects the dependent variable, standard errors of coefficients are likely to be inflated (for further details see Bardsley et al., forthcoming).

2) *The quality of emission factors, in particular those based on input-output analysis*: Even if we had more precise information on actual consumption, errors can occur in converting consumption to emissions due to limited quality of emission conversion factors. This conversion is relatively straightforward in relation to direct energy use, e.g. of gas or oil for heating or petrol and diesel for car travel, because reliable estimates of the average carbon content per volume of these fossil fuels are available. Conversion factors for electricity may already be less reliable because assumptions about the fuel mix for electricity generation within a certain period and region and losses within the grid have to be made. However, the greatest challenges occur in relation to conversion factors for emissions that are embedded in goods and services consumed. The most precise estimates could be achieved through product by product life cycle analysis that seeks to account for

fossil fuel inputs at every stage of production and consumption, including disposal (Hertwich, 2005). Since product-by-product life cycle analysis is extremely time and cost intensive, it has so far only been used to assess the carbon content of a small number of consumer items and it would be unfeasible to apply it to all the individual products and services consumed in one country. Therefore, studies in this area largely rely on environmental input-output analysis to attribute energy flows and the related carbon content to material and monetary flows between sectors within the economy (Reinders et al., 2003; Wier et al., 2001). Because it is possible to apply this to production and trade flows across national borders, input-output analysis can be used to estimate emissions related to the import of goods and services from abroad – and thus to estimate a country's emissions based on a consumption rather than production basis. Necessarily, input-output analysis can only provide estimates for average emissions of broad consumption categories because every firm and every sub-sector within certain industries operates in different ways. Input-output models are also often affected by time lags and inaccuracies of trade data and have to make simplifying assumptions about the production conditions in different countries (Baiocchi et al., 2010; Druckman and Jackson, 2009; Leontief, 1970). Gough et al. (2011) and Paul et al. (2010) discuss the input-output analysis applied by the Resources and Energy Analysis (REAP) Programme which is utilised in this paper.[3] An examination of the ways in which different input-output models influence results of research on the association between household emissions and socio-economic characteristics would be an important addition to the literature on individual input-output models but goes beyond the scope of this paper.

3) *Survey error*: The expenditure data collected through the survey may deviate from households' true expenditure due to survey errors. Survey errors that can occur at any stage of the survey, for example during survey design, collection of data and data processing (Groves et al., 2009). Survey error will differ for different consumption categories. For regular payments, LCF/EFS respondents are generally asked to provide bills (e.g. energy bills). We therefore assume survey error to be relatively small for electricity and gas payments. However, it is more difficult to estimate the size of survey error for two-week diary data (relevant for many transport related items and those that are included in "indirect" emissions). Any survey error will also affect CO_2 estimates. Given that the annual expenditure survey we use undergoes continuous quality checks, it is likely that survey error is small compared to errors introduced from 1) and 2).

Given that the literature has already addressed limitations of input-output analysis in studies on household emissions (Baiocchi et al., 2010; Druckman and Jackson, 2009; Gough et al., 2011) and

[3] The REAP database is based on a two-region input-output framework that models the embedded energy inputs and emissions for all the products and services consumed in the UK in 2006. It covers products and services from 178 sectors, based on the Standard Industrial Classification (SIC) System (Paul et al., 2010).

that end users can do little to address survey error 3), this paper focusses on issues discussed under 1). It concentrates on the question to what extent and in which ways different methods of converting expenditure into emissions affect distributional analysis of household CO_2 emissions in the UK within the bounds that we discussed above. In particular, we will focus on the question of how the level of detail employed in converting expenditure to emissions and ways of addressing the infrequency of purchase problem influence results:

- The first method uses a single emission conversion factor for all consumption categories; this is motivated by the question whether it is at all necessary to convert household expenditure into emissions, given that household expenditure and emissions are strongly correlated (e.g. Lenzen et al., 2006; Weber and Matthews, 2008). Would, therefore, research that converts expenditure into emissions add anything to insights provided by the extensive literature on the determinants of household expenditure (ONS, 2001-2011)? Previous research has shown that different consumption categories such as home energy or clothing have different energy or carbon contents per pound expenditure (e.g. Vringer and Blok, 1995: appendix), we would thus expect that conversion from expenditure to emissions does make a difference;
- The second method applies different conversion factors based on input-output analysis employed in the REAP programme to 57 consumption categories;
- The third, more detailed method, applies only 49 of the 57 input-output based conversion factors and uses various external data sources to estimate units of consumption more precisely for the remaining categories. It also employs a different measure of flight emissions that reduces the infrequency of purchase problem for flights to some extent. Since this method addresses some of the limitations discussed under 1) we would expect it to generate significantly different mean emission estimates, different measures of inequality and significantly different beta coefficients in multivariate regression. We would also expect the regression models based on this method to account better for variation in emissions.

Some of these methods, particularly the second, have been used in previous studies, for example, Gough et al. (2011) apply the same emission factors that we apply in the second method to estimate UK households' greenhouse gas emissions. Most of the studies mentioned above derive emission factors from input-output analysis that are then applied to household expenditures (Baiocchi et al., 2010; Cohen et al., 2005; Herendeen et al., 1981; Kerkhof et al., 2009; Lenzen et al., 2006; Reinders et al., 2003; Weber and Matthews, 2008; Wier et al., 2001), similar to the second method. Some studies have applied price data to estimate direct household CO_2 emissions but do not include indirect emissions (DEFRA, 2008; Druckman and Jackson, 2008; Fahmy et al., 2011) and one study has combined price data to estimate direct emissions with emission factors derived from input-output analysis to estimate indirect emissions (Druckman and Jackson, 2009), similar to our third method with remaining differences for estimating emissions from flights and public transport.

However, these methods have not yet been compared in the literature and both methods have not been compared to the first, 'single factor' method that simply scales up expenditure by one emission factor. From a policy perspective, it is important to examine how these different methods compare when applied in studies that examine inequalities of the distribution of emissions and associations between household characteristics and emissions. Due to a lack of alternative data on household CO_2 emissions, we can of course not examine how far or close estimates from any of these methods get to the actual amount, composition and distribution of household emissions, so we need to bear in mind that all of the approaches examined in this paper still only generate estimates of households' actual emissions.

3. Data sources and conversion methods

Each of the three methods of estimating UK household emissions is based on UK household expenditure data sourced the Expenditure and Food Survey (EFS) 2006-7 and the Living Costs and Food Survey (LCF) 2008-9.[4] The LCF/EFS is an annual survey, covering information on expenditure and socio-economic status of a representative sample of around 6,000 households per year (sample sizes vary by year). By combining 4 years, sample size is increased to 24,446 households. This also has the advantage that results are less influenced by economic circumstances during just one year. Expenditure data are collected on a large number of consumer items and services, through surveys at the household and individual level as well as expenditure diaries. The household survey is completed by the household representative (the person who pays for the mortgage/rent, or, if this is paid jointly, the person with the higher income) and covers a range of infrequently purchased goods and services as well as socio-economic characteristics of the household. In addition, each member of the household has to complete a survey on their income, benefits, taxes and other individual-level information. Each adult in the household has to keep an expenditure diary for two weeks whilst children aged 7-15 keep a simplified version of the expenditure diary. The LCF/EFS provides estimates of weekly household expenditure which we convert into annual expenditure to estimate annual household emissions.

To explore whether it is at all necessary to convert expenditure into emissions we created a 'single factor' emissions estimate for the first method by multiplying household expenditure by a constant CO_2 per £ expenditure factor. The factor is created by dividing total UK household emissions in 2006 from the input-output based Resources and Energy Analysis Programme (REAP) database by

[4] The LCF/EFS was introduced in 2001, combining the Family Expenditure and National Food Survey which had been conducted since 1957. The LCF replaced the EFS in 2008.

total UK household expenditure in 2006. This factor is applied to household expenditure for 2007-2009 corrected for inflation. The majority of REAP categories exactly match the expenditure categories in the LCF/EFS as both are based on the *Classification of Individual Consumption According to Purpose* (COICOP). In the appendix we discuss how we addressed divergences between both datasets.

The second method, which we denote the 'reap' method, uses emissions data from the REAP database to calculate CO_2 per £ expenditure factors for a range of different consumption categories, rather than just one single factor as in the method above. The REAP database provides us with annual UK household emissions in 2006 for 57 consumption categories, including home energy, motor fuels and other transport categories (see appendix for further details). We use this information to create CO_2/£ conversion factors for 57 consumption category in the LCF/EFS. The factors are created by dividing total UK household emissions for each of these categories (taken from REAP) by the total UK annual household expenditure for these categories (based on the LCF/EFS). The emissions factor *fi* for expenditure category *i* is given by:

$$f_i = \frac{CO2_i}{Exp_i} \tag{1}$$

where $CO2_i$ represents the annual UK household CO_2 emissions for consumption category i and Exp_i the total annual UK household expenditure for consumption category i. This is repeated for 57 different consumption categories – in contrast to the 'single factor' method where the same approach is employed just once, using total emissions and expenditure from all consumption categories combined.

The emission factors for each consumption category are then multiplied by households' expenditure for each of these categories and then summed to a household's total emissions. Since the LCF/EFS provides us with data on weekly spend, we multiply them by 52 to estimate a household's annual expenditure. That is, the CO_2 emissions for household j, $co2_j$, are estimated by:

$$co2_j = \sum_{i=1}^{m} exp_{ji}\ f_i \tag{2}$$

where exp_{ji} is the annual expenditure on consumption category i of household j, and i=1,…,m are the 57 COICOP consumption categories. Total annual UK household emissions can then be estimated by summing up the annual emissions of every individual household in the sample.

Since the REAP emissions data refer to 2006, we only use 2006 expenditure to create the emissions factors. Household expenditure for 2007-9 is then corrected for inflation using data from the Consumer Price Index for each of the 57 consumption categories.

The third, 'mixed' method applies more detailed information for estimating emissions of certain spending categories compared to the 'reap' method. The REAP-based emissions factors are still applied to 49 consumption categories but home energy (including electricity), motor fuels, public transport and flights are treated separately because we can exploit more detailed information from the survey as well as external data to generate more precise estimates. For *home energy* and *motor fuels*, we use price data to convert expenditure into quantities consumed (kWh and litres). Price data for home energy and motor fuels are sourced from the Department for Energy and Climate Change (DECC) Quarterly Energy Prices, Sutherland tables and AA motor fuel statistics.[5] Tables 2.2.3 and 2.3.3 of Quarterly Energy Prices provide annual domestic electricity and gas prices per kWh, including standing charge and VAT, for three payment methods (direct debit, credit and prepayment) and each electricity/gas region. Quarterly Energy Prices table 4.1.1 provides average monthly heating oil prices for the UK. Sutherland tables provide bi-annual prices for bottled gas, coal and wood for five different regions in the UK[6]. Price data per month and government region for petrol and diesel are sourced from AA statistics. DECC conversion factors were used to convert units of home energy and motor fuels into CO_2 emissions (DECC and DEFRA, 2011).

Using price data for home energy and motor fuels in the 'mixed' method means that we can account for regional and time variation of prices, as well as price differences between petrol and diesel (whilst we only have one figure for motor fuel emissions in the REAP database). In addition, this method takes into account that unit prices for home energy differ by payment method as data are provided for direct debit, credit (bill) and prepayment methods. The appendix provides further details on the price matching procedure.

For *public transport*, annual passenger kilometres for train, tube, bus and coach travel in Great Britain and Northern Ireland are used to create km/£ expenditure factors by dividing total passenger kilometres by total UK expenditure. Data on average annual passenger miles for train, tube, bus and coach journeys were provided by the National Travel Survey for Great Britain (table NTS0305) and the Northern Ireland Travel Survey (table 3.1). The travel factors can then be applied to household expenditure to estimate km travelled. DECC conversion factors per passenger kilometre are then applied to estimate emissions. Due to a lack of data, the 'reap' factors for ferries, road transport other than bus and coach journeys and "other transport" (e.g. cable cars and chairlifts) are also used in the "mixed method".

---

[5] See http://www.decc.gov.uk/en/content/cms/statistics/energy_stats/prices/prices.aspx for the most recent Quarterly Energy Prices tables and Oil and Petroleum price statistics; see http://www.theaa.com/motoring_advice/fuel/fuel-price-archive.html for the AA fuel price archive.

[6] See http://www.sutherlandtables.co.uk/.

For *flight emissions*, the 'mixed' method does not use expenditure data but information from the LCF/EFS interview questions on the number of flights the household purchased in the last year and the number of household members who were covered by the ticket.[7] Whilst the survey does not record exact flight destinations, it distinguishes UK, EU and non-EU flights. Information on average flight distances for flights from the UK to each of these broad regions is used to estimate average flight length. For flights within the UK we assume a mean return flight length of 1285 km, based on long distance journey data from the National Travel Survey 2006-9. According to the IPS 2006-2009, the average distance to destinations within the EU (but outside of the UK) and outside of the EU for private flights was 3,121 km and 16,502 km respectively. DECC conversion factors for flights were applied to flight kilometres to estimate emissions, including a factor of 1.09 proposed by DECC/DEFRA (2011) to account for additional distance flown during rise, cruise and descent.

