EUROMOD WORKING PAPER SERIES

EM12/14

Integrating VAT into EUROMOD. Documentation and results for Belgium

André Decoster, Richard Ochmann, and Kevin Spiritus

June 2104

Integrating VAT into EUROMOD. Documentation and results for Belgium[1]

André Decoster[a]
Richard Ochmann[b]
Kevin Spiritus[a]

[a] Centre for Economic Studies - Public Economics, Katholieke Universiteit Leuven
[b] German Institute for Economic Research (DIW Berlin)

Abstract

This paper documents the integration of microsimulation tools for direct taxation, indirect taxation, and social benefits in the context of the European tax and benefit simulator, EUROMOD. Integration has been developed in parallel for two countries: Belgium and Germany. The paper at hand documents the process and presents simulation results for the case of Belgium. An integrated database underlying EUROMOD that contains household-level information on income and consumption is generated. Consumption micro data from the 2009 cross section of the household budget survey for Belgium is used to impute information on spending for durable and non-durable commodities into EU-SILC data, applying regression-based imputation techniques. Engel curves are estimated at the household level for total non-durable spending, expenditures on durable goods, as well as non-durable expenditure share equations. The imputed household spending is then used to simulate the baseline VAT system in EUROMOD, for which we report an incidence analysis. Finally, several arbitrary policy reforms implementing VAT rate uniformity are analyzed with respect to their distributional impact.

JEL Classification: D12, D31, H24, H31
Keywords: Budget survey, expenditure estimation, Engel curves, EUROMOD.

Corresponding author:
Kevin Spiritus
E-mail: Kevin.Spiritus@kuleuven.be

---

[1]This paper was written as part of the SBO-project "FLEMOSI: A tool for ex ante evaluation of socio-economic policies in Flanders", funded by IWT Flanders. The project intends to build `FLEmish MOdels of SImulation'. See www.flemosi.be. The results presented here are based on EUROMOD version F6.36. EUROMOD is maintained, developed and managed by the Institute for Social and Economic Research (ISER) at the University of Essex in collaboration with national teams from the EU member states. The process of extending and updating EUROMOD is financially supported by the Directorate General for Employment, Social Affairs and Inclusion of the European Commission [Progress grant no. VS/2011/0445]. We are indebted to the many people who have contributed to the development of EUROMOD and to the European Commission for providing financial support for it. This paper greatly benefited from valuable discussions with Peter Haan, and the audience of the EUROMODupdate2 project meeting 2012 in Bucharest. We are grateful to Vincent Mouton from Statistics Belgium for his valuable support with the Household Budget Survey. The results and their interpretation are the authors' responsibility.

Contents

1. Introduction

This paper documents the integration of microsimulation tools for direct taxation, indirect taxation, and social benefits in the context of the European tax and benefit simulator, EUROMOD. Integration has been developed in parallel for two countries: Belgium and Germany. The paper at hand documents the process and presents simulation results for the case of Belgium.

Because the European microsimulation model EUROMOD[1] is already able to simulate direct taxation and social benefits, the exercise mainly consists of integrating indirect taxation into the existing environment of EUROMOD such that indirect tax policies are simultaneously available with direct tax policies and social benefits.

The integration essentially involves three components. First we need to enrich the SILC with detailed expenditure information at the household level. Second, we need to adjust the EUROMOD implementation, so that the user can change the VAT rates. And third, we need to integrate the change in consumer prices, resulting from a VAT-change, in the distributional analysis.

The major focus of this paper will be on the first component, namely generating an integrated database that contains household-level information on income and consumption, which is used by EUROMOD to run its simulations. We use consumption micro data from the household budget survey (HBS) for Belgium to impute information on spending for durable and non-durable commodities into the SILC data, which serves as an input for EUROMOD. This is necessary because there is no information on household consumption expenditures in the EU-SILC data. The general strategy involves regression-based imputation. Total non-durable spending, demand for durable goods, as well as non-durable expenditure share equations are estimated at the household level, using the 2009 cross section from the Belgian household budget survey (HBS)[2]. Coefficient estimates are applied to impute this information into the SILC database underlying EUROMOD.

The VAT legislation is integrated into EUROMOD in the form of a new policy sheet. Stata do-files have been developed in order to simulate household expenditures and VAT liabilities, given disposable incomes and indirect tax parameters from EUROMOD. These do-files are invoked from the new policy that was added to EUROMOD. We will not discuss these do-files and these changes to EUROMOD in the current paper, the interested reader is referred to Decoster and Spiritus (2014).

We conclude our paper with a number of simulations, further validating our exercise. First we simulate status quo tax legislation on imputed household spending and provide

[1] For an introduction into EUROMOD, see Sutherland and Figari (2013).

[2] More information about the Belgian HBS can be found on the website of Statistics Belgium: `http://statbel.fgov.be/fr/statistiques/collecte_donnees/enquetes/budget_des_menages/` (available in Dutch or in French)

an incidence analysis of baseline VAT. We find that VAT looks regressive when plotted against the income distribution – tax burdens decrease in income in relative terms – while VAT is slightly progressive when plotted against the expenditure distribution – tax burdens increase in spending in relative terms. Next we simulate a number of ad hoc policy reforms that affect the three areas of direct taxation and indirect taxation simultaneously. The reforms generally build on uniformity with respect to reduced VAT rates, and they can be implemented such that both revenue and distributional neutrality are guaranteed. Note that in the simulations, price elasticities are not taken into account. Only the effects of changing prices on real incomes are corrected for.

We do not perform a true welfare analysis in this paper. We do look at changes in tax liabilities, but we ignore the effects of changes in consumption patterns on individual welfare. In Decoster and Spiritus (2014) an example is given of how changes in consumer prices can be taken into account in a welfare analysis of tax reforms.

Since we started writing this paper, an additional feature has been implemented in the model. We are now also able to simulate excises, both specific and ad valorem. This feature is described in Decoster and Spiritus (2014)[3].

This paper is organized as follows. In section 2 we given an overview of the data used for our estimations and our imputations. Next we sketch the procedures followed for the estimations and the imputations in section 3. Sample descriptives are given in section 4, ensuring us that the covariates used in the estimations and the imputations are similarly distributed. The results of the estimations are discussed in section 5, where we given an overview of the estimation parameters and the demand elasticities. We discuss the results of the imputations in section 6. Distributional plots will evaluate the imputations for each of the single commodity groups. In section 7 we discuss some simulations.

2. Data

The consumption data for Belgium used in this analysis stems from the Belgian Household Budget Survey (HBS). The HBS is maintained by Statistics Belgium in repeated cross-sections. Households are drawn from the National Register every year, and invited to participate on a voluntary basis.

[3]For the future, we plan to implement a full demand system, taking into account also relative price effects, using e.g. Stone-Lewbel prices.

Households in the HBS data report detailed information on consumption. Expenditures are reported for over 1300 different commodities. We have aggregated these commodities into 15 commodity groups of non-durable expenditures, and one group for durable expenditures[4], for the purpose of this analysis. This aggregation follows the COICOP aggregation principles[5]. We consider 15 commodity groups of non-durable spending:

Commodity Groups of Non-durable Expenditures:

1. Food, non alcoholic beverages
2. Alcoholic beverages
3. Tobacco
4. Clothing and footwear
5. Home fuels and electricity
6. Rents (excluding imputed rents)
7. Household services
8. Health
9. Private transport
10. Public Transport
11. Communication
12. Recreation and culture
13. Education
14. Restaurants
15. Other goods and services

The total population covered by the HBS data is slightly restricted, as there are groups that are not covered: institutionalized people (i.e. military personnel in barracks, students in dormitories, elderly and disabled people in nursery homes or hospitals, nurses or migrant workers in residences, people in jails) and homeless people. The files for the 2009 cross section used in this analysis contain 3,599 households.

[4] Spending on durable goods includes purchases of furniture, floor covering, decorative objects, lamps, household goods made of textile such as mattresses, pillows and curtains, large household appliances such as heating systems, furnaces, freezers, washing machines, microwave ovens and coffee-makers, dinnerware, household goods such as scales, tools for the house and the garden, and vehicles such as cars, caravans, bicycles and motorcycles. Using such a broad definition limits the number of zero observations in durable expenditures.

[5] Also see Decoster *et al.* (2007).

A few observations have been dropped during the analysis, because of outlying consumption or disposable income. An automated procedure was implemented in order to achieve this. Both disposable income and total non-durable expenditures were fitted on a log-normal distribution, after which observations with extremely low probability were dropped. In this procedure, the logarithm of the concerned variables is taken. By subtracting their mean and dividing by their standard deviation, variables are obtained which, by assumption, follow a standard normal distribution. Observations with an extremely low probability, i.e. lower than $\frac{1}{2N}$, are dropped. This procedure of standard normalization and dropping outliers is repeated until no further outliers are found. For the HBS of 2009, this means that two observations were dropped because of extreme disposable incomes, and one observation because of extreme total non-durable expenditures.

Our EUROMOD simulations make use of the EU-SILC data from 2010, containing reported incomes for 2009. The target population is individuals living in private households in Belgium, except for people in institutions, i.e. for example soldiers living in military caserns, or old people living in nursing homes. The sample consists of 14,700 individuals, living in 6,132 households.

The standard output from EUROMOD contains individual-level data. This has been aggregated up to household-level data for the purpose of this analysis, such that the micro data structure between the budget survey, which is on household level, and the EUROMOD output data is compatible. For more details on the Belgian implementation of EUROMOD and the EU-SILC data used, see Hufkens *et al.* (2013).

Current disposable household income has been defined such that it follows a concept that is consistent between the two surveys applied. It *ex*cludes imputed rents for owner occupiers. Expenses for alimonies have been deducted for households observed paying them to other households, and included in turn in the income of the households observed receiving them. Because disposable incomes reported in the HBS diaries were found to be of low quality, it was decided to use reported "usual" monthly incomes. These correspond fairly well to incomes reported in EU-SILC.

Before we provide a more detailed overview of the sample descriptives of both datasets in section 4, we first introduce the estimation and imputation procedures, in the following section.