Based on this method, the same estimate of average emissions is applied to each individual flight within each of these three areas of destination. Whilst this does not account for variability of emissions for exact flight destinations, this is the only possibility that currently exists for estimating flight emissions based on the LCF/EFS survey information. We argue that this method is preferable to using the information on flight expenditure from the two-week diary because a) prices for flight tickets vary considerably for similar distances depending on the airline and time of booking and b) the diary window only captures flight expenditure of 1.2 per cent households in the sample (whilst 41 per cent of households had at least one flight in the previous year according to the survey) and is thus highly affected by the Infrequency of Purchase Problem. This paper will examine how these different assumptions influence the estimation of household emissions and their distribution.

Since the REAP database does not provide emissions for package holidays but includes them in other categories, the 'single factor' and 'reap' methods reallocate expenditure on "package holidays" to expenditure on holiday accommodation, public transport categories and flights to account for variability in household spending (see appendix for details). Since expenditure on package holidays is collected through the household survey covering the last three months, the re-allocation of package holiday spending to other travel categories increases the per cent of households with flight emissions in the 'single factor' and 'reap' methods to 14 per cent (from 1.2 per cent). Thus, it already minimises the infrequency of purchase problem for these areas. In the 'mixed' method, package holiday spending is only re-allocated to accommodation and public transport because package holiday flights are already captured in the household survey question on the number of flights which we employ in this method.

[7] However, we only know the number of household members included in the flight for EU and non-EU flights, not for UK flights. Here we have imputed the number of household members covered by the ticket using information from international flights. Every UK flight is treated as a return flight.

3.1 Methods of analysis

Relationships between household CO_2 emission estimates generated by the three methods described above are examined in a first step using scatter plots and correlations. Scatter plots provide a visual insight into how estimates from one estimation method relate to another. The scatter plots include zero emissions which is particularly important for comparing flight emission estimates because the 'mixed' method uses number of flights per year from the survey whilst the other two methods are based on flight expenditure, resulting in a much higher proportion of 'zero' flight emissions. The scatter plots are restricted to 80 tonnes CO_2 for total, 40 tonnes for transport and 10 tonnes for flight emissions to exclude outliers and ease comparability between the different estimates. This excludes less than 1% of the top emissions for each of the plots (see notes for details). They are also set to square shapes so that points on the 45° line indicate households for which both methods generated equal estimates. In the next step we present correlation coefficients to quantify the extent to which emission estimates are linearly associated. Zero emissions are excluded from all correlations to focus on how estimated positive values compare.

Mean and median comparison tests follow to examine whether the 'mixed' estimation method generates significantly different estimates compared to the 'reap' and 'single factor' methods for *average* emissions in the four emission domains. Standard deviations, coefficients of variation, 90/10 ratios and gini coefficients are also shown to examine whether the different methods of estimating emissions have effects on measures of variance and inequality. This is important for research on the inequality of emissions or underlying energy and resources – an area of the literature which occasionally applies the gini coefficient to measure resource or emissions inequality (e.g. Groot, 2010; Jacobson et al., 2005; Papathanasopoulou and Jackson, 2009). The gini coefficients measure the inequality of the emission distributions, based on Lorenz curves. Since the gini coefficient is sensitive to outliers, the 1^{st} and 99^{th} percentile of the distribution are excluded.

The following section examines how emission estimation methods influence analyses of associations between household characteristics and emissions. This is important because a range of papers examine these relationships based on different methods of estimation (e.g. Baiocchi et al., 2010; DEFRA, 2008; Gough et al., 2011; Kerkhof et al., 2009; Lenzen et al., 2006; Weber and Matthews, 2008; Wier et al., 2001). Comparing results from different methods of estimation directly will therefore provide important insights. First, we test whether mean and median emissions from 'mixed', 'reap' and 'single factor' methods differ significantly for different household groups.

Since many household characteristics are related to each other, it is also important to examine associations for individual factors whilst controlling for all other factors. For example, income and education or income and rural location are closely related in the UK. Are high education and rural location still significantly associated with household emissions after income is controlled for? Questions like these can be examined by applying OLS regression as several of the papers above have done, including Gough et al. (2011), DEFRA (2008), Baiocchi et al. (2010), Weber and Mathews (2008) and Lenzen et al. (2006). Therefore, we also test whether the 'mixed', 'reap' and 'single factor' methods result in significantly different beta coefficients from multivariate OLS regressions.

Complex survey design (clustering in sampling units and weights) is taken into account throughout.

4. Results

4.1 Comparing estimation methods – correlations and summary statistics

How do the emission estimates derived from the three methods relate to each other? Do the three methods of estimation provide us with significantly different estimates of mean and median emissions?

First of all we can calculate the proportion of households for whom any two of the three estimation methods generate the same emission estimates (for total emissions and any sub-category of emissions). Results showed that none of the estimates generated by the 'single factor' method is equal to estimates based on the 'mixed' and 'reap' methods (for total emissions and all sub-categories). 'Mixed' and 'reap' estimates also do not overlap for home energy, flight and motor fuel emissions. However, the 'mixed' and 'reap' methods generated a small proportion of equal estimates for total, transport and public transport emissions. This is because emissions from package holidays (apart from those allocated to flights), ferries, road transport other than bus and coach travel and 'other transport' are based on REAP factors in both approaches. The resulting overlap is small for total and transport emissions (1.6 and 0.3 per cent respectively) but 17 per cent for public transport emissions. For indirect emissions, the 'mixed' and 'reap' methods apply the same REAP emission factor to all but one sub-category. The category that is estimated differently are indirect emissions from home energy, resulting in different estimates for total indirect emissions for the majority of households. However, 18 per cent of households do not have an expenditure on gas (mostly because they do not have access to mains gas), they therefore also have equal values for indirect emissions in the 'mixed' and 'reap' approach.[8]

Scatter plots provide a useful further step for examining the relation between the different estimates visually. The first plot (figure 1) compares total CO_2 emission estimates from the 'mixed' and 'reap' methods, showing a strong linear association. The second plot which compares 'mixed'

[8] The variable for indirect home energy emissions differs in the two approaches because the REAP database combines the emissions for electricity and those for indirect emissions from gas, oil and other home energy fuels in one category. Therefore we generate per £ emission factors for electricity and home energy in the 'reap' method by dividing the total REAP emissions for that category by the summed up expenditure for electricity and all other fuels respectively. We cannot use that same CO2 per £ expenditure factor for indirect home energy emissions in the 'mixed' method because we would count electricity emissions twice (as we already estimate them using price data – the denominator for that factor would thus be too large). We therefore subtract the summed up amount of electricity emissions from the REAP figure for indirect home energy emissions and use this to calculate a separate CO_2 per £ emissions factor for the mixed method.

and ‘reap’ transport emissions (figure 2) demonstrates that estimates are still related but less closely than for total emissions. We can examine public transport and flight emission plots separately for further insights. The scatter plot that compares ‘mixed’ and ‘reap’ public transport emission estimates (figure 3) has a fan-like pattern with some estimates that are equivalent in both methods (on the 45° line) and others that systematically differ because different emission factors are used for sub-categories such as train/tube, road transport and ‘combined tickets’ travel. The scatter plot for ‘mixed’ and ‘reap’ flight estimates (figure 4) shows that estimates are not linearly associated. This is due to the different approaches applied to estimate flight emissions – one is expenditure based whilst the other employs information on the number of flights per year. Therefore, a range of households have an expenditure on flights in the two-week period but do not record a flight for the last year in the household survey and vice versa (the horizontal and vertical lines of dots at zero emissions for each category). The “step” shape of the scatter plot results from the fact that the ‘mixed’ method applies the same emissions factor within each of the three flight destination regions per flight. Due to space constraints only these four plots are presented. Other ‘mixed’ and ‘reap’ plots showed strong linear associations between estimates.

Table 1 provides correlation coefficients for each combination of the different estimates (apart from flights which are not linearly associated). In contrast to the scatter plots, zero emissions are excluded to focus on differences of positive emission estimates. The correlation matrix in table 1 confirms that the mixed and ‘reap’ estimates are highly correlated for most categories with Pearson’s r of 0.93 for total, 1.00 for indirect, 0.97 for home energy, and 0.99 for motor fuels and public transport emissions. As we would expect, Pearson’s r is considerably lower for total transport emission estimates with 0.65. Despite the high correlation of ‘reap’ and ‘mixed’ home energy and motor fuel emission estimates, none of the estimates is equal for a single household as explained above.

The correlations between the ‘mixed’ and ‘single factor’ method estimates are generally lower than for the ‘mixed’ and ‘reap’ methods, with r = 0.86 for total, 0.90 for indirect, 0.95 for home energy and 0.72 for public transport emissions – but it is higher for total transport with r=0.70 and the same for motor fuels. Correlation coefficients of the ‘reap’ and ‘single factor’ estimates are generally similar to those from the ‘mixed’ and ‘single factor’ correlations, but slightly lower for total emissions with r=0.83 and higher for total transport with r = 0.96.

However, since two variables that have the same distribution but differ by a factor are still perfectly correlated, it is crucial to examine whether or not estimates of mean and median emissions generated by the three estimation methods significantly differ. Table 2 shows total, mean and median household CO_2 emission estimates. Mean comparison tests show that all ‘single factor’ mean estimates are significantly different from the mixed estimates at the 1 per cent level. Whilst they are relatively close in size to the estimates from the other two methods for *total* mean and median CO_2

emissions, but almost twice as high for indirect and 6 or 2.5 times lower for home energy and transport emissions respectively. The ‘reap’ mean estimates are also significantly different from the ‘mixed’ estimates for total, indirect, motor fuels, flight and transport emissions (at 1 per cent or 5 per cent level, see table 2**Error! Reference source not found.**).

Estimated median emissions are considerably lower than mean emissions for all categories, indicating a positive skew of the emissions distribution. The estimates for total median emissions derived from the three estimation methods are not significantly different (using $p < 0.05$ as a threshhold). However, for all other emission domains, ‘single factor’ median estimates are significantly different from ‘mixed’ estimates with $p < 0.01$. ‘Reap’ median emission estimates are also significantly different from ‘mixed’ estimates for indirect, motor fuel and transport emissions. The mean/median ratio is relatively similar for the different methods of estimation with an average of 1.3. However, the ratio differs more for flights with 1.5 for the ‘mixed’ and 1.9 for the ‘reap’ method, indicating a higher skew for the latter (see Table 2).

We also examined whether the three methods of estimating household emissions influence measures of variance and inequality. The standard deviations based on ‘reap’ estimates are generally slightly higher compared to those of the ‘mixed’ estimates. Standard deviations based on the ‘single factor’ estimates are higher for indirect and total emissions but much lower for all other areas compared to those related to the ‘mixed ‘method. However, since the standard deviation depends on the unit of measurement, it is also important to compare the Coefficients of Variation (CV). This shows that ‘mixed’ estimates are more variable than ‘reap’ and ‘single factor’ estimates for most areas of emissions except for flight, and thus also for transport and total emissions.

Since the emission variables generated by the three methods show different levels of dispersion, we would also expect them to differ in relation to estimates of inequality such as the gini coefficient. However, gini coefficients generally only differ very marginally across the three methods with the exception of flights. The gini coefficient is very high with 0.93 for the ‘single factor’ and ‘reap’ estimates, even higher than the one associated with the ‘mixed’ estimate of 0.81.

4.2 The relationship between different emission estimates and household characteristics

Up to now, we only examined differences between emission estimates without taking household characteristics into account. However, most of the research employing household data is interested in the association between household characteristics and CO_2 emissions. The previous section demonstrates that the relationships between estimates for emission sub-categories such as home energy and transport are complex – whilst some estimates may be highly correlated they can relate to significantly different mean values and differences in variance. How do these similarities and differences play out when relationships between emissions and household characteristics are examined? How do results compare across different domains such as home energy, transport, indirect and total emissions? We examine these questions by first comparing mean and median values for the different types of emissions and for different groups of households as presented in table 4.

The variables are defined as follows: “high” and “low income” refers to households in the highest or lowest equivalised income quartile. “Age35” includes households with a reference person aged 35 and under, “age65” covers households with reference persons aged 65 and over. “>=16 years of education” includes households in which at least one person attended full time education for 16 years or more, and “<=11 years of education” identifies households in which none of the household members attended education for more than 11 years. “Rural” includes households in rural areas, defined by the LCF/EFS as settlements of under 10,000 inhabitants, and “urban” households those in all other areas. “Workless households” are defined as those that have at least one person of working age but no person of working age in employment, whilst “in employment” means that at least one household member of working age is in employment or self-employed. “Female” and “male head”’ means that the household reference person is female or male, and “white” and “ethnic minority” is defined by the household reference person’s self-declared ethnicity. “Children” and “no children” refers to households with/without children.

A comparison of mean values for total, indirect, home energy and transport emissions for different groups of households shows that most of the ‘single factor’ mean values are significantly different from the mean ‘mixed’ values at the 1 per cent level for all emission categories. ‘Single factor’ estimates are consistently higher for total and indirect emissions and consistently lower for home energy and transport emissions across most household groups. Many total and transport ‘reap’ estimates for different types of households are also significantly higher than the ‘mixed’ estimates. The only groups for which mean *total* emissions are not significantly different for all three methods of estimation are those for low income and workless households (see table 4 and table 5).