3. Estimation of the expenditure allocation

The model estimated at the household level involves the allocation of current disposable income to consumption in the current period, where several commodity groups will be differentiated, and the residual to savings. Generally, the model is constructed in two stages.

In the first stage, two fequations are estimated separately: one for total non-durable expenditures and one for durable expenditures. The estimation techniques will be different for the two equations. At the second stage, the allocation of total *non*-durable expenditures on to 15 commodity groups is modeled, in terms of budget share equations. Note that price elasticities are not being considered – although at the simulation stage effects of changing prices on real incomes are taken into account.

Total Non-Durable Expenditures

The equation for total household expenditures on non-durable commodities, estimated on the cross-section from consumption micro data, is specified as:

$$\ln(e_i^{nd}) = \alpha^{nd} + \gamma_1^{nd}\ln(y_i) + \gamma_2^{nd}(\ln(y_i))^2 + \gamma_3^{nd}(\ln(y_i))^3 + \mathbf{x_i'}\beta^{\mathbf{nd}} + \varepsilon_i \tag{1}$$

for households $i = 1, \ldots, N$. The dependent variable $\ln(e_i^{nd})$ is the logarithm of total monthly expenditures for non-durable consumption at the household level. The non-durable expenditures variable e_i^{nd} is strictly positive for all observations in the budget data, as is disposable income y_i, so Eq. (1) is defined for all households and the estimation can be conducted unconditionally for all observations. In the imputation stage, the prediction of $\ln(e_i^{nd})$ in EU-SILC has been adjusted with a common factor, the expected value of $\exp(\varepsilon_i)$.[6]

Among the explanatory variables, $\ln(y_i)$ denotes the logarithm of current disposable household income. $\mathbf{x_i}$ denotes a $K \times 1$-vector of household-specific characteristics, such as the number of adults and children in the household, the number of household members currently in work, and some demographic variables related to the household head, such as age, region, education, gender, and employment status. We also include interaction terms between the income function and different demographic variables, in order to take into account the different effects of these covariates at different income levels. The error term ε_i is assumed to be independent and identically distributed. A full overview of the covariates used in the Belgian model can be found in appendix C, containing all regression results.

The coefficient estimates $\widehat{\alpha}^{nd}$, $\widehat{\gamma}_1^{nd}$, $\widehat{\gamma}_2^{nd}$,$\widehat{\gamma}_3^{nd}$, and $\widehat{\beta}^{\mathbf{nd}}$ are used to predict Eq. (1) into the SILC data. In section 4 we show that disposable incomes from EUROMOD are different from those in the HBS. One of the reasons might be that not all tax reductions are simulated in EUROMOD, resulting in an upward bias of tax liabilities in EUROMOD, and hence lower disposable incomes. In order to correct for this, we have shifted and rescaled disposable income in EUROMOD such that its mean and variance correspond to the disposable income

[6] This is done because when taking the exponent of $\ln(e_i^{nd})$, the expected value of the error term $\exp(\varepsilon_i)$ is no longer zero. See Wooldridge (2003), pp. 207-210 and 276-280. Assuming normality of the errors, a consistent estimate of the expected value of $\sigma^2 = \exp(\varepsilon_i)$ can be derived from the squared standard error of the regression: $s^2 = SSR/(N-k)$, where SSR is the sum of squared residuals and $(N-k)$ denotes the degrees of freedom.

in the HBS, used for the estimation. This procedure is only done for disposable income as the regressor used in the imputation of total non-durable and durable expenditures. We then again use the original disposable income from EUROMOD to determine residual savings, such that adding-up conditions are fulfilled.

We further observed that for some households the disposable incomes simulated by EUROMOD are negative, for example because of losses incurred by self-employed individuals. Because our models are not defined for negative disposable incomes, and in order to avoid extreme imputed values when disposable incomes are positive but very small, we set imputed expenditures to zero when the regressor containing disposable income is smaller than one. This was done for four of the households in the SILC.

For the remaining covariates, it is assumed that their distributions are similar in both datasets. This will be further discussed in section 4.

Durable Expenditures and Savings

The estimation of durable expenditures is also undertaken in two stages. In the first stage, we take into account that a large number of households report zero spending on durable goods during the reporting period. In the second stage, demand for durable spending is estimated, conditionally on the outcome of the first stage. In both stages we use the same corrected income variable y_i as a covariate as the one which was used for the prediction of total non-durable expenditures.

In detail, in the first stage we use a Probit model to estimate the probability of strictly positive durable spending:

$$Pr(D_i^d = 1) \; = \; \Phi\left(\alpha_0^d \, + \, \gamma_{0,1}^d \, \ln(y_i) \, + \, \gamma_{0,2}^d \, (\ln(y_i))^2 \, + \, \gamma_{0,3}^d \, (\ln(y_i))^3 \; + \, \mathbf{x_i'} \, \beta_\mathbf{0}^\mathbf{d} \, + \, \xi_i\right) \tag{2}$$

where D_i^d denotes a dummy variable which is 1 for household i if demand for durable commodities is positive and zero otherwise. The function $\Phi\,(.)$ is the cumulative distribution function of the standard normal distribution. The covariates in the $\mathbf{x}$-vector are identical to the ones used in Eq. (1).

The coefficient estimates, $\widehat{\alpha}_0^d$, $\widehat{\gamma}_{0,1}^d$, $\widehat{\gamma}_{0,2}^d$, $\widehat{\gamma}_{0,3}^d$, and $\widehat{\beta}_\mathbf{0}^\mathbf{d}$ are used to predict the probability from Eq. (2) into the SILC data. Next a number is drawn from a uniform distribution between 0 and 1. If the number drawn is smaller than the predicted probability, D_i^d is coded as 1 in the imputation, otherwise it is coded as 0.

In the second stage, the demand equation for total household durable commodities is estimated, conditional on the outcome of the first stage. Specifically, for durable spending:

$$\ln(e_i^d) \;=\; \alpha^d \,+\, \gamma_1^d \ln(y_i) \,+\, \gamma_2^d\,(\ln(y_i))^2 \,+\, \gamma_3^d\,(\ln(y_i))^3 \,+\, \mathbf{x_i'}\,\beta^{\mathbf{d}} \,+\, \varepsilon_i \quad \text{if } \widehat{D}_i^d = 1 \tag{3}$$

for households $i = 1, \ldots, N$. The dependent variable $\ln(e_i^d)$ denotes the logarithm of total expenditures for durable consumption. It is defined for all households for which the condition $\widehat{D}_i^d = 1$ holds. All covariates in the $\mathbf{x}$-vector as well as the income variable are identical to those used in the preceding estimations, in Eq. (1) and to Eq. (2). Again, at the imputation, the prediction of $\ln(e_i^d)$ is adjusted by the expected value of $exp(\varepsilon_i)$, and disposable income is corrected in order to better correspond to the distribution in the HBS.

Given estimates for Eqs. (1) and (3), we define savings residually as the difference between the non-corrected disposable income y_i and the sum of total non-durable spending and total durable expenditures:

$$\widehat{s}_i \;=\; \mathrm{y}_i \,-\, \widehat{e}_i^d \,-\, \widehat{e}_i^{nd} \tag{4}$$

where, as in the HBS data, $\widehat{s}_i$ is theoretically unbound in the open interval $]-\infty, +\infty[$, i.e. dissavings are explicitly allowed for.[7]

Commodity Shares from Non-Durable Expenditures

Among the 15 non-durable commodity groups, estimations are conducted differently for two sub-groups. The first sub-group consists of commodities that are typically exposed to many zero-expenditures in the data, which in our case are tobacco, renting, public transport, and education. The second sub-group consists of the remaining 11 non-durable commodities.

The application of a different estimation strategy for these two groups is supposed to deal with the observation of many zero expenditures in the sample for the first-group commodities. Applying two-step estimation techniques can improve the goodness of fit when expenditures are imputed into the target data under these circumstances.

The estimation strategy for the first group closely follows the approach for durable spending (see previous subsection). The estimation proceeds in two steps. In the first step, a Probit model is estimated for the probability of positive demand for the respective commodity:

$$Pr(D_i^k = 1) \;=\; \Phi\left(\alpha_0^k \,+\, \gamma_{0,1}^k \,\ln(e_i^{nd}) \,+\, \gamma_{0,2}^k\,(\ln(e_i^{nd}))^2 \;+\, \mathbf{x_i'}\,\beta_{\mathbf{0}}^{\mathbf{k}} \,+\, \nu_i^k\right) \tag{5}$$

for households $i = 1, \ldots, N$, where D_i^k denotes a dummy variable that is 1 for household i if demand for non-durable commodity k is positive, and zero otherwise. Covariates in the $\mathbf{x}$-

[7] This degree of freedom allows households to have spending exceed their current income temporarily, i.e. either borrow against their future income or run down their assets. Negative savings are in fact observed for a significant number of households in the HBS (see table 4 in the next section) and will also be predicted into the SILC data (Table 5).

vector are again identical to those used in Eq. (1), except for the fact that now no interaction terms are taken into account. The reason for not using the interaction terms here is that doing so actually worsened the resemblance of the imputed expenditures to those observed in the HBS. The term α_0^k denotes a commodity-specific constant, e_i^{nd} denotes total non-durable expenditures and ν_i^k is an error term, assumed to be independent and identically distributed. Eq. (5) is estimated separately for each of the four non-durable commodities among the first sub-group.

Similarly to durable demand, the coefficient estimates, $\widehat{\alpha}_0^k$, $\widehat{\gamma}_{0,1}^k$, $\widehat{\gamma}_{0,2}^k$, and $\widehat{\beta}_0^{\mathbf{k}}$ are used to predict Eq. (5) into the SILC data, for each of the $K = 4$ commodities, assuming that the explanatory variables are distributed similarly in both surveys. Before using total non-durable expenditures (e_i^{nd}) as a covariate, it is corrected in order to more closely resemble the distribution of total non-durable expenditures observed in the Household Budget Survey, just as we explained for disposable income in a previous section. However, if this correction yields a negative value for total non-durable expenditures, as is the case for two observations, the non-corrected variable is used. After imputing the probabilities for positive expenditures, a number is drawn for each commodity from a uniform distribution between 0 and 1. If the number drawn is smaller than the predicted probability, $\widehat{D}_i^k$ is coded 1, and 0 otherwise.