Figure 5 plots cumulative total emissions against total emissions for high and low income households and each estimation approach. This not only shows us how unequally emissions are distributed across different income groups but also that the three approaches generate more similar estimates for low income than for high income households. This is plausible because the differences in estimates multiply with higher incomes and thus higher emissions. Furthermore, the ‘single factor’ approach overestimates emission inequalities between low and high income households (represented by the outer dark grey lines).

As explained above, a range of papers apply OLS regression to examine relationships between household characteristics and emissions, conditional on other factors (e.g. Baiocchi et al., 2010; DEFRA, 2008; Gough et al., 2011; Lenzen et al., 2006; Weber and Matthews, 2008). Therefore it is important to investigate whether different methods of estimation impact on regression results, including ‘effect’ sizes, level of significance and overall performance of the model to account for variability in emissions. To this end, we run OLS regressions using variables from the three estimation methods with log-transformed total, indirect, home energy and transport emissions as dependent variables. Zero emissions and outliers for emissions and income (defined as the 1st and 99th percentile of the distribution) are excluded to reduce the influence of outliers.[9] Variables that differ from the ones described above for unconditional analysis are log income (based on disposable household income), dummy variables for the number of adults and children in the household, e.g. “adult2” is coded 1 for households with at least two adults and 0 otherwise, “adult3” is coded 1 for households with at least 3 adults and 0 otherwise. “Age” provides age in years. Since the relationship between age and emissions has an inverse u-shape, an age-squared term is also used (‘age2/100 – age squared divided by 100). The age variable is top coded at 80 in the LCF/EFS. Therefore, a dummy variable is included, coded 1 for households with reference persons aged 80 and over and 0 otherwise. “Edu 12-15” is coded 1 if at least one household member attended education for 12 to 15 years and 0 otherwise, “edu 11” as defined above is the reference category. “Rural missing” is coded 1 if information on rural location is missing which is mainly households in Northern Ireland.

Tables 6 and 7 and Tables A6 and A7 in the appendix show OLS regression results. The former two present results from the ‘full’ model that includes all the variables described above. The latter two tables present results from a restricted model that only includes income and household size. Beta coefficient comparison tests showed that overall, the ‘mixed’ and ‘reap’ approaches did not generate significantly different results for home energy and indirect emissions. Results are also similar for other emission areas. In the full model, coefficients for high education differ significantly between the ‘reap’ and ‘mixed’ approaches for total and transport emissions (see tables 6 and 7). In the restricted model, none of the ‘reap’ beta-coefficients significantly differs from the ‘mixed’ coefficients apart from that for income in the transport model (see

[9] Outliers from each dependent variable are excluded in all models to achieve equal sample sizes.

Table A6). However, all of the 'mixed' models result in slightly higher R-squares than the 'reap' models, particularly for total and transport emissions.

However, there are more differences between the 'mixed' and the 'single factor' regression results: a range of beta coefficients are significantly different due to the different underlying mean values for emissions in different areas (for home energy, this only applies to the extended model that includes housing and heating type, see Table A8 in the appendix). The 'single factor' models generate higher R-squares than the 'mixed' models for total, indirect and home energy emissions.

5. Discussion

The results above provide interesting insights, some of which unexpected. This section discusses how we can explain the findings and outlines limitations.

The scatter plots provided a first insight into relationships between estimates from the 'mixed' and 'reap' emissions, showing that the different methods of estimating flight emissions generated very different, uncorrelated estimates for individual households. This also translated into a weaker match of total transport emission estimates compared to estimates for total emissions. Correlations confirmed that estimates from all three methods are highly correlated for all other emission domains (excluding 'zero' values).

However, the three methods of estimation often generated significantly different mean emissions at the 1 or 5 per cent level. This is due to different emission factors being applied. In particular, the 'single factor' method generates considerably higher estimates for indirect and considerably lower estimates for home energy and transport emissions than the other two methods. 'Reap' based estimates were also significantly different from 'mixed' estimates for most domains except for motor fuels and public transport.

For *home energy*, it can be shown that the 'mixed' method can take some 'real world' variation of household expenditure into account that the other two methods miss. The application of price data by payment method in the 'mixed' approach accommodates for the fact that electricity and gas are more expensive per kWh for households on prepayment meters than for those who pay by direct debit or quarterly bill. As a result, home energy emissions for households who are on prepayment schemes for both electricity and gas are significantly lower than mean emissions for those who pay by credit or direct debit (for both electricity and gas) in the 'mixed' method whilst they are not significantly different in the 'reap' method (see table 3**Error! Reference source not found.**). Furthermore, unlike the 'reap' approach, 'mixed' estimates take regional and monthly price differences for petrol and diesel into account. As a result, mean CO_2 emissions for petrol are significantly different from mean diesel emissions in the 'mixed' approach, but not in the 'reap' approach (at the 1 per cent level, standard errors are calculated taking clustering of data in sampling units into account) (see table 3**Error! Reference source not found.**).

Whilst the three methods generated relatively similar measures of variance and inequality for most emission domains, those related to flight emissions differed much more (with impacts on measures for transport and total emissions). For several emission domains, ‘mixed’ estimates had a higher CV than ‘reap’ or ‘single factor’ estimates, including for indirect, home energy, motor fuel and public transport emissions. This can be explained by the fact that the ‘mixed’ approach applies more detailed emission factors, generating a more variable sample of estimates compared to the other two methods. First of all, the ‘mixed’ method applies different emission factors for six types of home energy fuels whereas the ‘reap’ method only uses two factors, one for electricity and one for all other types of home energy. The ‘mixed’ method also applied different emission factors for petrol and diesel where ‘reap’ only uses one. In addition, the ‘mixed’ approach exploits more differentiated price data as explained above.

However, for flight and transport emissions, ‘reap’ and ‘single factor’ estimates have a higher CV than the ‘mixed’ estimates, reflecting that the ‘mixed’ approach only employs three different “emission factors” depending on whether the flight was in the UK, within the EU (but outside of the UK) or outside of the EU. In contrast, the estimates for the ‘reap’ and ‘mixed’ estimates are based on expenditure where single values considerably vary around the mean. This is not compensated by the fact that we have a much larger share of households with zero flight emissions in the ‘reap’ and ‘single’ factor method than in the ‘mixed’ method (see above).

The different gini coefficients for flights result from the much lower proportion of estimated ‘zero’ flight emissions in the ‘mixed’ method – when zero emissions are excluded, the gini coefficient for flights is more similar for the three estimates with 0.54 for ‘mixed’ and 0.53 for the other two methods. We also have to assume that the gini coefficients for other emission categories are likely to be inflated due to the infrequency of purchase problem. This particularly affects emission subcategories which contain certain proportions of ‘false zero’ emissions, including home energy, motor fuel and public transport emissions. Yet, we can conclude that *within* these bounds, the choice of estimation method generally does not have significant implications for this highly policy-relevant estimate for emission domains other than flights and total transport.

However, if mean *total* emissions and cumulative distribution plots for different income groups are compared, it becomes clear that the ‘single factor’ method is likely to overestimate emission inequalities between contrasting household groups. This is likely to be due to the fact that the ‘single factor’ approach applies a much higher emissions factor to indirect emissions than the other two methods. Indirect emissions typically constitute a higher share of overall emissions for rich than for poorer households (54.3 per cent compared to 50.5 per cent) and they increase by 0.7 per cent for every 1 per cent increase in income (unconditional on other factors (for further details see Buchs and Schnepf, 2013). Differences in estimation are multiplied with rising emissions, thus leading to significantly different mean estimates for high emission households in particular.

A comparison of OLS regression results showed that, against our expectation, coefficients from the 'mixed' and 'reap' based regressions were not significantly different (apart from high education for total and transport emissions for which the 'mixed' approach showed larger effect sizes). However, 'mixed' method based regressions generate slightly higher R-squares than 'reap' based regressions. We expected this to be the case because the 'mixed' approach has a better coverage of flight emissions and takes more of the variability of emissions into account by matching domestic energy and motor fuel expenditure to more fine-grained price data.

A comparison of 'mixed' and 'single factor' regression results confirmed that a range of coefficients differed significantly. This provides further evidence that a 'single factor' approach would mis-specify the relationship between emissions and household characteristics based on OLS regressions – not only for sub-categories of emissions but also for total emissions.

However, the 'single factor' regressions resulted in a higher R-square for total and indirect emissions. At first sight, this questions the argument that the 'mixed' method provides superior emission estimates than the other methods as argued above. The 'single factor' based regression actually examines associations between household characteristics and expenditure since expenditure is simply scaled up by one emission factor. Therefore, the higher R-square for the 'single factor' method indicates that *expenditure* on goods and services is more closely related to the household characteristics than the *emissions* resulting from this spending. This is plausible, for instance, some high expenditures may relate to relatively low emissions per pound expenditure and vice versa (e.g. £100 spent on clothing has a lower carbon content than £100 spent on home energy in the 'mixed' and 'reap' models whilst they have the same emissions in the 'single factor' model) whilst characteristics like high household income may be related to higher expenditure more generally. In other words, we would expect household characteristics to predict expenditure better than emissions.

6. Conclusion

Research on household CO_2 emissions and their association with household characteristics is highly relevant to provide insights into potential distributional effects of climate policies. Since data on total CO_2 emissions arising from the whole range of consumption items is not currently available at the household level, this research relies on expenditure datasets to estimate CO_2 emissions. Whilst expenditure-based research on household CO_2 emissions faces a range of limitations, this paper focused on the question of whether, within these general bounds, the choice of estimation techniques makes a difference to estimating mean emissions for different emission domains as well as to results from multivariate analysis on household CO_2 emissions. More specifically, this paper examined two questions: 1) is it necessary to convert household expenditure into emissions, given that household expenditure and emissions are strongly correlated, and 2) do more detailed methods of converting household expenditure into CO_2 emissions (mainly addressing the first of the three limitations set out in section 2) generate significantly different results compared to simpler approaches? Since the values of “true” emissions per household are unknown, we cannot determine which of these estimates is closer to reality but we can test whether different methods generate significantly different estimates of mean household CO_2 emissions, measures of inequality, and coefficients and model performance of multivariate OLS regressions.

Regarding the first question we conclude that it is not sufficient to apply just one single emissions factor per pound expenditure because this method generates biased estimates of total and mean emissions for all emission sub-categories as it does not take variation of carbon intensities of different consumption categories into account. If applied in OLS regression, this method also generates significantly different beta-coefficients for a range of predictor variables, particularly for total, indirect and transport emissions. This demonstrates that estimating household emissions in a more detailed fashion does not simply replicate results from OLS regressions on expenditure and that household characteristics relate in different ways to expenditure than to emissions. Studies in this area thus make an important contribution to knowledge about the distribution of emissions and their associations to household characteristics.

This leaves us with the question of the level of detail that should be applied to estimate household emissions based on expenditure data. We analysed this question by comparing the ‘mixed’ and ‘reap’ methods of estimating household emissions. Our results provide some evidence that the more detailed, ‘mixed’ method of estimating household emissions generates significantly different mean and median emission estimates, apart from those for home energy and public transport. This is also true when we compare total and transport mean and median emissions for different household groups (see tables 4 and 5**Error! Reference source not found.**) which is of policy relevance as one or the other method is likely to under- or over-estimate emissions for specific groups. Whilst the dispersion of data (based on the coefficient of variation) is lower for the ‘mixed’ flight estimates and thus also for ‘mixed’ total emissions, ‘mixed’ estimates tended to be more dispersed for all other categories because it applies more differentiated emissions factors. Gini coefficients examining the overall inequality of the emissions distribution were fairly similar across the three methods of estimation but markedly lower for ‘mixed’ compared to ‘reap’ and ‘single factor’ flight emissions when zero emissions were included. Lower dispersion and measures of inequality for ‘mixed’ flight emissions result from an entirely different method of estimating flight emissions in the ‘mixed’ model that utilises the number of flights during the past 12 months rather than expenditure on flights during the two-week diary period used in the ‘reap’ (and ‘single factor’) method. The ‘mixed’ methods flight emission estimate thus addresses the infrequency of purchase problem (part of limitation 1)) and is thus likely to generate more robust estimates than the ‘reap’ method.

We also expected that the ‘mixed’ and ‘reap’ methods would perform differently in OLS regression because the ‘mixed’ method can take differences in home energy payment methods and regional fuel price differences into account and employs a more comprehensive measure of estimating flight emissions. Whilst almost none of the beta coefficients were significantly different, slightly higher R-squares for the ‘mixed’ method might provide some evidence to confirm our assumption.

Overall, we conclude that expenditure-based studies on household emissions need to be transparent about the method of emission estimation applied. The more detailed approaches of matching expenditure data with external data on prices, passenger kilometres and conversion factors are, the more precision we can expect for measures of spread and inequality. To generate more reliable estimates, further investments into more detailed household consumption surveys will be necessary, ideally providing not only data on expenditure but also on actual consumption, covering longer periods for at least part of the survey to address the infrequency of purchase problem which can have significant effects on results as shown above for flight emission estimates.

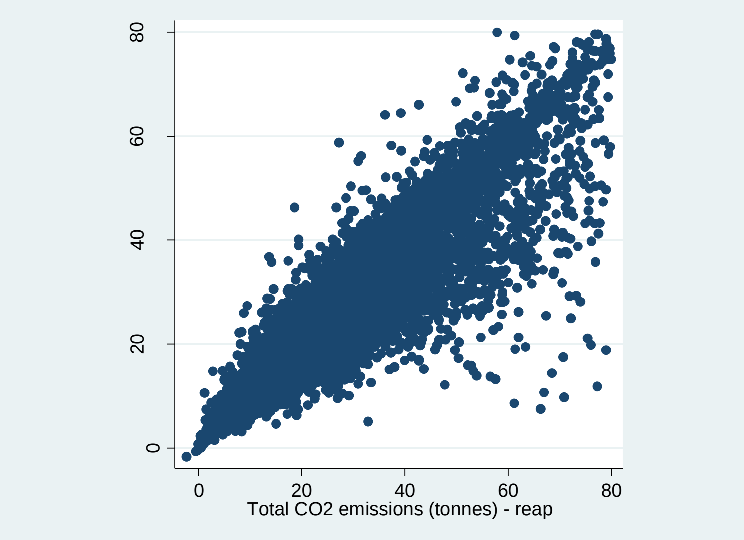
Tables and figures

Table 1: Correlation coefficients for the three estimates

	'Mixed'-'reap'	'Mixed'-'single factor'	'Reap'-'single factor'
total	0.93	0.86	0.83
indirect	1.00	0.90	0.90
home energy	0.97	0.95	0.96
transport	0.65	0.70	0.96
motor fuels	0.99	0.99	1.00
public transport	0.99	0.72	0.76

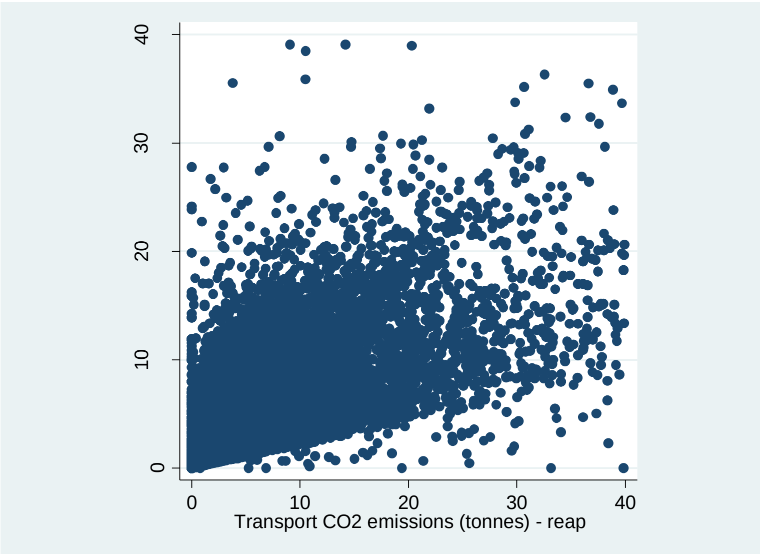
Note: Zero emissions were excluded for all estimates. Sample sizes were: 24446 for total and indirect emissions, 23105 for home energy emissions, 20840 for transport emissions, 15943 for motor fuel emission, 11885 for public transport emissions,

Figure 1: Scatterplot comparing 'mixed' and 'reap' total household CO_2 emissions



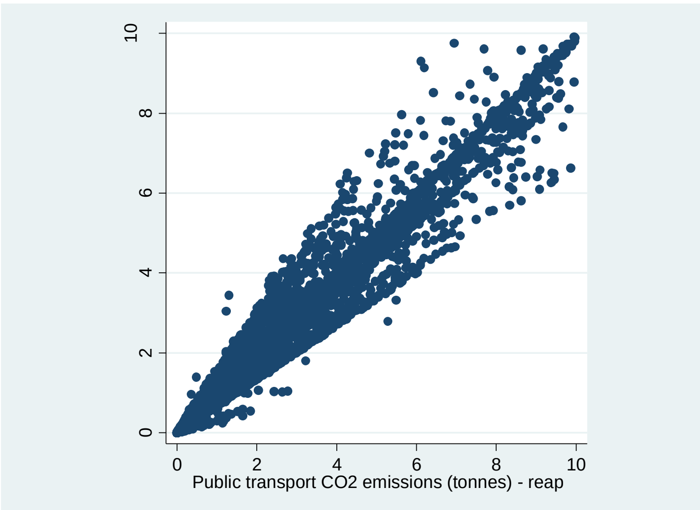
Note: Emissions over 80 tonnes per year are excluded for comparability. This excludes less than 1 per cent of the top emissions (99th percentile are 68 tonnes for 'mixed', and 79 for 'reap').

Figure 2: Scatterplot comparing 'mixed' and 'reap' transport CO_2 emission estimates



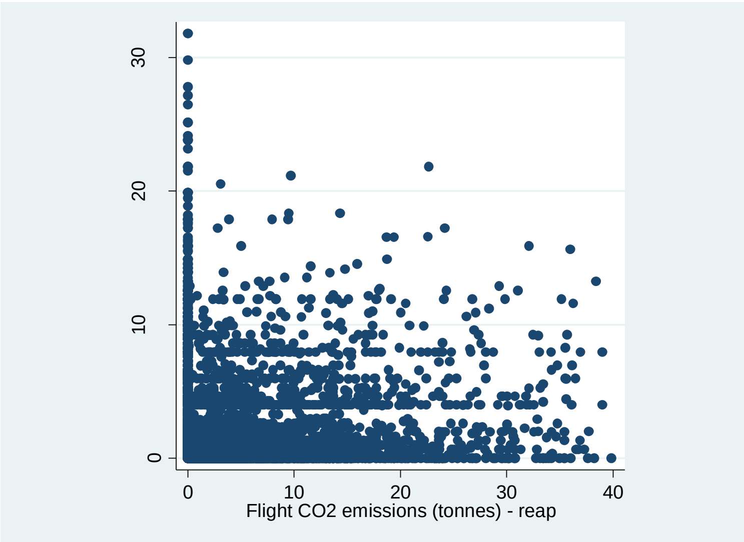
Note: Emissions over 40 tonnes a year are excluded for better comparability. This excludes less than 1 per cent of top transport emissions (99th percentile are 22 tonnes for 'mixed' transport and 39 for 'reap' transport).

Figure 3: Scatterplot comparing 'mixed' and 'reap' public transport CO_2 emissions



Note: Public transport emissions of over 10 tonnes a year are excluded. This excludes less than 1 per cent of top public transport emissions which are 9 tonnes for both estimates.

Figure 4: Scatterplot comparing 'mixed' and 'reap' flight CO_2 emissions



Note: Flight emissions over 40 tonnes a year are excluded – which excludes less than 1 per cent of top flight emissions. Flight emissions at the 99th percentile are estimated at 11 tonnes for 'mixed' and 27 tonnes for 'reap'.

Table 2: Total, average and variance estimates of annual CO_2 emissions by method and emission category

	Total UK CO2 (mio tonnes)	Median (tonnes)	Mean (tonnes)	Mean - median ratio	Standard deviation	CV	90/10 ratio	Gini coef-ficient
Total (mixed)	513	17.13	20.18	1.18	14.61	72.4	6.2	0.33
Total (reap)	538	17.52	21.17***	1.21	16.55	78.2	6.6	0.35
Total (single factor)	559	18.02	21.97***	1.22	17.76	80.8	7.3	0.36
Indirect (mixed)	271	8.69	10.67	1.23	9.25	86.7	6.8	0.35
Indirect (reap)	279	8.99***	10.98**	1.22	9.38	85.4	6.7	0.35
Indirect (single factor)	492	15.72***	19.33***	1.23	16.34	84.5	7.7	0.37
Home energy (mixed)	130	4.48	5.11	1.14	3.98	78.0	4.7	0.35
Home energy (reap)	132	4.53	5.17	1.14	4.01	77.4	4.9	0.36
Home energy (single factor)	22	0.77***	0.86***	1.12	0.61	70.2	4.2	0.33
Motor fuel (mixed)	61	1.60	2.38	1.49	2.94	123.8	6.4	0.59
Motor fuel (reap)	69	1.83***	2.70***	1.48	3.33	123.4	6.3	0.59
Motor fuel (single factor)	23	0.62***	0.92***	1.48	1.14	123.4	6.3	0.59
Public trans (mixed)	23	0.00	0.89	–	2.21	247.6	32.3	0.78
Public trans (reap)	24	0.00	0.93	–	2.28	244.0	30.3	0.78
Public trans (single factor)	12	0.00	0.48***	–	1.04	215.4	25.7	0.78
Air (mixed)	29	0.00	1.13	–	2.44	216.5	20.2	0.81
Air (reap)	35	0.00	1.38***	–	5.94	428.5	15.2	0.93
Air (single factor)	1	0.00	0.37***	–	1.59	428.5	15.2	0.93
Transport (mixed)	112	2.97	4.40	1.48	4.98	113.2	17.4	0.53
Transport (reap)	127	2.72***	5.02***	1.85	8.27	165.0	19.4	0.60
Transport (single factor)	45	1.04***	1.78***	1.71	2.51	141.2	13.8	0.57

Note: *** significantly different at 1 per cent level from 'mixed', ** significantly different at 5 per cent level from 'mixed'.
Standard error calculation takes clustering of data into account.
The 90/10 ratio for all transport categories only applies to those with non-zero emissions.

Table 3: Mean annual home energy and motor fuel emissions (CO_2 in tonnes) for different payment methods and types of fuel

	Mixed	Reap
Home energy		
Direct debit	5.69 (0.03)	5.56 (0.03)
Credit	5.10 (0.05)	5.47 (0.05)
Prepay	**4.73 (0.09)**	5.25 (0.11)
Motor fuels		
Petrol	3.30 (0.03)	3.83 (0.03)
Diesel	**3.76 (0.06)**	3.99 (0.07)

Notes: Home energy - zero expenditures/emissions are excluded. The payment methods apply to both electricity and gas payments. Standard errors in parentheses. Bold figure is significantly different from the other two mean values for 'mixed' method at 1 per cent level (when taking clustering into account).
Motor fuels - since many households have purchased both petrol and diesel, these figures are only referring to motor fuel emissions for those who have only bought petrol (n= 11,791) or diesel (n= 2,382) but not both (n= 1,746). Bold figure for diesel is significantly different from petrol at 1 per cent level, taking clustering into account.

Table 4: Mean and median total CO2 emissions for different household groups

	Mixed Mean	**Reap** Mean	**Single factor** Mean	**Mixed** Median	**Reap** Median	**Single factor** Median
High income	29.8	31.6***	35.3***	27.3	27.5	30.7***
Low income	12.0	12.5	11.7	10.3	10.6	9.1***
Age35	19.1	19.7	23.9***	17.1	17.3***	20.5***
Age65	14.0	15.0***	12.4***	12.1	12.5	9.9***
>=16 years education	27.7	28.6	32.4***	25.1	25.2	27.8***
<= 11 years education	15.7	16.7***	16.1	13.8	14.3	13.2***
Urban	19.2	20.1***	21.5***	16.5	16.8	17.4***
Rural	23.2	24.4**	23.7	19.9	20.3	19.5
Employed	21.2	22.3***	23.3***	18.7	19.0	19.4***
Workless	12.7	13.3	12.6	10.6	11.0	9.8**
Male head	22.3	23.4***	24.4***	19.8	20.2	20.3
Female head	16.8	17.6**	18.1***	14.3	14.7	14.4
White	20.3	21.3***	22.1***	17.6	18.1**	18.0
Ethnic minority	19.0	19.4	21.1**	16.5	15.7	17.1
No children	18.2	19.2***	19.1***	15.4	15.8**	15.0
Children	25.0	25.9	28.8***	22.7	22.9	25.2***

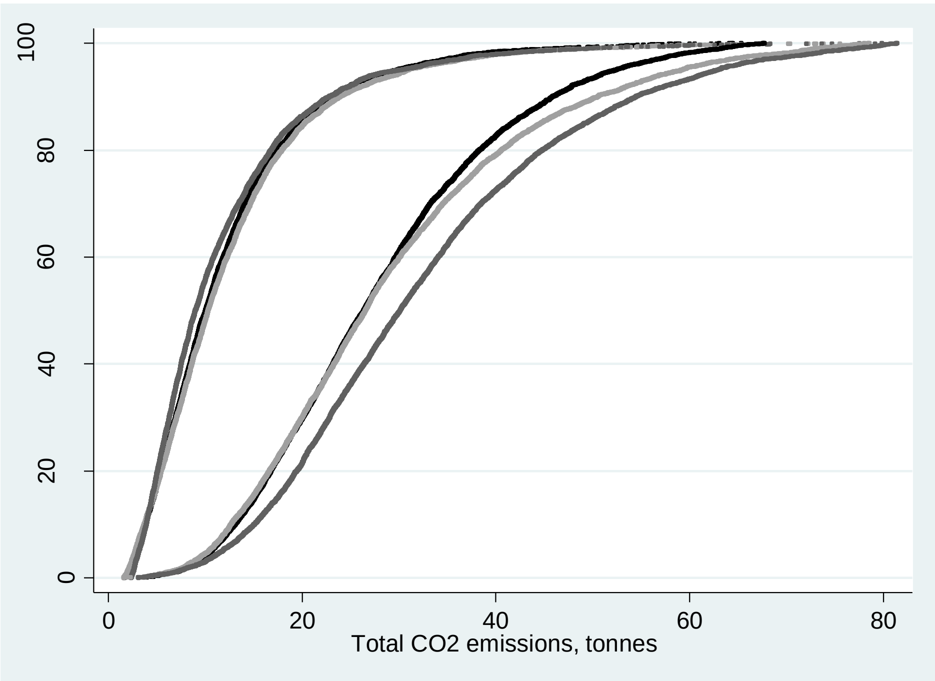
Note: *** indicates significant difference from 'mixed' estimate at 1 per cent level. ** indicates significant difference from 'mixed' estimate at 5 per cent level. Clustering is taking into account. Confidence interval for median uses binominal exact option.