In the second step, we estimate budget share equations for each of the four non-durable commodities in the first sub-group, conditional on the outcome of the first step. Specifically, for households for which demand is estimated to be positive ($\widehat{D}_{ik}^{nd} = 1$), we estimate

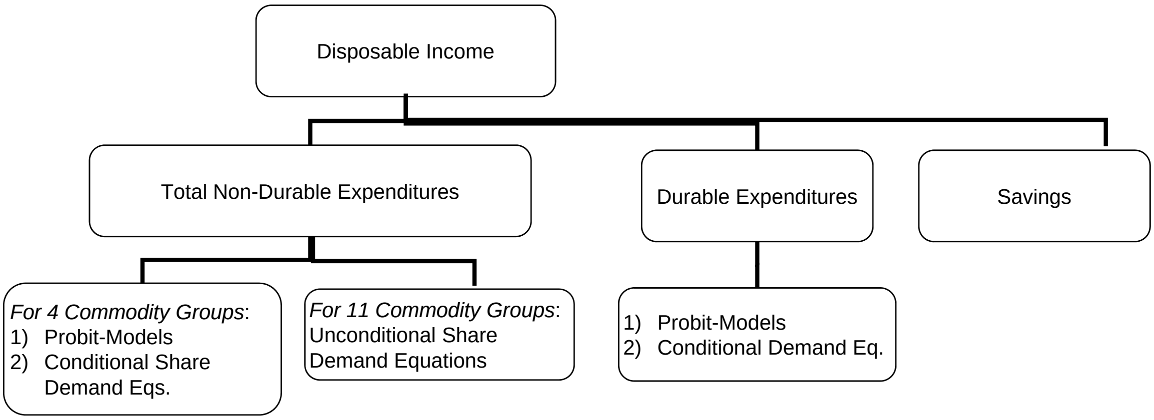
$$w_i^k \;=\; \alpha^k \;+\; \gamma_1^k \ln(e_i^{nd}) \;+\; \gamma_2^k \,(\ln(e_i^{nd}))^2 \;+\; \mathbf{x_i'}\,\beta^{\mathbf{k}} \;+\; \eta_i^k \quad \text{if } \widehat{D}_i^k = 1 \tag{6}$$

for households $i = 1, \ldots, N$ and commodities $k = 1, \ldots, K$. The dependent variable w_i^k is the share of expenditures for non-durable consumption in commodity group k. The latter appears on the right-hand side in a quadratic-log functional form. It is defined for all households for which the condition $\widehat{D}_i^k = 1$ holds. Eq. (6) has again the same demographic controls as in Eq. (5), so no interaction terms are included. Here α^k is a commodity-specific constant, and the commodity-specific error term η_i^k is assumed to be independent and identically distributed. Again, before imputing the expenditure shares into EU-SILC, we adjust total non-durable expenditures such that their distribution resembles the distribution in the HBS.

For some commodities, negative predictions have been set to zero, for reasons of consistency. This was applied for less than 5% of all households. The other budget shares were rescaled, to comply with adding-up conditions. We then calculate expenditures for these four subgroups by multiplying these imputed expenditures shares with the *non-corrected* total non-durable expenditures. This allows us to comply with adding-up conditions.

The estimation strategy for the 11 remaining commodities in the second group is differ-

Figure 1: The Two-Stage Structure of the Model for Household Expenditures



ent from the first group. For these commodities, no special treatment of zero spending is necessary because the population of households with zero spending among each of these commodities is relatively small. Thus, budget share demand equations for the second group are estimated unconditionally, for all households, in a single step:

$$w_i^j \;=\; \alpha^j \,+\, \gamma_1^j \,\ln(e_i^{rnd}) \,+\, \gamma_2^j\,(\ln(e_i^{rnd}))^2 \,+\, \mathbf{x_i'}\,\beta^{\mathbf{j}} \,+\, \eta_i^j \tag{7}$$

for households $i = 1, \ldots, N$ and commodities $j = 1, \ldots, J$. The relevant covariate for the second sub-group (e_i^{rnd}) is now the remaining non-durable spending. This means after spending on the four commodities from the first sub-group has been deducted:

$$e_i^{rnd} \;=\; e_i^{nd} \;-\; \sum_{k=1}^{K} w_i^k * e_i^{nd} \tag{8}$$

Thus, the dependent variable w_i^j in Eq. (6) is the budget share, expressed with respect to non-durable consumption in commodity group j from the *remaining* non-durable expenditures (e_i^{rnd}). The latter again appears on the right-hand side in a log-linear-quadratic functional form. It is defined for all households because for all households it is observed that $e_i^{rnd} > 0$. Eq. (7) has the same demographic controls as the previous estimation equations, now including interaction terms. The term α_j is a commodity-specific constant. The commodity-specific error term η_i^j is assumed to be independent and identically distributed.

The estimated parameters are used to impute expenditures on these remaining categories into EU-SILC. Again a corrected version of the total remaining expenditures is used as covariate, as was done before for total non-durable expenditures and disposable income. In this case the corrected values are never found to be negative, so no further corrections are necessary. For some commodities, negative predictions again have been set to zero, and budget shares were rescaled to comply with adding-up conditions.

The two-stage structure of the entire model is visualized in Figure 1. Each of the three single models at the two stages in Eq. (1)-(7) is estimated on the single cross section for 2009 from the Belgian household budget survey data (HBS, see Section 2).

Table 1: Sample Descriptives – EUROMOD and Budget Survey (HBS)

	EUROMOD		Budget Survey	
	Mean	Median	Mean	Median
Income (Euros):	.		.	
disposable income (not imputed)	2543.2	2168.1	2725.2	2355
Region (Fraction):	.		.	
flanders	0.56	1	0.57	1
wallonia	0.32	0	0.33	0
Age:	.		.	
age head	51.4	50	51.2	51
Demographics (Fraction):	.		.	
male	0.62	1	0.61	1
education of head secondary	0.51	1	0.43	0
education of head high	0.36	0	0.48	0
Household composition:	.		.	
number of household members	2.28	2	2.35	2
number of children	0.37	0	0.41	0
number of HH members working	0.91	1	0.96	1
Economic status head (Fraction):	.		.	
self employed	0.055	0	0.085	0
employed	0.52	1	0.50	1
unemployed	0.065	0	0.076	0
pensioner	0.29	0	0.31	0
Observations	6,132		3,595	

Notes: Income in Euros per month. Disposable income is defined in Chapter 2. Among economic status, the group of nonactives has been omitted as a dummy. Among the regions, Brussels has been omitted as a dummy. Disposable incomes from EUROMOD simulations, switching on take-up correction for income support. Data weighted by population weights. All amounts are downrated to 2009.

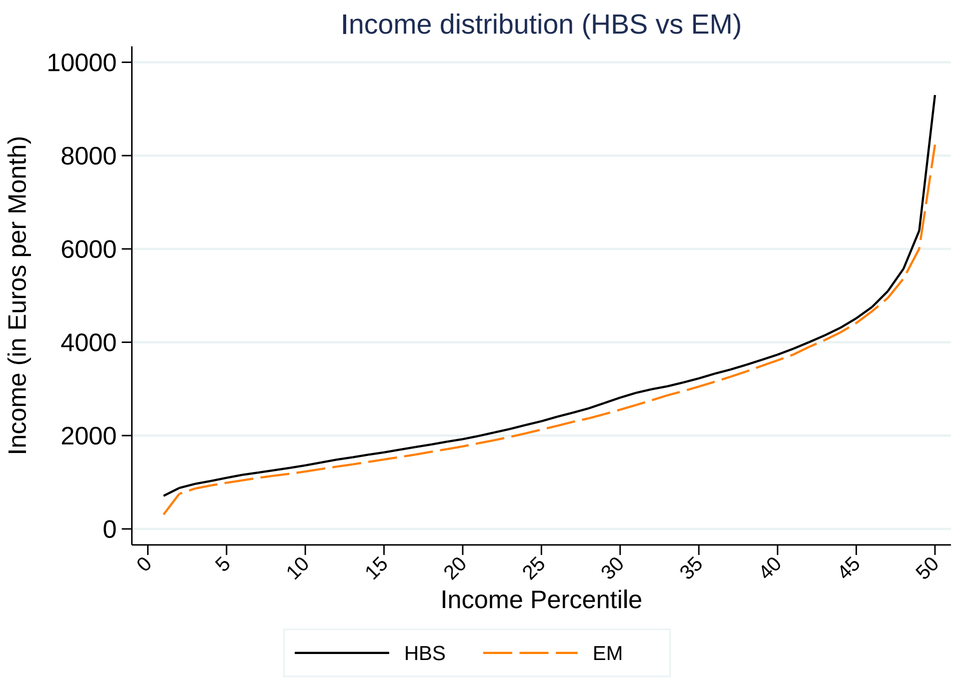
Source: Output of Belgian EUROMOD baseline policy for 2012, using EU-SILC data 2010 as input, and Belgium HBS data (2009).

4. Sample Descriptives

To ensure that the imputations for expenditures are not biased, it is important that the income variable and the demographic variables that have been applied as controls in the $\mathbf{x_i}$-vectors of the estimated demand equations are similarly distributed in the source data set (HBS) and in the target data set (EU-SILC data). Table 1 shows descriptive statistics on these demographic variables and on the income variable from the two data sets. It is obvious that all applied demographic controls have mean and median values of largely the same size in the two data sets[8].

However, median monthly disposable household income in the budget survey (HBS) amounts to 2,355 Euros, which exceeds considerably disposable income as simulated by EUROMOD: 2,168 Euros. Mean incomes show similar differences in the HBS (2,725 Euros) and EUROMOD (2,543 Euros).[9] We found a better correspondence when we compared disposable incomes between the HBS and the original SILC data, implying that the difference is probably caused by EUROMOD. Indeed, a number of tax deductions and tax credits have not been implemented in EUROMOD, because insufficient information is available in the SILC data. This difference between the distributions is also visible in (Figure 2). Over the entire distribution, and slightly more so at the top, there is an under-simulation of household incomes in EM compared to the reported HBS incomes.