Table 5: Mean and median transport CO_2 emissions for different household groups

	Mixed Mean	Reap Mean	Single factor Mean	Mixed Median	Reap Median	Single factor Median
High income	7.4	8.7***	3.0***	6.1	5.5***	2.1***
Low income	1.9	2.0	0.8***	0.8	0.7	0.3***
Age35	4.5	4.7	1.8***	3.2	2.8***	1.1***
Age65	2.2	2.9***	1.0***	1.3	1.4	0.5***
>=16 years education	6.8	7.3	2.7***	5.5	4.6***	1.8***
<= 11 years education	3.1	3.8***	1.3***	1.9	1.9	0.7***
Urban	4.2	4.8***	1.7***	2.8	2.5***	1.0***
Rural	5.1	5.9***	2.0***	3.8	3.5	1.3***
Employed	4.7	5.4***	1.9***	3.4	3.1***	1.1***
Workless	2.2	2.4	1.0***	0.9	0.7	0.4***
Male head	5.1	5.8***	2.0***	3.7	3.4***	1.3***
Female head	3.3	3.7***	1.4***	2.0	1.8***	0.7***
White	4.4	5.1***	1.8***	3.0	2.8**	1.1***
Ethnic minority	4.7	4.6	1.8***	3.2	2.2***	1.0***
No children	3.9	4.6***	1.6***	2.6	2.4**	0.9***
Children	5.5	6.0**	2.1***	4.1	3.7***	1.4***

Note: *** indicates significant difference from ‘mixed’ estimate at 1 per cent level. ** indicates significant difference from “mixed estimate at 5 per cent level. Clustering is taking into account. Confidence interval for median uses binominal exact option.

Figure 5: Cumulative versus total CO_2 emissions for high and low income households



Note: The three lines towards the left represent low income households (below or at the 25th income decile), the three lines to the right represent high income households (at or above the 75th income decile). The black lines are based on ‘mixed’ estimates, the light grey lines ‘reap’ estimates and the outer middle grey lines are based on the ‘single factor’ estimates.

Table 6: OLS regression total and transport CO_2 emissions, full model

	Total emissions			Transport emissions		
	'Mixed'	'Reap'	'Single factor'	'Mixed'	'Reap'	'Single factor'
Log income	0.400***	0.403***	**0.494*** **	0.601***	0.567***	**0.512*** **
	(0.008)	(0.008)	(0.008)	(0.017)	(0.019)	(0.017)
adult2	0.264***	0.269***	**0.222*** **	0.291***	0.305***	0.264***
	(0.009)	(0.010)	(0.009)	(0.021)	(0.024)	(0.021)
adult3	0.111***	0.115***	**0.080*** **	0.055**	*0.080****	***0.135**** ***
	(0.011)	(0.012)	(0.010)	(0.025)	(0.029)	(0.025)
adult4	0.077***	0.089***	0.055***	0.076	*0.141****	*0.140****
	(0.020)	(0.021)	(0.020)	(0.048)	(0.052)	(0.046)
adult5	0.110**	*0.083**	0.106**	0.015	-0.049	-0.043
	(0.045)	(0.050)	(0.044)	(0.094)	(0.107)	(0.093)
child1	0.094***	0.099***	0.078***	-0.114***	-0.091***	-0.086***
	(0.009)	(0.010)	(0.009)	(0.021)	(0.024)	(0.022)
child2	0.075***	0.071***	0.072***	0.043	0.043	0.021
	(0.011)	(0.012)	(0.011)	(0.027)	(0.029)	(0.026)
child3	0.055***	0.054***	***0.002***	-0.083**	-0.046	-0.021
	(0.015)	(0.016)	(0.015)	(0.039)	(0.042)	(0.037)
age	0.023***	0.022***	**0.017*** **	0.039***	0.038***	0.028***
	(0.002)	(0.002)	(0.002)	(0.004)	(0.004)	(0.004)
$age^2/100$	-0.021***	-0.021***	-0.022***	-0.040***	-0.038***	-0.030***
	(0.002)	(0.002)	(0.002)	(0.004)	(0.004)	(0.004)
Age > 80	-0.087***	-0.096***	-0.102***	-0.229***	-0.210***	-0.128***
	(0.018)	(0.019)	(0.018)	(0.050)	(0.055)	(0.048)
Edu 16+	0.111***	**0.074*** **	0.129***	0.227***	***0.062**** **	**0.113*** **
	(0.009)	(0.010)	(0.009)	(0.021)	(0.024)	(0.021)
Edu 12-15	0.079***	0.067***	0.092***	0.151***	0.109***	0.108***
	(0.008)	(0.009)	(0.008)	(0.019)	(0.021)	(0.019)
Edu missing	-0.040***	-0.050***	***0.008***	-0.066*	*-0.096***	***-0.073**** **
	(0.015)	(0.016)	(0.014)	(0.038)	(0.043)	(0.037)
Female reference person	0.015**	0.015**	*0.000*	-0.103***	-0.108***	-0.051***
	(0.007)	(0.008)	(0.007)	(0.016)	(0.019)	(0.017)
Workless	-0.017	-0.010	***-0.096**** ***	-0.241***	-0.277***	-0.173***
	(0.013)	(0.014)	(0.014)	(0.032)	(0.034)	(0.030)
Ethnic minority	-0.103***	-0.138***	**-0.150*** **	-0.002	***-0.226**** ***	***-0.122**** ***
	(0.015)	(0.017)	(0.016)	(0.033)	(0.038)	(0.032)
Rural	0.081***	0.081***	**0.045*** **	0.177***	0.212***	0.131***
	(0.008)	(0.009)	(0.008)	(0.017)	(0.019)	(0.017)
Rural missing	0.152***	0.179***	***0.015***	0.122***	0.191***	0.117***
	(0.013)	(0.013)	(0.011)	(0.023)	(0.027)	(0.024)
Constant	-0.438***	-0.431***	-0.587***	-3.634***	-3.440***	-3.751***
	(0.050)	(0.054)	(0.051)	(0.119)	(0.132)	(0.115)
Observations	21,658	21,658	21,658	19,133	19,133	19,133
R-squared	0.548	0.501	0.640	0.327	0.252	0.252

Note: Zero emissions, 1st and 99th percentile of all dependent emission variables and the income variable are excluded in the models. *** is significant at 1 per cent level, ** significant at 5 per cent level. Bold beta coefficients are significantly different from "mixed method" coefficients at least at 5 per cent level. Coefficients in italics have different level of significance or different sign compared to the 'mixed' model.

Table 7: OLS regression on indirect and home energy emissions, full model

	Indirect emissions			Home energy emissions		
	'Mixed'	'Reap'	'Single factor'	'Mixed'	'Reap'	'Single factor'
Log income	0.439***	0.432***	**0.504*****	0.166***	0.159***	0.146***
	(0.008)	(0.008)	(0.008)	(0.009)	(0.010)	(0.009)
adult2	0.271***	0.270***	**0.215*****	0.209***	0.220***	0.201***
	(0.009)	(0.009)	(0.009)	(0.012)	(0.012)	(0.011)
adult3	0.117***	0.116***	**0.071*****	0.110***	0.117***	0.127***
	(0.011)	(0.011)	(0.011)	(0.015)	(0.015)	(0.014)
adult4	0.071***	0.071***	*0.044***	0.062**	0.060**	0.062**
	(0.022)	(0.022)	(0.021)	(0.025)	(0.026)	(0.025)
adult5	0.115**	0.112**	*0.119****	0.117*	*0.120***	*0.133***
	(0.050)	(0.050)	(0.045)	(0.060)	(0.059)	(0.060)
child1	0.123***	0.125***	**0.088*****	0.172***	0.181***	0.162***
	(0.010)	(0.010)	(0.010)	(0.013)	(0.014)	(0.012)
child2	0.083***	0.083***	0.078***	0.087***	0.084***	0.081***
	(0.012)	(0.012)	(0.011)	(0.015)	(0.015)	(0.014)
child3	0.051***	0.053***	***-0.004***	0.099***	0.102***	0.101***
	(0.015)	(0.015)	(0.016)	(0.020)	(0.021)	(0.019)
age	0.020***	0.020***	0.016***	0.023***	0.023***	0.024***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
$age^2/100$	-0.019***	-0.019***	-0.022***	-0.016***	-0.016***	-0.018***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Age > 80	-0.126***	-0.120***	-0.100***	0.022	0.020	0.026
	(0.018)	(0.019)	(0.019)	(0.023)	(0.024)	(0.021)
Edu 16+	0.114***	0.112***	0.136***	0.027**	0.025**	0.026**
	(0.010)	(0.010)	(0.010)	(0.012)	(0.013)	(0.011)
Edu 12-15	0.088***	0.086***	0.097***	0.036***	0.030***	0.030***
	(0.008)	(0.008)	(0.008)	(0.010)	(0.011)	(0.010)
Edu missing	-0.040***	-0.040***	*0.012*	-0.030*	-0.027	***0.080******
	(0.015)	(0.015)	(0.015)	(0.018)	(0.019)	(0.016)
Female reference person	0.021***	0.023***	*0.002*	0.048***	0.053***	0.042***
	(0.007)	(0.007)	(0.007)	(0.009)	(0.009)	(0.008)
Workless	-0.022	-0.019	***-0.111******	0.039**	*0.055****	*0.059****
	(0.013)	(0.013)	(0.015)	(0.016)	(0.016)	(0.015)
Ethnic minority	-0.198***	-0.189***	-0.164***	-0.040**	*-0.017*	-0.036**
	(0.017)	(0.016)	(0.017)	(0.018)	(0.019)	(0.017)
Rural	0.064***	0.061***	**0.034*****	0.037***	*0.021**	0.064***
	(0.009)	(0.009)	(0.009)	(0.011)	(0.012)	(0.010)
Rural missing	0.141***	0.148***	***-0.012***	0.143***	0.178***	**0.196*****
	(0.012)	(0.012)	(0.011)	(0.021)	(0.022)	(0.017)
Constant	-1.276***	-1.204***	-0.737***	-0.503***	-0.467***	-2.112***
	(0.053)	(0.054)	(0.055)	(0.066)	(0.070)	(0.062)
Observations	21,658	21,658	21,658	21,658	21,658	21,658
R-squared	0.554	0.546	0.621	0.172	0.159	0.178

Note: See note for table 10.

Appendix

Table A 1: Overview of studies on the distribution of household emissions/energy requirement using expenditure data

UK

Study	**Data**	**Outcome variable / policy**	**Level**	**Area**	**Country/region**	**Analysis**	**Variables**	**Methods discussion ii**
Distribution of emissions / energy								
(Baiocchi et al., 2010)	ACORN/CACI dataset, EFS 2004, Input-output analysis	CO_2	Households by 56 ACORN types	Total CO2	UK	Descriptive and “pooled” OLS regression	Descriptive: 17 ACORN groups and by type of dwelling Regression: hh size, income, education, large houses, presence of children, pensioner, single pensioner, use of internet, social housing, NT membership	a, f
(DEFRA, 2008)	EFS 2003/4-2005/6; home energy and motor fuel price data	CO_2 emissions	Per adult (based on household data)	Direct emissions (sum of home energy and motor fuels)	UK	OLS regression	Various EFS variables	b
(Druckman and Jackson, 2008)	EFS 2004-5, Census 2001, home energy price data, emissions factors	CO_2	7 Output Area Classification groups; small areas	Household energy	UK	appendix: correlation with hh and per capita emissions	Income deciles, OAC groups (descriptive); income, hh composition (correlation)	e
(Druckman and	EFS, Census 2001, home	CO_2	7 Output Area	Total CO2, 9 or	UK	Descriptive	OAC groups	a

Jackson, 2009)	energy and motor fuel price data, input-output data (flights are expenditure based)		Classification groups	4 consumption categories				
(Fahmy et al., 2011)	EFS 2004-7, NTS, APS, EHCS; source of conversion factors unclear	CO_2	Households	Direct emissions (home energy, transport)	UK	Descriptive and ANOVA	Income, hh type, tenure, number of workers, empl status, age, socio-economic group, settlement type, number of cars, type of heating fuel, government region	c
(Girod and De Haan, 2010)	Swiss income and expenditure survey 2002-2005; input-output data; physical units combined with LCA data	CO_2e	Per capita	Total and split up into various consumption categories	Switzerland	No	n/a	d
(Gough et al., 2011)	EFS 2006, input-output data (REAP) (flights are expenditure based)	CO_2e	Per capita (based on household data)	Total, home energy, food, consumables, public and private services	UK	OLS regression (total and separately for the 5 consumption categories, sample sizes unclear.	Income, family type (composition and age), employment status	a, b
(Papathanasopoulou and Jackson, 2009)	Family Expenditure and Family Spending 1968-2000, input-output data	Mill. Tonnes of oil equivalent	Country/households	Total fossil resources	UK	Gini coefficients, descriptive	Income quintiles	d

Non-UK

Study	Data	Outcome variable	Level	Area	Country/region	Analysis	Variables	Methods discussion ii
Distribution of emissions / energy								
(Burney, 1995)	UN energy statistics (coun-	Electrici-	Per capita per	Electricity	World	OLS regression		

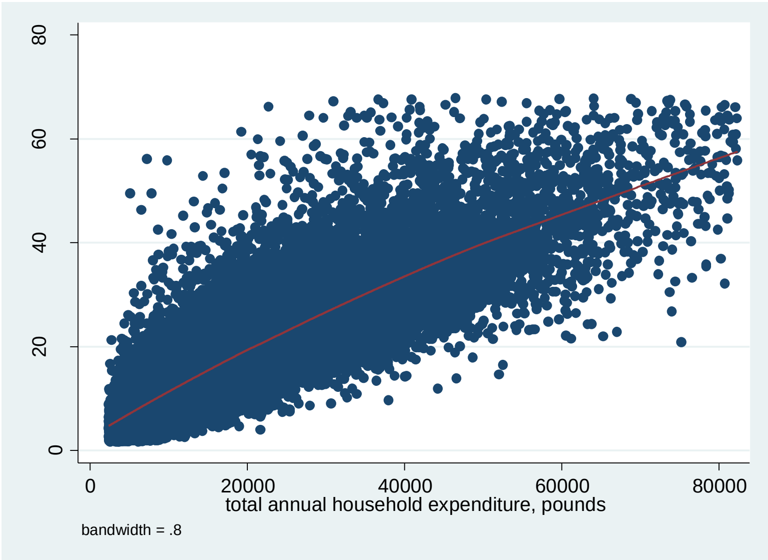
	try level)	ty (kWh)	country					
(Cohen et al., 2005)	IBGE household expenditure survey 1995-6, input-output data	Energy intensity (MJ/US$ PPP)	Households	total	11 capital cities in Brazil	Descriptive distribution. OLS regression of hh energy on expenditure and capital city.	Income groups (descriptive) Expenditure (regression)	n/a
(Duarte et al., 2010)	Household Budget Continuous Survey 1999	CO_2 emissions	Households and per capita	Split up for different types of fuels and total (including from consumption?)	Spain	Descriptive over income bands	Income	n/a
(Herendeen and Tanaka, 1976)	Consumer Expenditure Survey 1960-61; input-output analysis	Energy in Btu	Households	Total, some results for 11 consumption categories	US	n/a		a, c,
(Herendeen et al., 1981)	Bureau of Labor Statistics Consumption Survey 1972-3	Total energy from consumption in BTU	Households	Total, Direct	US	OLS regression	Hh size, hh expenditure, rural/urban, number of rooms	(b)
(Heinonen and Junnila, 2011)	Finnish Consumer Survey 2006	CO_2	Households	Total	Finland	Descriptive	Rural/urban	a
(Kerkhof et al., 2009)	Household expenditure datasets from 4 countries, input-output data	CO_2	Households	Total and 12 consumption categories	Sweden, Norway, Netherlands, UK	descriptive	Income groups	d, f
(Larivière and Lafrance, 1999)	Electricity consumption data, dataset with socio-demographic data	Electricity MWh	Per capita per city	Electricity	Canadian cities	OLS regression	Variables are at the city level! (density, average age, temperature, wealth per inhabitant	n/a
(Lenzen et al., 2006)	Expenditure surveys (different countries), input-output data	Energy intensity (MJ/$PPP)	Per capita	Total energy requirements	Australia, Brazil, Denmark, India and Japan	OLS regression (total)	Expenditure, hh size/type, urbanity, age, employment, education	(a)

(Lyons et al., 2012)	Household budget survey Ireland, Input output data	CO_2 and various other pollu-tants	Per person and different household types	Total	Ireland	Descriptive	Income, number of persons, ru-ral/urban, hh composition, number of disa-bled residents	n/a
(O'Neill and Chen, 2002)	Residential transport (and energy) consumption sur-veys	Energy (Btu)	Per capita	Residential and transport sepa-rately and com-bined	US	Descriptive and OLS regression (residential and transport sepa-rately) – but for indi-vidual independent vars separately	Hh size, age, presence of chil-dren	n/a
(Pachauri, 2004)	National Sample Survey 1993-4, input-output anal-ysis, data on physical di-rect energy consumption converted to MJ	Energy (MJ)	Households	Total, 9 con-sumption cate-gories	India	OLS regression (total)	HH income / ex-penditure most important. Size of dwelling, age, literacy and hh size, employment status, region	n/a
(Reinders et al., 2003)	Household expenditure surveys from 11 countries, input-output data	Energy (GJ)	Households	Total (12 con-sumption cate-gories)	B, DK, EL, E, I, L, NL, P, FIN, S, UK	OLS regression (mean total household/per capita emissions for 11 countries)	Expenditure	d
(Vringer and Blok, 1995)	Household expenditure survey 1990, combined with input-output and process data.	Energy (GJ)	Households	Total, 13 con-sumption cate-gories	Netherlands	Descriptive	Income, hh size, age	(a)
(Weber and Perrels, 2000)		Energy (GJ), CO2	11 household types	Total; 6 con-sumption cate-gories	G, F, N	Descriptive	11 hh types	a, d
(Weber and Matthews, 2008)	Consumer Expenditure Survey 2004, input-output data	CO_2	Households	Total emissions, some split up into 13 con-sumption cate-gories for de-scriptive anal.	US	OLS regression (total emissions only)	Income, expendi-ture (separately), # children, # adults	a, b

(Wier et al., 2001)	Consumer survey, Statistics Denmark, input-output data	CO_2	Households	Total	Denmark	Descriptive and correlation	Income, expenditure household size	n/a

Note: a = issues related to input-output analysis, b = discussion of infrequency of purchase; c = standard errors / confidence intervals of mean emissions reported, d = functional vs. monetary units to estimate emissions; e = issues related to matching expenditure data to external data (e.g. prices), f = issues related to survey errors

Figure A 1: Lowess regression of total emissions (tonnes) on household expenditure



Note: Total emissions estimate is based on the 'mixed' method

'mixed' method home energy price matching

For the 'mixed' method we estimated home energy emissions by matching expenditure with price data. Whilst this approach cannot capture the entire variation of price differences in different areas or by different suppliers and tariffs, it can capture some variation which makes estimation more precise. For electricity and mains gas, the LCF/EFS provides information on government area, payment method and year and month of the interview which can be used for price matching. For electricity, gas and heating oil, we used DECC energy price statistics (Quarterly Energy Prices, tables 2.2.1, 2.3.1 and 4.1.1). The DECC statistics provide annual electricity and gas prices for three payment methods – credit (account), direct debit and prepayment, as well as each electricity and gas region. The differentiation by payment method is relevant because electricity and gas are more expensive per unit for households on prepayment meters than for the other two methods and usually cheapest for households who pay by direct debit. However, price variations within a year, between different providers and tariffs are neither captured in the price statistics nor in the LCF/EFS. The DECC electricity and gas price data include unit cost, standing charges and VAT whilst the LCF/EFS expenditure also includes meter rent and installation cost if applicable.

How did we match prices for different regions to the LCF/EFS? The electricity and gas regions[10] have broadly similar boundaries to government regions but they are not identical. Whilst previous studies (e.g. DEFRA, 2008; Druckman and Jackson, 2008) have treated electricity/gas regions and government regions as equivalent we sought to account for the differences in the areas covered by generating a unit price for each government region that represents the proportion of gas and electricity meters from different electricity or gas regions covered within each government region. This was achieved by using DECC statistics on sub-national gas and elec-

[10] There are 14 electricity regions which correspond to the previous Public Electricity Supplier (PES) regions and 12 gas regions, also called local distribution zones, LDZ.

tricity sales and consumers[11] which includes data on the number of electricity and gas meters at district level (local authority unit level 1) in each government region. Each district was matched to the electricity and gas region that they belong to, using information on the postcode areas covered by each electricity and gas region[12]. This enabled us to calculate the proportion of meters from different electricity and gas region in each government region for each year. Those proportions were then used to weight the unit prices that the DECC statistics provide for each electricity and gas region. For example, in 2008, the government region South-East included 10.06 per cent of meters from the gas region London, 47.27 per cent from the gas region South-East and the remaining 42.68 per cent from the gas region Southern. The unit prices of 3.60p for London, 3.58p for South-East and 3.54p for Southern were weighted by those proportions and then combined to the unit price of 3.54 which we applied to the government region South-East for 2008. The same procedure was applied to all government regions in each year.

Further assumptions needed to be made to match the LCF/EFS electricity and gas payment methods with those from the DECC price statistics. The DECC statistics provide unit prices for payment by credit, direct debit and prepayment.[13] Electricity and gas expenditure is recorded in four different variables in the derived household expenditure file in the LCF/EFS: "account", "standing order", "prepayment" and "second dwelling". Whilst at first sight those variables seem to correspond to the DECC categories as seems to be suggested by Druckman and Jackson (2008: 3180), they do not straightforwardly match the variables that record payment method (a128 for gas and a130 for electricity).

An additional complication is that the expenditure variable labelled 'slot meter' also records rebates for all other payment methods. It would therefore not be accurate to match the 'slot meter' variable expenditures with 'prepayment' prices. Furthermore, the payment method variables were recoded in 2009. Since then, direct debit payments all feed into the "account" expenditure variable even though direct debit payments are usually discounted. It would therefore be misleading to match the 'account' expenditure variables with 'credit' prices. Therefore, we did not use the four expenditure variables to match payment methods but the more detailed payment method categories of variables a128 and a130 as set out in table A2.

[11] Available from the DECC webpage http://www.decc.gov.uk/en/content/cms/statistics/energy_stats/regional/regional.aspx, last accessed 23 September 2011

[12] Information on postcodes covered by each electricity region is provided by energylinx http://energylinx.co.uk/electricity_distribution_map.htm and on postcodes in each gas regions by Xoserve http://www.xoserve.com/ (under "Postcode - Exit Zone Data").

[13] Prices are also provided for "largest", "average" and "smallest" bills. We have taken unit prices based on "average" bills.

Table A 2: Matching LCF/EFS payment methods to DECC price statistics payment methods

LCF payment methods variables a128 and a130, 2009	LCF/EFS payment methods variables a128 and a130, 2006-8	Matched to DECC payment methods
Direct debit	Budgeting scheme	Direct debit
Standing order		Direct debit
Monthly quarterly bill	Account	Credit
	Slot meter	Prepayment
Pre-payment (keycard or token) meters	Electricity card, disc, token or electr	Prepayment
Fuel direct from benefits	DSSs pay the whole or part of the bill	n/a
Included in rent	n/a	Credit
Frequent cash payment method	n/a	Prepayment
n/a	Paid direct by someone outside the house	n/a
Fixed annual bill	n/a	Credit
Other (please specify)	Some other method	Credit
n/a	Or by c.o.c.d. (NI and elec only)	Direct debit

This matching method leads to a reasonably good fit between DECC and LCF/EFS payment methods for 2006 and 2008. Based on this matching method, we allocate only 1.75 per cent more households to credit payments for electricity, 2.93 per cent fewer to direct debit and 1.16 per cent more to prepayment as compared to DECC data (as set out in DECC's Quarterly Energy Price statistics, tables 2.4.2 for electricity and 2.5.2 for gas). This is similar for gas as we allocate 0.7 per cent more households to credit payment, 2.42 per cent fewer to direct debit and 1.75 per cent more to the prepayment method in the LCF/EFS compared to DECC. However, for 2009 the matching process works less well: for electricity, 9.85 per cent fewer households are allocated to credit payments, 9.57 more to direct debit and 0.2 fewer to prepayment in our data compared to DECC statistics. The differences are again similar for gas payment methods as 10.28 per cent fewer households are allocated to credit payment, 10.77 per cent more to direct debit and 1.64 fewer to prepayment.

Prices for other heating fuels including bottled gas, coal and wood were sourced from the Sutherland tables.[14] Prices for wood were complemented with data from John Willoughby[15] as the Sutherland tables only provided *estimates* for wood prices for 2006 and 2007). The Sutherland tables provide unit cost for coal, bottled gas and wood (2008-9) on a bi-annual basis (whereby the prices reflect the average price for the six (or sometimes seven) months previous to the publication of the table). Different prices are provided for five different regions in the UK: Northern England, Midlands, South-East England, South-West and Wales, Scotland, and Northern Ireland. Since Sutherland tables could not provide any information on the postcodes covered by

[14] See http://www.sutherlandtables.co.uk/.