Figure 2



[8]Individual demographic variables, such as age, gender and education, are given for the head of household. Since we noticed that the distributions for the given variables were rather different when we looked at the reference persons in both surveys, we decided to determine the head-of-household following the same procedures as those followed in EUROMOD: the head of the household is the person with the highest disposable income; if this results in several persons with equal disposable income, the oldest among them will be designated head of household; if still no unique head-of-household has been determined, the person with the lowest identifier in the dataset is picked among the remaining candidates. Following this identical procedure in both datasets, gives better correspondance between the sample descriptives.

[9]Note that disposable incomes from the HBS are reported values, whereas in EM they are simulated incomes. In the program version for 2012, the EM simulations for Belgium apply a correction for non-take-up of income support, assuming that a non-negligible fraction of eligible households do not claim receipt of these benefits. For details, see Hufkens *et al.* (2013).

Mean and median age of the household head are almost identical (51) in both data sets. For the other demographic variables, only the mean can be interpreted meaningfully. Both datasets contain about 61% male household heads. Household heads in the HBS report slightly higher education rates than in the SILC, and there are 3 percent points more self employed persons. Overall, the demographic variables correspond rather well.

5. Results for the Estimations

The full estimation output is available in appendix C. Although many of the covariates used are not significant in the regression tables, their presence significantly improves the quality of the imputations, and thus have been retained. In the remainder of this section, we will discuss the expenditure elasticities that can be derived from the estimation results.

5.1. Expenditure elasticities

We can estimate the expenditure elasticity of the non-durable consumption categories using the expenditure data in the household budget survey. We define this elasticity as follows for commodity group k:

$$\eta_i^k \equiv \frac{d \log e_i^k}{d \log e_i^{nd}}$$

where e_i^k denotes expenditures of household i on good k. This can be rewritten as follows:

$$\eta_i^k = \frac{d \log w_i^k}{d \log e_i^{nd}} + 1 \tag{9}$$

where w_i^k here denotes the budget share of good k out of total non-durable expenditures e_i^{nd} for household i. In order to determine these elasticities for all commodity groups, we cannot use the estimation coefficients obtained in section 3, because for most groups expenditure shares out of total non-durable expenditures minus expenditures on groups with many zeroes, were used. In order to obtain elasticities for all commodity groups, we estimate the following equation:

$$w_i^k \;=\; \alpha^k \;+\; \gamma_1^k \ln(e_i^{nd}) \;+\; \gamma_2^k (\ln(e_i^{nd}))^2 \;+\; \mathbf{x_i'} \beta^{\mathbf{k}} \;+\; \nu_i^k \tag{10}$$

where now no interaction terms are included in the covariates $\mathbf{x_i'}$. As usual, the error term ν_i^k is assumed to be identically and independently distributed. Using Eq. (9) we can use the estimated coefficients in order to determine for each household i the expected expenditure elasticity, given total non-durable expenditures:

$$\eta_i^k = 1 + \frac{\gamma_1^k + 2\gamma_2^k \ln(e_i^{nd})}{w_i^k}$$

For each quintile of equivalized total non-durable expenditures and for each commodity group, we plug into this formula the average expenditures and budget shares. We did the same for the entire population. The resulting elasticities are reported in table 2. We see that food and non-alcoholic beverages, water and energy products, rents, communication and public transport are necessities, and more so for the higher quintiles. Also tobacco is a necessity, although it is less so for the higher quintiles. Note that we calculated the elasticities for the entire population, not just for those who have positive expenditures on the commodities under investigation.

From the table it also follows that clothing, health care and education are luxury goods: higher quintiles spend larger budget shares on them. Among the luxury goods are also the usual suspsects, such as restaurants and hotels, recreation and culture and private transport. Also household services and other goods and services are luxuries.

Table 2: Elasticities w.r.t. non-durable expenditures in HBS

	Q1	Q2	Q3	Q4	Q5	ALL
Food, non-alcoholic beverages	0.65	0.59	0.52	0.48	0.32	0.53
Alcoholic beverages	1.03	1.05	1.05	1.07	1.08	1.06
Tobacco	0.65	0.68	0.71	0.72	0.79	0.72
Clothing and footwear	1.61	1.49	1.40	1.43	1.40	1.46
home fuels, electricity and water	0.56	0.45	0.40	0.33	0.26	0.44
Rents (excluding imputed rents)	0.97	0.90	0.84	0.80	0.73	0.84
Household services	1.19	1.20	1.21	1.21	1.20	1.21
Health	1.45	1.27	1.20	1.14	1.06	1.20
Private transport	1.87	1.58	1.42	1.33	1.23	1.42
Public Transport	0.87	0.75	0.76	0.74	0.70	0.76
Communication	0.76	0.65	0.57	0.47	0.31	0.56
Recreation and culture	1.28	1.21	1.20	1.19	1.18	1.21
Education	0.73	1.38	1.88	2.20	2.52	1.89
Restaurants and hotels	2.02	1.85	1.72	1.70	1.62	1.76
Other goods and services	1.06	1.05	1.05	1.04	1.04	1.05

Notes: Quintiles for equivalized non-durable expenditures, using OECD scale.
Data weighted by population weights.
Source: Own calculations based on HBS data (2009).

6. Results for the Imputations

In this section we present results for the imputations of expenditures into EUROMOD. First we compare imputed expenditures on the different non-durable commodity groups and on durable goods to reported HBS spending in terms of fit along their distributions. Next we evaluate the imputations across the distribution of household disposable income between the two data sets.

6.1. Distributions of Expenditures

We compare imputed expenditures with the reported expenditures in the source data (HBS) in order to evaluate the performance of the imputation. We start by looking only at mean and median values. Table 3 reports the imputed expenditures for EUROMOD (columns 1 and 2) and the corresponding variables for the HBS (columns 3 and 4).

Table 3: Spending Descriptives – EUROMOD (Imputed) and Budget Survey (HBS)

	EUROMOD		Budget Survey	
	Mean	Median	Mean	Median
Income (Euros):	.		.	
disposable income (not imputed)	2553.1	2172.5	2725.2	2355
Spending (Euros)	.		.	
Total non-durable expenditures	2221.5	2050.6	2232.0	1984.6
Durable expenditures	288.4	302.1	267.0	32.9
Savings (%):	.		.	
Household Savings Rate	-8.71	-5.25	-0.63	9.26
Fraction with Pos. Exp. (%):	.		.	
Smoking	26.0	0	24.4	0
Renting a dwelling	65.7	100	61.7	100
Using public transport	28.6	0	29.9	0
In Education	16.8	0	17.1	0
Non-durab. Exp. Shares (%):	.		.	
Food, non-alcoh.	18.0	17.6	18.2	17.2
Alcoholic beverages	1.86	1.81	1.95	0.82
Tobacco	1.01	0	0.95	0
Clothing and footwear	5.07	4.85	5.03	2.71
Home fuels, electricity and water	10.1	9.30	10.2	8.64
Rents (excl. imputed rents)	11.0	10.7	9.78	1.19
Household services	2.51	2.38	2.53	1.35
Health	6.36	5.81	6.34	4.07
Private transport	8.01	8.18	8.22	5.32
Public Transport	0.63	0	0.64	0
Communication	4.06	3.86	3.88	3.20
Recreation and culture	8.68	8.78	8.50	6.90
Education	0.78	0	0.78	0
Restaurants and hotels	9.30	9.59	9.52	6.49
Other goods and services	12.7	13.2	13.5	11.7
Observations	6,112		3,595	

Notes: Income and expenditure variables in Euros per month. Data weighted by population weights. All amounts are downrated to 2009.

Source: Own calculations using the EU-SILC data 2010 for Belgium and HBS data (2009).

Total non-durable expenditures have very similar mean values in the HBS (2,232 euros) and when imputed into EUROMOD (2,222 euros). The difference is similarly small for durable spending (267 euros and 288 euros, respectively).

At the median these expenditure variables differ more between the two data sets, which is naturally related to the fact that expenditures have been imputed into the EM data based on regressions, by which a significant share of the variance is lost so that as a result the distribution tends to be smoother for the imputed variables compared to the observed variables.

The residual household savings rate is higher in the HBS (-0.63%) than in the imputation (-8.71%). This is because observed disposable incomes in the HBS are higher than the ones simulated by EUROMOD. Because in our model saving is a mere residual, this is not important for our purposes. The imputed probabilities for belonging to the four zero-spending commodities are fairly close to the observed probabilities in the budget survey. For example, the probability for a household to be observed as renting a dwelling in the HBS is 61.7 percent and it is imputed to be 65.7 percent in EM. Differences are smaller for the other three commodities from this zero-spending subgroup. Note that the median values have little significance here, as they are just true or false.

Mean budget shares for non-durable expenditures are also very similar between the HBS and EM, for all of the 15 commodities. For example, the observed share for food is on average 18.2 percent in the HBS and it has been imputed to 18.0 percent in EM. The observed share for home fuels and electricity is 10.2 percent, and the imputed share is 10.1 percent. Only for rents there is a difference slightly larger than one percent point.

Tables 4 and 5 show the distributions of expenditures in more detail for the commodity groups that have been imputed into EUROMOD, in order to point out which groups are over- or under-imputed in the tails of the distributions. Table 4 presents the distributions of the observed expenditures from the HBS data, and Table 5 the respective predicted expenditure distributions that have been imputed into EM.

Taking a look at the distributions of the spending variables observed in the budget survey reveals that total non-durable expenditures and total durable expenditures are distributed very differently. The distribution of total non-durable expenditures (Figure A.1 in Appendix A) increases largely in a linear form, with a convex shape at the higher end of the distribution. There are no zeros observed at total non-durable spending. They increase almost linearly until about the 80th percentile. At the top of the distribution, the curve increases significantly with a convex shape.

On the contrary, the distribution of total durable expenditures (Figure A.2 in Appendix A) looks totally different. It is almost flat across large parts of the distribution and spikes heavily at the very top of the upper tail.