[15] See http://www.johnwilloughby.co.uk/, last accessed 23 September 2011. In 2008-10, the wood prices provided by John Willoughby, which are only provided for one area in the UK, were on average 12.93 per cent higher than the average price provided by Sutherland. We thus deducted 12.93 per cent of the prices provided by Willoughby to estimate average prices for 2006-7.

those regions, we were not able to weight unit prices in the same way we did for electricity and gas. Therefore, we matched the Sutherland regions to the government regions as detailed in table A3.

Table A 3 – Matching Sutherland regions to Government regions (LCF/EFS)

Sutherland region	**Government Office Region**
NA – Northern England	North East
NA – Northern England	North West
NA – Northern England	Yorkshire and the Humber
MA = Midlands	East Midlands
MA = Midlands	West Midlands
SEA – South East England	Eastern
SEA – South East England	London
SEA – South East England	South East
SWWA – South West and Wales	South West
SWWA – South West and Wales	Wales
SCA - Scotland	Scotland
NIA – Northern Ireland	Northern Ireland

After the completion of price matching for all different types of home energy, households expenditure on electricity, gas and other fuels was then divided by the unit cost to estimate kWh or unit consumed (e.g. kilograms of coal and wood and litres of bottled gas). DECC/DEFRA (2011) emission factors for electricity[16], gas, heating oil, bottled gas and coal were then applied to estimate household emissions from home energy.

Mixed method motor fuel price matching. To estimate emissions from motor fuels in the 'mixed' approach we utilised data from AA on the unit cost of motor fuels to estimate units consumed. DECC conversion factors for petrol and diesel were then applied to estimate emissions. Employing price data on motor fuels again does not capture the full variation of prices across time and space as they are only available in aggregated form. However, the AA motor fuel statistics employed in this study are provided per month and government regions thus capturing price variations for these dimensions (DECC also provides monthly motor fuel statistics but not broken down by region). A comparison of AA and DECC prices (after we had weighted the AA regional data by population size) showed that the AA prices were on average 0.6 pence per litre higher for petrol and 0.5 pence higher for diesel than the DECC data. Estimated emissions will therefore be slightly lower using the AA data compared to applying DECC data. Both price datasets are based on surveys of the main motor fuel providers in the UK but AA has a larger proportion of the big supermarkets which also provide petrol and diesel. The AA price data were merged with the LCF/EFS dataset using month and government region variables to estimate litres of petrol or diesel consumed per week. DECC emission factors for premium unleaded petrol and diesel cars per litre were then applied to estimate CO_2 emissions.

[16] Different factors were used for electricity for each year because emissions from electricity generation vary over time depending on the underlying technology and fuel mix. We applied the "scope 2" electricity emissions factor for electricity as consumed at the point of usage, i.e. losses in the grid are included. However, scope 2 emissions do not include emissions arising from producing the fuels used in electricity generation (they are included in scope 3 but this is only expressed in greenhouse gases, not CO2 separately).

Fit between REAP and LCF/EFS categories

As described in the main text, emissions data from the REAP database were used to create emissions factors for the 'reap' and 'mixed' methods (here only for indirect emissions and three transport categories). While both the LCF/EFS and the REAP database categorise expenditure according to the COICOP (Classification of Individual Consumption According to Purpose) typology, there is no perfect fit between categories in each dataset. Most differences are relatively minor, for instance, REAP sometimes combines several COICOP categories in one single category or allocates individual consumption items to a different category than the LCF/EFS. For example, 'sugar' belongs to the 'chocolate and confectionary' category in the LCF/EFS but to 'other foods' in REAP. Tables A4 and A5 below set out in more detail how we matched expenditure categories.

The most relevant differences between the LCF/EFS and REAP relate to transport categories, particularly for "combined fares" and "package holidays". The REAP database does not include emissions for "combined fares" or "package holidays". Instead, the REAP team had allocated emissions for these categories to others such as rail travel, flights and accommodation. However, no information was available regarding the proportions that had been used to allocate emissions to these categories [email correspondence with Anne Owen, 27 June and 21 July 2011, and Tommy Wiedmann on 5 August 2011]. Whilst this re-allocation within REAP does not make a difference to estimating total UK CO_2 emissions, it introduces an inaccuracy for estimating emissions at the household level because we cannot match expenditure to emissions (instead, every household who has an expenditure for the categories to which these emissions have been distributed is effectively allocated an equal share of these emissions).

Since only 458 households or 2 per cent of our sample had an expenditure on "combined fares" we have not re-allocated expenditure to other categories in the 'reap' method.[17] However, we reallocated package holiday spending to other categories because 3897 or 16 per cent of the sample had an expenditure on package holidays which are also likely to include emission-intensive flights for holidays abroad. Re-allocating spending decreases the emissions per £ expenditure factor generated for these different categories because it increases the denominator of that factor. At the same time, this lower factor is then multiplied with the total travel/accommodation expenditure that a household has, thus correctly allocating emissions to those households who have had an expenditure on package holidays rather than distributing emissions evenly to all households with expenditures for these other categories. To re-allocate expenditure for package holidays we used data from the National Travel Survey for Great Britain, the Northern Ireland Transport Statistics for Northern Ireland and the International Passenger Survey[18] to calculate proportions of different types of travel. For the 'reap' estimate, expenditure on package holidays was re-allocated to flights/travel and accommodation. For package holidays abroad we assumed people mainly travel by plane whilst small proportions of the expenditure was also allocated to train, bus and ferry travel (based on data from the International Passenger Survey on flight/ferry passengers and expenditure on different types of travel within the LCF/EFS).[19] We then calculated

[17] In the "mixed" method we apply a factor for the kilometres travelled per £ expenditure, weighting the factors for rail/tube and bus/coach travel by the proportions of total passenger kilometres for these modes of travel for each year. A weighted emissions factor for train/tube and bus/coach travel is then applied.

[18] Sources: Table 0305 National Travel Survey for Great Britain and table 3.1 Northern Ireland Travel Survey 2007-2009 In Depth Report.

[19] Expenditure for package holidays abroad was allocated to the following categories: 2 per cent to rail, 2 per cent to road, 7 per cent to ferry, 46.4 per cent to flights and 42.4 per cent to accommodation. Expenditure for

the proportion of spending on plane tickets to expenditure on accommodation abroad using data from LCF/EFS on holiday spending. Expenditure on package holiday abroad was then allocated to non-UK flights, accommodation, train, ferry and road travel. The same procedure was applied to spending on holiday packages in the UK, using data from the National Travel survey to calculate the proportion of non-local bus/coach, train and ferry travel and the proportion of travel/UK accommodation spending from the LCF/EFS.

For the 'mixed' estimate, we did not reallocate any package holiday spending to flights because package holiday flights are already captured in the survey question on the number of flights in the last year. However, we still allocated the same proportions of package holiday spending to train, road and ferry travel as well as accommodation.

Table A 4: LCF/EFS and REAP transport category match

LCF/EFS	REAP
Rail and tube	Rail and tube
Bus and coach	Road services
Taxis and hired cars with drivers	Road services
Hire of self-drive cars, vans, bicycles	Road services
Car leasing	Road services
School travel	Road services
Other personal travel and transport services	Road services
Combined fares	A weighted emissions factor combining train/tube and road services has been applied
Air fares (within UK)	Air transport
Air fares (international)	Air transport
Water travel	Ferry (water transport)
Other transport services	Ancilliary Transport

Table A 5: LCF/EFS – REAP match for all other consumption

LCF/EFS	**REAP**
Bread	Bread, biscuits and pastry
Buns, cakes, biscuits etc	Bread, biscuits and pastry
Pastry (savory)	Bread, biscuits and pastry
Other breads and cereals (this contains only cereals)	Grains and starch products
Rice	Grains and starch products
Pasta products	Grains and starch products
Beef (fresh, chilled or frozen)	Meat, excl. poultry
Pork (fresh, chilled or frozen)	Meat, excl. poultry
Lamb (fresh, chilled or frozen)	Meat, excl. poultry
Bacon and ham	Meat, excl. poultry
Other meats and meat preparations	Meat, excl. poultry
Poultry (fresh, chilled or frozen)	Poultry
Fish	Fish
Milk, cheese and eggs:	Dairy products
Oils & fats	Oils and fats
Fruit	Fruit and Vegetables

package holidays in the UK was allocated to the following categories: 17.7 per cent to rail travel, 33 per cent to road travel (mainly bus/coach), 1.3 per cent to travel by ferry and 48 per cent to accommodation.

Vegetables	Fruit and Vegetables
Sugar, jam, honey, chocolate and confectionary	Chocolate, cocoa and confectionary (excluding sugar)
Food products not elsewhere specified	Other foods (including sugar)
Non-alcoholic beverages:	Non-alcoholic beverages
Alcoholic beverages	Alcoholic beverages
Tobacco and narcotics	Tobacco (excl. narcotics)
Clothing	Clothing
Footwear	Footwear
Actual rentals for housing (net)	Housing: expenditure on rent
Expenditure on mortgages (capital repayments, interest and protection premium payments)	Housing: Expenditure on mortgages
Maintenance and repair of dwelling	House maintenance and repair
Water supply and miscellaneous services relating to the dwelling	Private services: water utilities
Furniture and furnishings, carpets and other floor coverings	Furniture and furnishings, incl carpets
Household textiles	Textiles
Household appliances	Household appliances
Glassware, tableware and household utensils	Glassware and household utensils
Tools and equipment for house and garden	Garden equipment and household tools
Goods and services for routine household maintenance	Goods and services for routine household maintenance
Medical products, appliances and equipment	Medical products, appliances and equipment
Hospital services (excl. outpatient services)	Hospital services
Outpatient services	Out-patient services
Purchase of vehicles	Purchase of vehicles
Operation of personal transport equipment (motor fuels have been allocated to transport)	Expenditure of running a vehicle
Postal services	Postal services
Telephone and telefax equipment	Telephone and telefax equipment
Telephone and telefax services	Telephone and telefax services
Audio-visual, photographic and information processing equipment	Audio-visual and photo processing equipment
Other major durables for recreation and culture	Items for recreation and culture (major durables)
Other recreational items and equipment, gardens and pets	Other recreational equipment
Recreational and cultural services	Recreational and cultural services
Newspapers, books and stationery	Newspapers, books and stationery
Package holidays (distributed to accommodation services and transport)	Package holidays
Education fees	Private services: education
Payments for school trips, other ad-hoc expenditure	Private services: education
Catering services	Catering services
Accommodation services	Accommodation services
Personal care	Personal care

Personal effects n.e.c	Jewellery and personal items (this is equivalent to personal effects n.e.c.)
Social protection	Social protection
Insurance	Insurance
Bank, building society, post office, credit card charges	Financial Services
Other services nec	Other business services (equivalent to "other services nec")

Table A6: OLS regression on total and transport CO_2 emissions, restricted model

	Total emissions			**Transport emissions**		
	Mixed	Reap	Single factor	Mixed	Reap	Single factor
lnincome	0.462***	0.452***	**0.623*****	0.786***	**0.713*****	**0.637*****
	(0.00611)	(0.00654)	(0.00679)	(0.0150)	(0.0164)	(0.0141)
adult2	0.262***	0.270***	**0.212*****	0.300***	0.325***	0.264***
	(0.00850)	(0.00924)	(0.00897)	(0.0204)	(0.0231)	(0.0198)
adult3	0.134***	0.139***	**0.0737*****	0.0754***	0.0959***	**0.147*****
	(0.0106)	(0.0116)	(0.0108)	(0.0248)	(0.0284)	(0.0250)
adult4	0.0757***	0.0819***	0.0605***	0.0810*	*0.115***	*0.132****
	(0.0212)	(0.0221)	(0.0213)	(0.0490)	(0.0528)	(0.0468)
adult5	0.0558	0.0191	0.0420	-0.0474	-0.177	-0.117
	(0.0451)	(0.0504)	(0.0478)	(0.0918)	(0.109)	(0.0949)
child1	0.113***	0.108***	**0.192*****	-0.0475**	*-0.0686****	*-0.0348**
	(0.00882)	(0.00939)	(0.00882)	(0.0204)	(0.0231)	(0.0203)
child2	0.0766***	0.0728***	0.0713***	0.0599**	0.0581**	*0.0303*
	(0.0114)	(0.0120)	(0.0110)	(0.0271)	(0.0296)	(0.0259)
child3	0.0391***	0.0355**	***-0.0177***	-0.119***	*-0.103***	*-0.0541*
	(0.0151)	(0.0160)	(0.0154)	(0.0411)	(0.0431)	(0.0377)
Constant	-0.223***	-0.140***	-1.153***	-3.889***	-3.483***	-3.903***
	(0.0349)	(0.0371)	(0.0382)	(0.0884)	(0.0955)	(0.0820)
Observations	21,664	21,664	21,664	19,140	19,140	19,140
R-squared	0.523	0.477	0.596	0.295	0.225	0.234

Note: Zero emissions, 1st and 99th percentile of all dependent emission variables and the income variable are excluded in the models. *** is significant at 1 per cent level, ** significant at 5 per cent level. Bold beta coefficients are significantly different from “mixed method” coefficients at least at 5 per cent level. Coefficients in italics have different level of significance or different sign compared to the ‘mixed’ model.

Table A7: OLS regression on indirect and home energy CO_2 emissions, restricted model

	Indirect emissions			Home energy emissions		
	Mixed	Reap	Single factor	Mixed	Reap	Single factor
lnincome	0.511***	0.499***	**0.643*****	0.140***	0.127***	0.121***
	(0.00657)	(0.00654)	(0.00730)	(0.00764)	(0.00803)	(0.00716)
adult2	0.268***	0.265***	**0.203*****	0.200***	0.211***	0.193***
	(0.00899)	(0.00897)	(0.00960)	(0.0111)	(0.0117)	(0.0104)
adult3	0.133***	0.134***	**0.0615*****	0.153***	0.161***	0.163***
	(0.0115)	(0.0113)	(0.0115)	(0.0146)	(0.0149)	(0.0142)
adult4	0.0653***	0.0648***	*0.0494***	0.0514**	0.0525**	0.0552**
	(0.0234)	(0.0231)	(0.0227)	(0.0256)	(0.0259)	(0.0249)
adult5	0.0382	0.0382	0.0522	0.0917	*0.104**	*0.112**
	(0.0504)	(0.0502)	(0.0499)	(0.0621)	(0.0606)	(0.0630)
child1	0.152***	0.150***	**0.213*****	0.0992***	0.110***	0.102***
	(0.00945)	(0.00938)	(0.00940)	(0.0119)	(0.0122)	(0.0113)
child2	0.0802***	0.0801***	0.0760***	0.0889***	0.0854***	0.0850***
	(0.0120)	(0.0119)	(0.0116)	(0.0148)	(0.0151)	(0.0143)
child3	0.0273*	0.0303*	***-0.0255***	0.0908***	0.0987***	0.0961***
	(0.0157)	(0.0155)	(0.0162)	(0.0203)	(0.0211)	(0.0191)
Constant	-1.219***	-1.112***	-1.416***	0.429***	0.499***	*-1.207****
	(0.0372)	(0.0371)	(0.0410)	(0.0429)	(0.0450)	(0.0402)
Observations	21,664	21,664	21,664	21,664	21,664	21,664
R-squared	0.528	0.520	0.575	0.142	0.131	0.147

Note: See note for table A6.

Table A8: OLS regression on home energy, extended model

	Home energy emissions		
VARIABLES	Mixed	Reap	Single factor
lnincome	0.079***	0.069***	0.076***
	(0.009)	(0.009)	(0.009)
adult2	0.143***	0.146***	0.151***
	(0.011)	(0.011)	(0.010)
adult3	0.104***	0.109***	0.117***
	(0.014)	(0.014)	(0.014)
adult4	0.031	0.026	0.035
	(0.024)	(0.024)	(0.024)
adult5	0.055	0.057	0.071
	(0.060)	(0.058)	(0.061)
child1	0.115***	0.118***	0.115***
	(0.013)	(0.013)	(0.012)
child2	0.041***	*0.036***	0.042***
	(0.015)	(0.015)	(0.014)
child3	0.070***	0.070***	0.071***
	(0.019)	(0.020)	(0.019)
age	0.011***	0.010***	0.015***
	(0.002)	(0.002)	(0.002)
age2_100	-0.008***	-0.007***	-0.013***
	(0.002)	(0.002)	(0.002)
agetop	0.043**	0.048**	0.042**
	(0.020)	(0.021)	(0.019)
Female	0.050***	0.053***	0.044***
	(0.008)	(0.009)	(0.008)
hhedu16_m	-0.003	-0.004	-0.001
	(0.011)	(0.012)	(0.011)

hhedu1215_m	0.012	0.007	0.008
	(0.010)	(0.010)	(0.010)
edum	-0.027*	-0.023	***0.082****
	(0.016)	(0.017)	(0.015)
wlhh	0.039**	0.050***	0.053***
	(0.015)	(0.016)	(0.015)
refeth	0.016	*0.039***	0.010
	(0.017)	(0.017)	(0.016)
rur	-0.007	-0.006	0.007
	(0.011)	(0.011)	(0.010)
rural_m	0.028	*0.106****	*0.064***
	(0.032)	(0.034)	(0.026)
own_out	0.069***	0.063***	0.041***
	(0.013)	(0.013)	(0.012)
own_mort	0.070***	0.053***	**0.034*** **
	(0.012)	(0.012)	(0.011)
h_mis	0.171***	0.179***	0.168***
	(0.033)	(0.034)	(0.031)
detached	0.273***	0.275***	0.245***
	(0.018)	(0.019)	(0.017)
semid	0.201***	0.205***	0.174***
	(0.016)	(0.017)	(0.015)
terraced	0.156***	0.165***	0.133***
	(0.016)	(0.017)	(0.015)
flatconv	0.083***	0.086***	0.077***
	(0.028)	(0.030)	(0.027)
central_elec	-0.218***	**-0.365*** **	***-0.014***
	(0.017)	(0.018)	(0.017)
central_oil	0.044	-0.015	*0.079****
	(0.031)	(0.034)	(0.025)
heat_other2	-0.176***	**-0.263*** **	**-0.088*** **
	(0.021)	(0.021)	(0.019)
bedroom	0.107***	0.110***	0.102***
	(0.006)	(0.006)	(0.005)
Constant	0.011	0.101	-1.768***
	(0.066)	(0.068)	(0.064)
Observations	21,658	21,658	21,658
R-squared	0.258	0.264	0.238

Note: See note for table A6.

References

Baiocchi, G., Minx, J., Hubacek, K., 2010. The Impact of Social Factors and Consumer Behavior on Carbon Dioxide Emissions in the United Kingdom. Journal of Industrial Ecology 14, 50-72.

Baker, P., Blundell, R., Micklewright, J., 1989. Modelling Household Energy Expenditures Using Micro-Data. The Economic Journal 99, 720-738.

Baker, P., Micklewright, J., 1987. Fuel and the Family Expenditure Survey. Working Paper No. 164, Department of Economics, Queen Mary College, University of London.

Bardsley, N., Buchs, M., Schnepf, S.V., forthcoming. From expenditure to emissions II? Examining the infrequency of purchase problem in research on household CO2 emissions, to be published as S3RI working paper.

Brand, C., Boardman, B., 2008. Taming of the few - The unequal distribution of greenhouse gas emissions from personal travel in the UK. Energy Policy 36, 224-238.

Brand, C., Preston, J.M., 2010. '60-20 emission'-The unequal distribution of greenhouse gas emissions from personal, non-business travel in the UK. Transp. Policy 17, 9-19.

Buchs, M., Schnepf, S.V., 2013. UK households' carbon footprint: a comparison of the association between household characteristics and emissions from home energy, transport and other goods and services, IZA Discussion Paper No. 7204, Bonn: IZA, available from http://ftp.iza.org/dp7204.pdf.

Burney, N.A., 1995. Socioeconomic development and electricity consumption A cross-country analysis using the random coefficient method. Energy Econ. 17, 185-195.

Cohen, C., Lenzen, M., Schaeffer, R., 2005. Energy requirements of households in Brazil. Energy Policy 33, 555-562.

Deaton, A., Irish, M., 1984. Statistical models for zero expenditures in household budgets. Journal of Public Economics 23, 59-80.

DECC, DEFRA, 2011. 2011 Guidelines to Defra / DECC's GHG Conversion Factors for Company Reporting. London: Department for Environment, Food and Rural Affairs and Department for Energy and Climate Change.

DEFRA, 2008. Distributional Impacts of Personal Carbon Trading. London: Department for Environment, Food and Rural Affairs, March 2008, online http://www.defra.gov.uk/environment/climatechange/uk/individual/carbontrading/pdf/pct-distributional-impacts.pdf, download 1 July 2008.

Dresner, S., Ekins, P., 2006. Economic instruments to improve UK home energy efficiency without negative social impacts. Fiscal Studies 27, 47-74.

Druckman, A., Jackson, T., 2008. Household energy consumption in the UK: A highly geographically and socio-economically disaggregated model. Energy Policy 36, 3177-3192.

Druckman, A., Jackson, T., 2009. The carbon footprint of UK households 1990-2004: A socio-economically disaggregated, quasi-multi-regional input-output model. Ecol. Econ. 68, 2066-2077.

Duarte, R., Mainar, A., Sanchez-Choliz, J., 2010. The impact of household consumption patterns on emissions in Spain. Energy Econ. 32, 176-185.

Fahmy, E., Thumim, J., White, V., 2011. The distribution of UK household CO2 emissions: Interim report. JRF programme paper: climate change and social justice, University of Bristol and Centre for Sustainable Energy.

Girod, B., De Haan, P., 2010. More or Better? A Model for Changes in Household Greenhouse Gas Emissions due to Higher Income. Journal of Industrial Ecology 14, 31-49.

Gough, I., Abdallah, S., Johnson, V., Ryan-Collins, J., Smith, C., 2011. The distribution of total greenhouse gas emissions by households in the UK, and some implications for social policy. CASE paper 152, Centre for Analysis of Social Exclusion, London: London School of Economics.

Groot, L., 2010. Carbon Lorenz curves. Resour. Energy Econ. 32, 45-64.

Groves, R.M., Fowler, F.J., Couper, M.P., Lepkowski, J.M., Singer, E., Tourangeau, R., 2009. Survey Methodology. John Wiley, Hoboken, New Jersey.

Heinonen, J., Junnila, S., 2011. A Carbon Consumption Comparison of Rural and Urban Lifestyles. Sustain. 3, 1234-1249.

Herendeen, R., Tanaka, J., 1976. Energy cost of living. Energy 1, 165-178.

Herendeen, R.A., Ford, C., Hannon, B., 1981. Energy cost of living, 1972–1973. Energy 6, 1433-1450.

Hertwich, E.G., 2005. Life Cycle Approaches to Sustainable Consumption: A Critical Review. Environmental Science & Technology 39, 4673-4684.

Jacobson, A., Milman, A.D., Kammen, D.M., 2005. Letting the (energy) Gini out of the bottle: Lorenz curves of cumulative electricity consumption and Gini coefficients as metrics of energy distribution and equity. Energy Policy 33, 1825-1832.

Kerkhof, A.C., Benders, R.M.J., Moll, H.C., 2009. Determinants of variation in household CO2 emissions between and within countries. Energy Policy 37, 1509-1517.

Larivière, I., Lafrance, G., 1999. Modelling the electricity consumption of cities: effect of urban density. Energy Econ. 21, 53-66.

Lenzen, M., Wier, M., Cohen, C., Hayami, H., Pachauri, S., Schaeffer, R., 2006. A comparative multivariate analysis of household energy requirements in Australia, Brazil, Denmark, India and Japan. Energy 31, 181-207.

Leontief, W., 1970. Environmental repercussions and the economic structure: an input-output approach. The Review of Economics and Statistics 52, 262-271.

Lyons, S., Pentecost, A., Tol, R.S.J., 2012. Socioeconomic distribution of emissions and resource use in Ireland. Journal of environmental management 112, 186-198.

O'Neill, B.C., Chen, B.S., 2002. Demographic determinants of household energy use in the United States. Popul. Dev. Rev. 28, 53-88.

ONS, 2001-2011. Family Spending, Office of National Statistics, http://www.ons.gov.uk/ons/rel/family-spending/family-spending/index.html, editions 2001-2011.

Pachauri, S., 2004. An analysis of cross-sectional variations in total household energy requirements in India using micro survey data. Energy Policy 32, 1723-1735.

Papathanasopoulou, E., Jackson, T., 2009. Measuring fossil resource inequality-A case study for the UK between 1968 and 2000. Ecol. Econ. 68, 1213-1225.

Paul, A., Wiedmann, T., Barrett, J., Minx, J., Scott, K., Dawkins, E., Owen, A., Briggs, J., Gray, I., 2010. Introducing the Resources and Energy Analysis Programme (REAP). Stockholm Environment Institute.

Reinders, A.H.M.E., Vringer, K., Blok, K., 2003. The direct and indirect energy requirement of households in the European Union. Energy Policy 31, 139-153.

Tiffin, R., Arnoult, M., 2010. The demand for a healthy diet: estimating the almost ideal demand system with infrequency of purchase. Eur. Rev. Agric. Econ. 37, 501-521.

Vringer, K., Blok, K., 1995. The direct and indirect energy requirements of households in the Netherlands. Energy Policy 23, 893-910.

Vringer, K., Blok, K., 1997. Do higher incomes pay more? The effect of price-income relations on the direct and indirect energy requirements of households in the Netherlands. Report, Department of Science, Technology and Society, University of Utrecht, The Netherlands.

Weber, C., Perrels, A., 2000. Modelling lifestyle effects on energy demand and related emissions. Energy Policy 28, 549-566.

Weber, C.L., Matthews, H.S., 2008. Quantifying the global and distributional aspects of American household carbon footprint. Ecol. Econ. 66, 379-391.

Wier, M., Lenzen, M., Munksgaard, J., Smed, S., 2001. Effects of Household Consumption Patterns on CO2 Requirements. Econ. Syst. Res. 13, 259 - 274.