Spending for durable goods is a rare event. In the relatively short interview period of three months, many respondents will be observed as not purchasing a durable good. However, because we aggregate various durable commodities under the group of durable goods here,

Table 4: Spending Descriptives – Distribution in the Budget Survey (HBS)

	p1	p5	p10	p25	p50	p75	p90	p95	p99
Income (Euros):	.								
disposable income (not imputed)	725.0	967.0	1128.0	1534.8	2355.0	3513.0	4603.3	5561.4	8687.0
Spending (Euros)	.								
Total non-durable expenditures	575.0	801.7	983.7	1341.4	1984.6	2798.4	3741.0	4544.6	6510.2
Durable expenditures	0.0	0.0	0.0	4.2	32.9	130.0	445.5	1323.5	4707.4
Savings (%)	.								
Household Savings Rate	-213.4	-89.3	-50.6	-17.0	9.3	30.5	47.3	57.3	71.1
Fraction with Pos. Exp. (%):	.								
Smoking	0.0	0.0	0.0	0.0	0.0	0.0	100.0	100.0	100.0
Renting a dwelling	0.0	0.0	0.0	0.0	100.0	100.0	100.0	100.0	100.0
Using public transport	0.0	0.0	0.0	0.0	0.0	100.0	100.0	100.0	100.0
In Education	0.0	0.0	0.0	0.0	0.0	0.0	100.0	100.0	100.0
Non-durab. Exp. Shares (%):	.								
Food, non-alcoh.	4.9	7.6	9.2	12.6	17.2	22.8	28.6	32.3	41.1
Alcoholic beverages	0.0	0.0	0.0	0.0	0.8	2.6	5.4	7.6	14.0
Tobacco	0.0	0.0	0.0	0.0	0.0	0.0	3.5	6.5	12.2
Clothing and footwear	0.0	0.0	0.0	0.3	2.7	7.5	13.1	17.9	29.0
Home fuels, electricity and water	2.1	3.3	4.2	5.9	8.6	12.9	17.9	22.0	32.3
Rents (excl. imputed rents)	0.0	0.0	0.0	0.0	1.2	18.3	31.7	37.8	51.2
Household services	0.0	0.1	0.2	0.6	1.4	3.0	5.9	8.9	16.5
Health	0.0	0.0	0.0	1.2	4.1	8.3	15.4	21.1	37.1
Private transport	0.0	0.0	0.0	2.2	5.3	11.2	20.5	27.0	39.5
Public Transport	0.0	0.0	0.0	0.0	0.0	0.3	1.7	3.6	10.2
Communication	0.0	0.0	0.3	1.5	3.2	5.3	8.0	10.3	16.1
Recreation and culture	0.0	1.2	1.9	3.7	6.9	11.4	17.4	21.9	31.4
Education	0.0	0.0	0.0	0.0	0.0	0.0	1.6	4.9	16.6
Restaurants and hotels	0.0	0.0	0.0	1.9	6.5	14.0	23.4	29.3	45.3
Other goods and services	1.1	3.3	4.6	7.6	11.7	17.2	23.9	29.5	44.0
Observations	3,595								

Notes: Income and expenditure variables in Euros per month. Data weighted by population weights. All amounts are downrated to 2009.
Source: Own calculations using the EU-SILC data 2010 for Belgium and HBS data (2009).

the number of households with zero durable consumption is significantly reduced. In fact, only about 20% of all households remain with zero durable consumption[10]. Nevertheless, the majority of the households have a relatively low level of durable consumption. The distribution is almost flat until the 60th percentile, and it only increases marginally until about the 90th percentile. From then on, the increase is slightly greater, and it is followed by a heavy spike at the top end of the distribution. As a result, finding a good fit for the estimation of total durable consumption turns out to be more difficult than for non-durable consumption, due to its largely skewed distribution, even when conditioning on positive durable consumption. Overall, comparing distributions of spending variables, we see that much of the variance of expenditures in the HBS is lost in the imputation.

[10] A better result could possibly have been achieved by looking at imputed usage costs of the durable goods instead of looking at the entire purchase price of the durable goods at a single moment of time. We have not applied this procedure for Belgium.

Table 5: Spending Descriptives – Distribution in EUROMOD (Imputations)

	p1	p5	p10	p25	p50	p75	p90	p95	p99
Income (Euros):	.								
disposable income (not imputed)	498.7	879.3	1022.4	1389.7	2172.5	3377.1	4524.4	5372.7	7437.6
Spending (Euros)	.								
Total non-durable expenditures	737.5	969.3	1096.3	1400.2	2050.6	2851.1	3551.7	4039.8	5349.2
Durable expenditures	0.0	0.0	0.0	84.2	302.1	449.1	542.6	609.0	823.7
Savings (%)	.								
Household Savings Rate	-62.4	-32.0	-24.8	-14.9	-5.3	5.2	15.3	22.0	37.3
Fraction with Pos. Exp. (%):	.								
Smoking	0.0	0.0	0.0	0.0	0.0	100.0	100.0	100.0	100.0
Renting a dwelling	0.0	0.0	0.0	0.0	100.0	100.0	100.0	100.0	100.0
Using public transport	0.0	0.0	0.0	0.0	0.0	100.0	100.0	100.0	100.0
In Education	0.0	0.0	0.0	0.0	0.0	0.0	100.0	100.0	100.0
Non-durab. Exp. Shares (%):	.								
Food, non-alcoh.	10.3	12.3	13.5	15.3	17.6	20.3	23.0	24.8	29.4
Alcoholic beverages	0.0	0.5	0.7	1.2	1.8	2.5	3.1	3.3	3.8
Tobacco	0.0	0.0	0.0	0.0	0.0	1.8	4.3	5.1	6.2
Clothing and footwear	0.6	1.9	2.5	3.5	4.9	6.5	8.0	8.8	10.2
Home fuels, electricity and water	4.6	6.0	6.5	7.5	9.3	12.2	14.9	16.3	20.0
Rents (excl. imputed rents)	0.0	0.0	0.0	0.0	10.7	18.4	24.9	28.7	37.1
Household services	0.7	1.4	1.5	1.9	2.4	3.0	3.6	4.2	5.2
Health	1.1	2.8	3.6	4.6	5.8	7.7	9.9	11.4	14.9
Private transport	0.0	0.1	1.5	4.6	8.2	11.5	13.7	15.1	17.6
Public Transport	0.0	0.0	0.0	0.0	0.0	1.1	2.5	3.1	4.0
Communication	1.2	1.8	2.2	2.9	3.9	5.1	6.2	6.8	8.6
Recreation and culture	1.5	4.8	5.9	7.4	8.8	10.0	11.5	12.2	14.3
Education	0.0	0.0	0.0	0.0	0.0	0.0	3.6	5.9	9.5
Restaurants and hotels	0.0	2.0	3.5	6.5	9.6	12.1	14.5	15.8	18.8
Other goods and services	2.0	6.3	8.6	11.3	13.2	14.9	16.1	16.8	18.4
Observations	6,112								

Notes: Income and expenditure variables in Euros per month. Data weighted by population weights. All amounts are downrated to 2009.
Source: Own calculations using the EU-SILC data 2010 for Belgium and HBS data (2009).

6.2. Spending across the Income Distribution

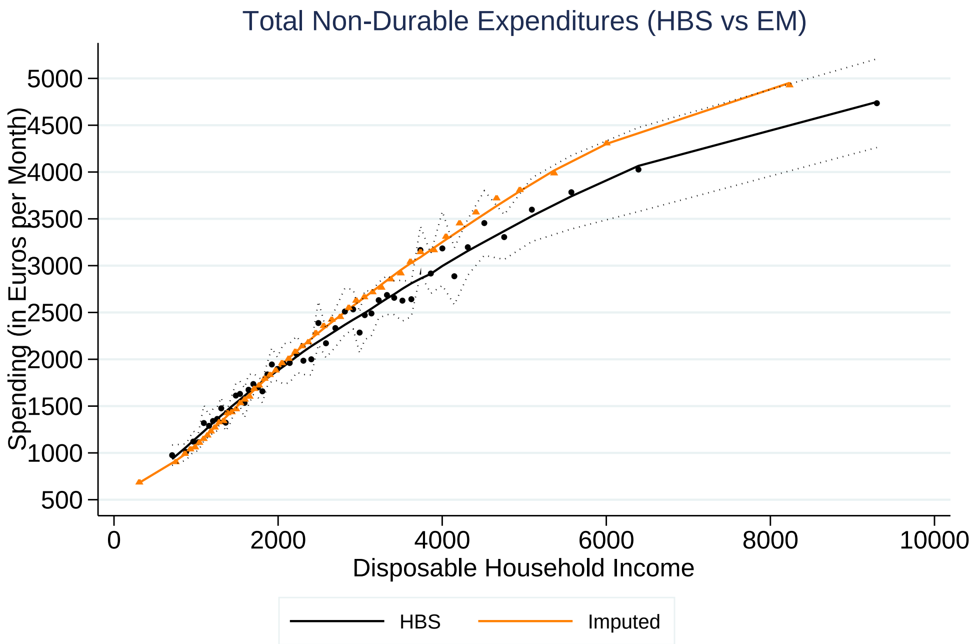
Since the main aim of extending EUROMOD to incorporate VAT, is to perform distributional analysis of indirect tax changes, it is important to get the expenditure patterns across the distribution right. In this subsection, we therefore evaluate the single spending variables across the distribution of disposable household income.

Figure 3 plots total non-durable spending, Figure 4 durable spending, and Figure 5 the resulting residual savings variable across the income distribution. Respective plots for the 15 non-durable commodity shares are relegated to Appendix A.[11]

The plot for total non-durable spending (Figure 3) looks very similar to the respective plot across the spending distribution: spending increases almost linearly in income up to the 85th percentile, after which spending rises faster in consecutive percentiles. The fit of the imputations appears to be very good, although it is slightly shifted to the left because of the corrections to disposable income described in section 3. The respective plot for durable

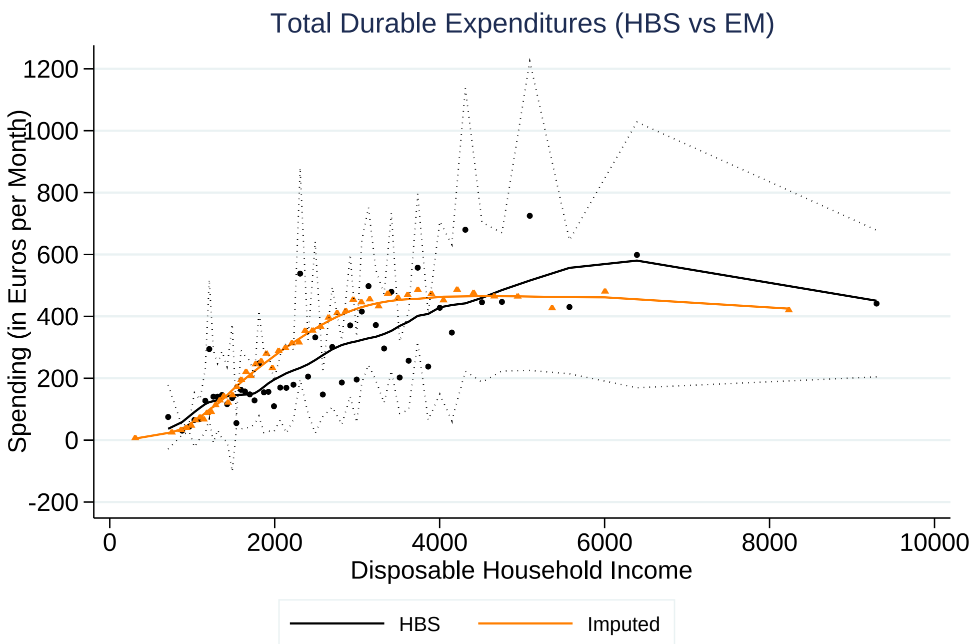
[11]Confidence intervals in all figures for the income distribution are 5% confidence intervals for the mean respective spending value at the respective percentile of the income distribution. Each dot in the figure represents two percentiles. All amounts are downrated to 2009.

Figure 3



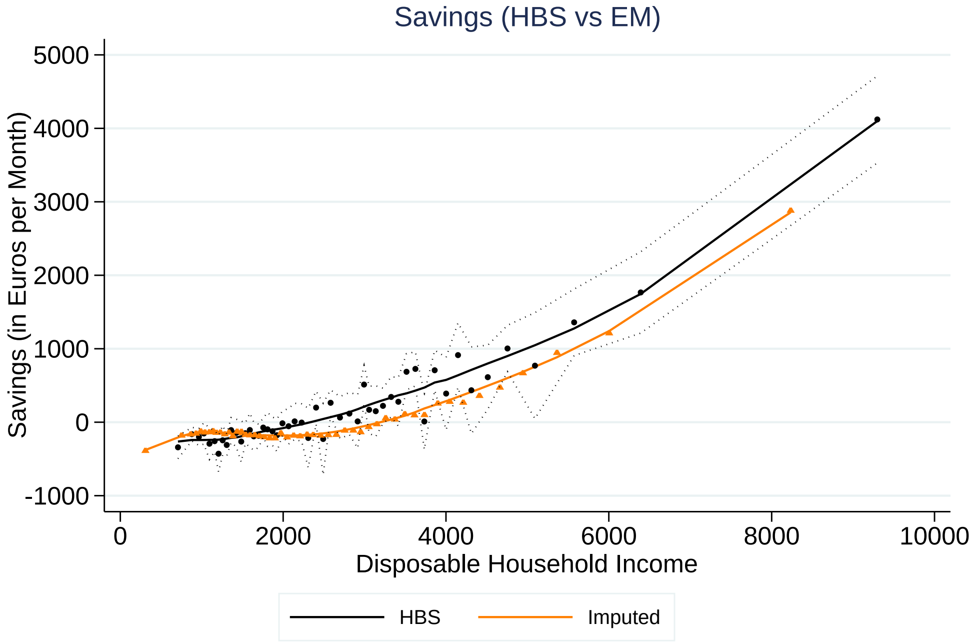
Notes: Weighted by population weights. Based on Belgian HBS data (2009) and EUROMOD output (2012 policy, based on EU-SILC 2010).

Figure 4



Notes: Weighted by population weights. Based on Belgian HBS data (2009) and EUROMOD output (2012 policy, based on EU-SILC 2010).

Figure 5



Notes: Weighted by population weights. Based on Belgian HBS data (2009) and EUROMOD output (2012 policy, based on EU-SILC 2010).

spending (Figure 4) also looks very good, with a similar shift to the left. As both non-durable and durable spending look good in the imputations, residual savings also fit quite well across the income distribution, again slightly shifted to the left (Figure 5).

The same holds for imputations on most of the 15 non-durable commodity shares plotted in Appendix A. Naturally, variation in the HBS is somewhat greater for most commodities. Still, for most commodities, imputations into EM look relatively good across the income distribution when compared to the respective observed shares in the HBS. Only expenditure shares on communication fall faster with income than observed in the HBS, although they stay near the 95% confidence interval for the highest quantile.

7. VAT Simulations

Based on these expenditure imputations, simulations of VAT within EUROMOD are now possible. We first present results for VAT simulations for the baseline tax legislation. We show aggregate revenues and incidence analysis across the income distribution. Next, we present results for two ad hoc policy reforms, and we discuss effects on VAT incidence. We start by moving all non-zero rates to the standard rate of 21%, and look at the distributional and budgetary impact. Then we introduce a compensating reform, which balances the government budget, but does not compensate the distributional effects. These simulations are chosen as mere illustrations of what the model can do. For use in policy debates, it may be interesting for example to introduce compensating reforms in the benefit system, in order to also compensate for the distributional effects. Another potentially interesting simulation would be to simulate a move from social security contributions to higher VAT rates, as is recurrently suggested in Belgium.

In preceding sections, all simulated amounts were downrated to the HBS year 2009 in order to enable the comparison between the imputations and the HBS. In the current section, all amounts are uprated to 2012, since this is the policy year used in EUROMOD. This also enables the comparison with national account data.

Note again that relative price effects are not taken into account in the simulations, only real income effects are incorporated. Taking relative price effects into account necessitates the estimation of a complete demand system, based on sufficient relative price variation. We do not have this at our disposal in the budget survey. Moreover, we think that the effects of incorporating relative price effects on the distributional analysis would only be of second order.

7.1. VAT incidence in the baseline

This section addresses incidence analysis of the baseline VAT rate structure across households for Belgium. The baseline structure of VAT rates refers to the policy year 2012. In the baseline, the regular VAT rate in Belgium is 21 percent, there are reduced rates of 6% and 12%, and there are a number of exempt goods. In our model, where the production sector is absent, we can safely assume that VAT-exempt goods are taxed at a rate of 0 percent. For some of these goods input tax deduction is not allowed, while for others it is for social or cultural reasons.

Table 6: VAT Rates and Aggregate Revenues in the Baseline (HBS)

	Total	0 %	6 %	12 %	21 %
Expenditures (bn Euros):	148.8	36.1	30.5	7.2	75.0
VAT Revenues (bn Euros):	15.5	0.0	1.7	0.8	13.0
VAT/(EXP-VAT) (in %):	11.6	0.0	6.0	12.0	21.0
Expenditure Share (in %):	100.0	24.3	20.5	4.8	50.4
VAT Share (in %):	100.0	0.0	11.1	5.0	83.9

Notes: Data weighted by population weights.
Source: Own calculations using the HBS for Belgium, i.e. HBS data (2009), uprated to 2012.

To give a picture of the statutory rates in the baseline, we first calculate VAT liabilities on the HBS itself. The results are shown in Table 6. Of a total of 149 bn euros, including VAT, households in Belgium spend the greatest part (75 bn euros or 50%) on commodities that are taxed at the regular rate. The other half is spend on goods that are VAT-exempt (36 bn euros or 24%) and goods that are taxed at the reduced rate of 6% (31 bn euros or 21%) or 12% (7 bn euros or 5%). This information is derived from the HBS for Belgium for 2009, uprated to 2012.

Overall, VAT revenues observed in the Belgium HBS for private households sum up to 15.5 bn euros for the year 2012. 13 bn euros (84 percent) are related to commodities taxed at the regular rate, 1.7 bn euros and 0.7 bn euros refer to the 6% and 12% rates respectively. VAT revenues from national accounts for 2012 aggregate up to 26.8 bn euros. As a result, about 58 percent of aggregate VAT revenues are simulated in EUROMOD. The major reason for this discrepancy is the fact that several groups that pay significant amounts of VAT are not covered in HBS. Among these groups are private households that are not covered by the HBS, such as people in dormitories, jails, or retirement homes. Also VAT paid by the government is not included in the HBS. There are also hospitals and business enterprises such as financial companies that are themselves exempt from VAT but have to pay the input VAT from all previous production stages.

Given the statutory VAT rates, implicit or effective VAT rates can be calculated. The implicit tax rate relates the tax liability to net total spending, i.e. spending excluding tax liability. Statutory rates are given in terms of producer prices, not consumer prices. The resulting implicit VAT rate in Belgium, on average for all commodities, is 11.6 percent.

Table 7 breaks the implicit tax rate down by the different commodity categories and relates them to gross spending and absolute as well as relative VAT revenues. These figures are based on the imputations of expenditures into EUROMOD. This is why there is a difference between the figures on aggregate spending and VAT revenues imputed into EUROMOD in Table 7, and the respective figures in Table 6 that relate to observed values from the HBS. Both aggregate expenditures and VAT revenues are about 3 percent higher when imputed into EUROMOD. These relatively small differences are caused by deviations at the imputation of the different expenditures, as discussed in section 3.

Aggregate VAT revenues for the population of private households, as simulated in EUROMOD, sum up to 15.9 bn euros for Belgium in 2012, and aggregate consumption expenditures for durable and non-durable goods sum up to 152.5 bn euros). Related to aggregate VAT revenues from national accounts (26.8 bn euros), VAT revenues simulated in EUROMOD make up about 59 percent when summed up for the population covered in EUROMOD. This coverage is about the same as for the HBS. It is still significantly lower than 100 percent, which is due to the same reasons as apply to the HBS, namely that several groups that pay significant amounts of VAT are not covered in the surveys.

A large share of aggregate VAT revenues simulated in EUROMOD relate to durable goods (19%). This is followed by electricity (12%) and private transport (11%). All other commodity groups account for less than 10 percent of overall revenues each, with lowest shares

Table 7: Implicit Tax Rates in the Baseline (EUROMOD – Imputed)

	Expenditures incl. VAT (mln Euros)	VAT Revenue (mln Euros)	Implicit Tax Rate (Baseline)	Share in VAT Revenue (%)
Total	152,462.0	15,972.6	11.7	100.0
Food	1,915.7	110.3	6.1	8.3
Alcohol	209.3	36.2	20.9	2.7
Tobacco	99.6	17.3	21.0	1.3
Clothing	644.0	110.9	20.8	8.3
Electricity	994.7	159.1	19.0	12.0
Rents	1,030.5	54.1	5.5	4.1
HH Services	278.2	31.6	12.8	2.4
Health	693.0	22.6	3.4	1.7
Prv. Transport	1,069.4	153.3	16.7	11.5
Pub. Transport	73.8	4.2	6.0	0.3
Communication	393.4	65.9	20.1	4.9
Recreation	1,034.9	118.0	12.9	8.9
Education	141.6	0.0	0.0	0.0
Restaurants	1,159.0	122.9	11.9	9.2
Others	1,509.3	71.5	5.0	5.4
Durables	1,458.8	253.2	21.0	19.0

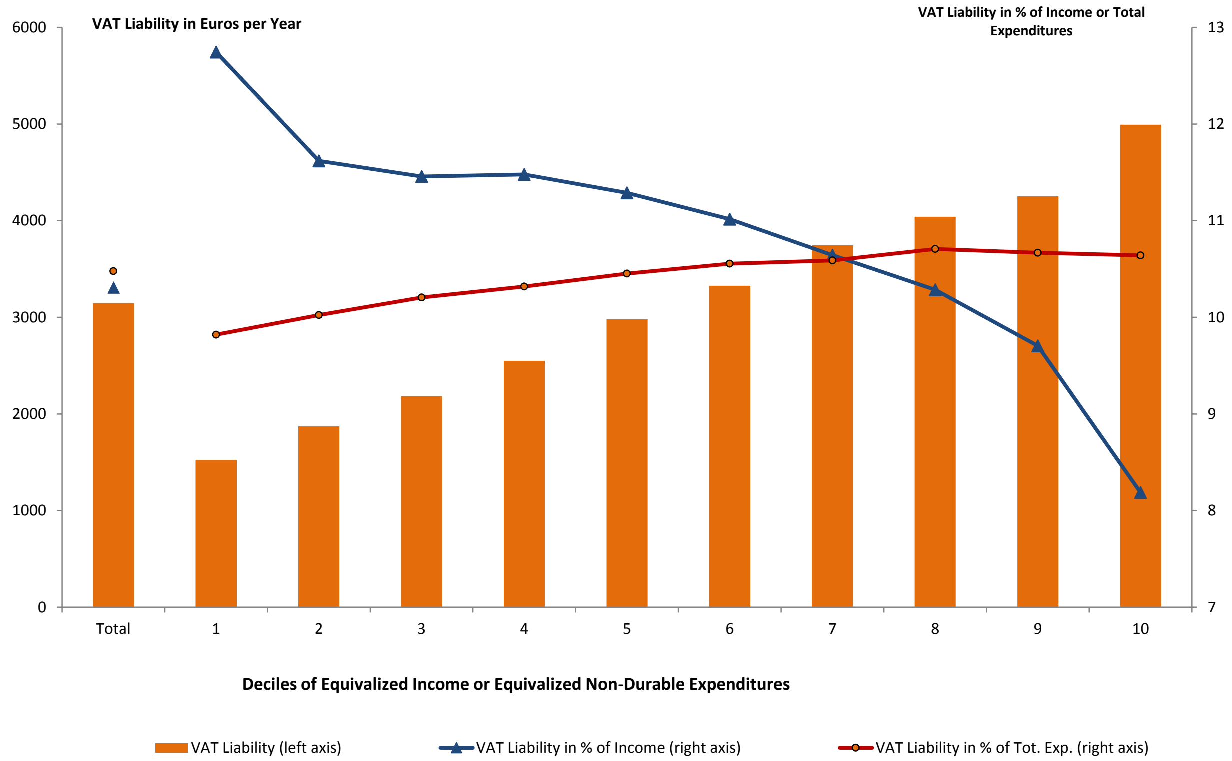
Notes: Data weighted by population weights.
Source: Own calculations using EUROMOD 2012 policies for Belgium based on EU-SILC data 2010 and imputations based on HBS data (2009).

out of revenues being levied on education (0%) and public transport (0.3%).

Table 7 also contains implicit VAT rates, broken down by commodity group. While the average implicit VAT rate is 11.7 percent (about the same as in the HBS), the implicit rate varies largely across the single commodity groups, according to their composition with respect to goods taxed at the regular VAT rate and at the reduced rates. Commodity groups that consist only of goods that aref taxed at the regular rate, such as alcohol, tobacco, clothing, or communication have an implicit VAT rate that matches the regular rate of 21 percent. Groups that consist of some goods that are taxed at one of the reduced rates have an implicit VAT rate that is lower than the regular rate, such as private transport, restaurants or household services. Commodity groups that largely consist of goods with reduced rates have an implicit rate around 6 percent (food, public transport), and groups with mostly exempt goods have implicit rates near zero (education, health care).

Now we turn to the analysis of the VAT incidence. The distribution of VAT liability for the baseline VAT legislation is presented in Figure 6 across the respective distributions of household income and total non-durable spending. Decile median VAT payments are plotted in Euros per year as well as related to income or total non-durable spending in percent. The deciles are based on equivalized incomes, respectively non-durable expenditures, weighted by the OECD-modified scale.

Figure 6: VAT Incidence in the Baseline



Notes: Weighted by population weights. Based on HBS data (2009) and Belgian EUROMOD output (2012 policy, based on EU-SILC 2010). Deciles for nominal VAT liability and VAT liability in % of total expenditures are based on equivalized non-durable expenditures. Deciles for liability in % of income are based on equivalized disposable income.

We can see from Figure 6 that VAT payments increase in income in absolute terms. On average across all deciles, households pay over 3,150 euros per year for VAT. While households in the lowest decile only pay 1,540 euros, households in the 10th income decile pay about 5,100 euros.

However, when we consider VAT payments in percent of household net income (right axis of Figure 6) the picture suggests that VAT is regressive. Tax liabilities decrease in income in relative terms. They amount to some 13% of income in the 1st income decile and decrease down to some 8% for the 10th decile. On average, households pay about 10.3% of their net income for VAT.

If on the contrary, VAT liabilities are related to total spending, instead of net income, the picture turns around (Figure 6). Now VAT is slightly progressive, in the sense that tax liabilities increase in spending in relative terms. Households in the lowest expenditure decile spend 9.7% of their total expenditures on VAT, whereas households in the highest expenditure decile pay almost 10.6%. The variation though is not as great as it is when related to income and plotted across the income distribution. The average household pays some 10.5% of total spending for VAT.

The reason for the different pictures when VAT liabilities are plotted against the income distribution and the spending distribution lies in the distribution of savings, i.e. that part of current income that is not spent in the current period but saved for future consumption. It is not subject to VAT in the period under consideration. Because households with higher incomes tend to save a greater part of their income (see Figure 5), VAT liabilities make up a smaller fraction of income for rich households than for poor ones. Once the picture abstracts from savings and relates VAT liabilities to total spending only, we see a slight increase in tax liabilities for higher spending deciles. This increase is related to the fact that poor households tend to spend a larger part of their total expenditures on commodities that are VAT-exempt or subject to reduced VAT rates, such as food.[12]

7.2. Uniformity without Compensation (Reform A)

In order to demonstrate the use of our model, we now simulate a number of possible reforms. We start by moving the reduced VAT rates towards the regular rate, without introducing any compensating measures, and keeping zero rates and exemptions (Reform A).

Table 8 displays the changes of the implicit tax rates resulting from this reform[13]. The overall implicit VAT rate increases by 3.6 percentage points to 15.3 percent. This implies an overall price increase by 3.2 percent. Broken down by commodity groups, we can see that uniformity implies price increases for groups that consist of goods that are taxed at reduced rates in the baseline tax legislation. Prices increase more than on average for groups that have relatively large shares of these goods, such as food and public transport, whereas increases are lower for groups that have only smaller fractions of reduced-rate goods, such as recreation, restaurants, or health. Prices are not affected for commodity groups that do not contain any reduced-rate goods, such as durables and tobacco, which only contain goods that are taxed at the regular VAT rate. There is also no price effect on groups that consist only of regular-rate goods and VAT-exempt goods, such as rents, household services, and

[12]Similar pictures for VAT incidence in Germany are reported e.g. in Adam *et al.* (2011). Interestingly, however, Adam *et al.* (2011) also find that the slightly progressive effect of VAT when plotted against the spending distribution does not occur in all countries. In many countries, VAT is largely distributionally neutral when related to expenditures.

[13]In order to calculate the new implicit tax rates, the new VAT rates are applied at the detailed level of the 1300 commodities, assuming constant producer prices. Assuming that budget shares on the different commodities within the commodity groups remain constant, the new implicit rates on the groups can be calculated as the proportions of the new total tax liabilities and expenditures on the groups. This is explained in more detail in Decoster and Spiritus (2014).

education.

Table 8: Implicit Tax Rates in the Baseline and Under Uniformity

	Implicit Tax Rate (Baseline)	Implicit Tax Rate (Uniformity)	Implied Price Change (%)
Total	11.7	15.3	3.2
Food	6.1	21.0	14.0
Alcohol	20.9	21.0	0.1
Tobacco	21.0	21.0	0.0
Clothing	20.8	21.0	0.2
Electricity	19.0	21.0	1.6
Rents	5.5	5.5	0.0
HH Services	12.8	12.8	0.0
Health	3.4	7.6	4.1
Prv. Transport	16.7	16.9	0.2
Pub. Transport	6.0	21.0	14.2
Communication	20.1	20.1	0.0
Recreation	12.9	16.3	3.1
Education	0.0	0.0	0.0
Restaurants	11.9	16.3	3.9
Others	5.0	5.0	0.0
Durables	21.0	21.0	0.0

Notes: Data weighted by population weights.
Source: Own calculations using EUROMOD 2012 policies for Belgium based on EU-SILC data 2010 and imputations based on HBS data (2009).

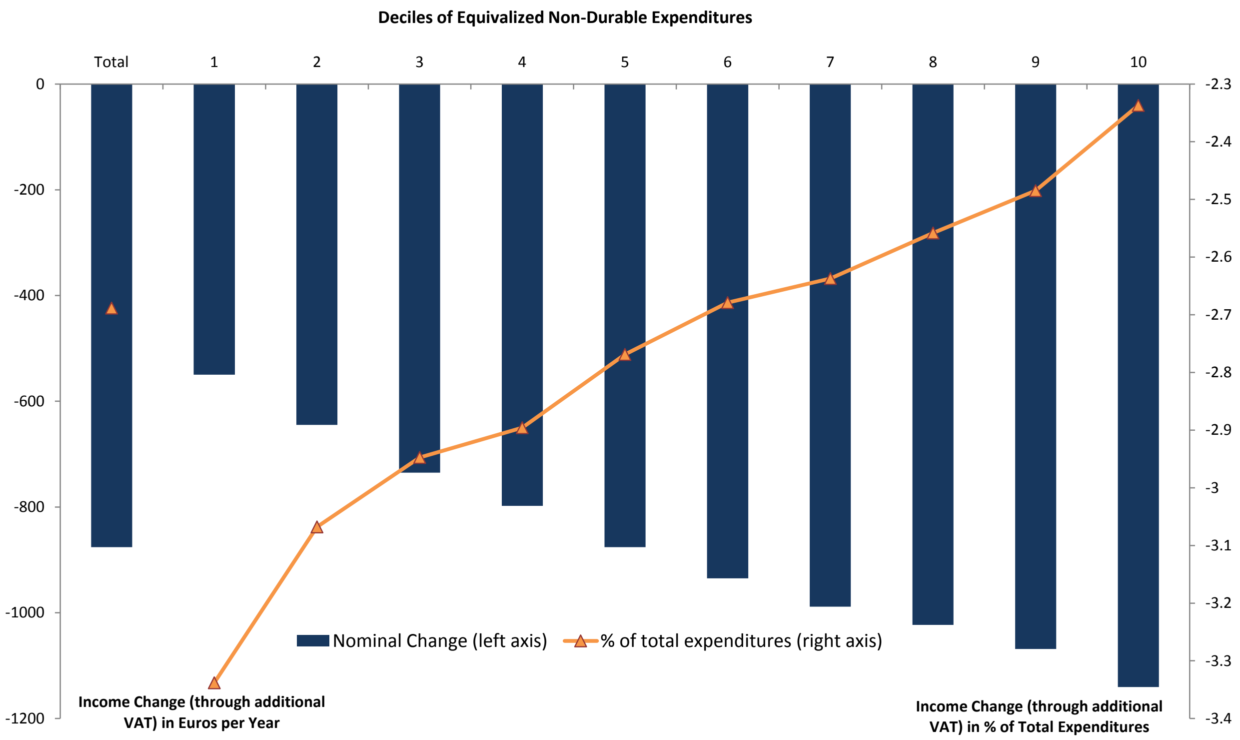
Implementing uniform VAT rates has huge effects on aggregate tax revenues. VAT revenues increase by 4.86 bn euros in the course of this reform if behavioral effects are assumed absent (i.e. assuming constant quantities). If we account for behavioral responses at the household spending structure (simulating responses implied by structures of estimated Engel curves)[14], VAT revenues only increase by 4.10 bn euros. This implies that households alter their spending structure (in the current period) such that they move away from goods that show implied price increases towards goods that became relatively cheaper, in the sense that their VAT rates were unaffected by uniformity.

We now come to the distributional effects of this reform, taking into account behavioral effects. From Figure 7 it is clear that the effects of moving VAT rates towards uniformity, without any form of compensation, are regressive. On average, households bear an extra

[14]Note that when behavioral effects are taken into account in the simulation, we assume savings to be the same as in the baseline, and we assume that quantities consumed of durable goods do not change, as modeling these would require a more complicated, intertemporal model. Also note that, while we use the Engel estimations to take into account changes in real disposable income when simulating the new budget shares, we did not incorporate the effects of price elasticities.

burden of about 880 euros per year in terms of additional VAT liabilities. This burden in absolute terms increases across the spending distribution from some 550 euros in the lowest spending decile to almost 1,140 euros in the 10th decile. However, in proportional terms, it decreases across the spending distribution. On average, it makes up 2.7% of total spending. For households in the lowest decile it is 3.3 percent, while the richest households only bear an additional tax liability of 2.3 percent of their total expenditures. We conclude that a move towards uniform VAT rates, without compensation for redistributional effects, is regressive, in the sense that the proportionally burdens are higher for poor households than for rich ones.

Figure 7: Distributional Reform Effects (Uniformity of VAT Rates)

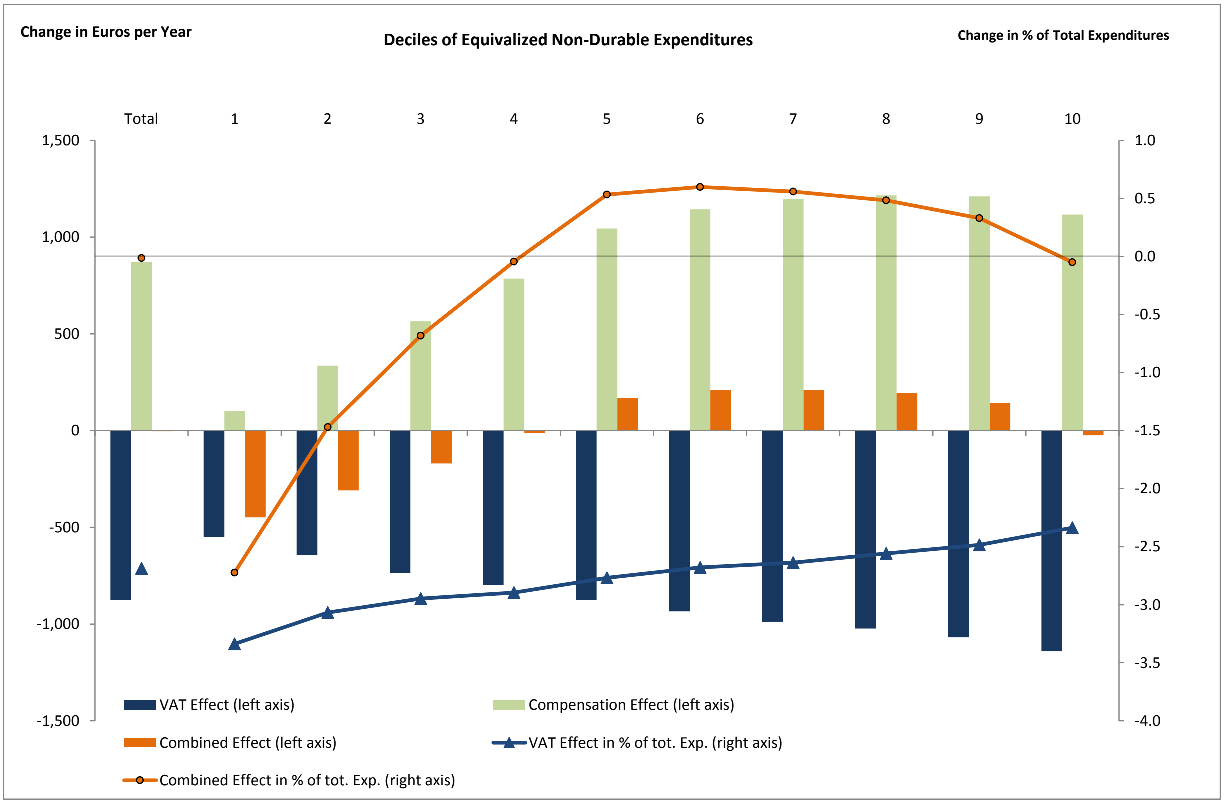


Notes: Weighted by population weights. Based on HBS data (2009) and Belgian EUROMOD output (2012 policy, based on EU-SILC 2010).

7.3. Uniformity with Compensation in Benefits and PIT (Reform B)

We now build on the results of the previous subsection, introducing an additional reform in the personal income tax system, making the total reform largely revenue-neutral, and on average compensating households for their increase in tax liabilities. We raise the tax-exempt allowance in the personal income tax system by about 41%, from 6,800 euros to 9,600 euros (Reform B).

Figure 8: Distributional Reform Effects (Uniformity, Compensated at Social Benefits and Taxes)



Notes: Weighted by population weights. Based on HBS data (2009) and Belgian EUROMOD output (2012 policy, based on EU-SILC 2010).

The aggregate cost of raising the tax-exempt allowance amounts to about 4.08 bn euros per year. This implies that Reform B is almost exactly revenue-neutral. From the distributional plot in figure 8 we can see though that households in the 5th-9th spending deciles slightly gain from the reform, while households in the 1st-3rd deciles bear some additional burdens (see Figure 8). This is related to the fact that these lower deciles benefit less that on average from the compensation, and that their VAT liabilities are relatively higher.

8. Conclusion

We have imputed expenditure information at the household level from household budget survey data into EUROMOD, documented in this paper for the case of Belgium. The goodness of fit seems to be acceptable to build on this integrated data base a microsimulation model that combines devices of direct taxation and social benefits with indirect taxation, in the context of EUROMOD.

VAT simulations based on imputed spending into the SILC data for Belgium reveal the typical incidence results in the baseline scenario of current VAT legislation. VAT looks regressive when plotted against the income distribution – tax burdens decrease in income in relative terms – while VAT is progressive when plotted against the expenditure distribution – tax burdens increase in spending in relative terms.

Policy reforms that build on uniformity with respect to reduced VAT rates can be implemented such that revenue neutrality is largely guaranteed.

References

Adam, S., Phillips, D., Smith, S., Bettendorf, L., Boeters, S., Kox, H., Straathof, B., Stuut, K., Baldini, M., Ferrari, M., Giannini, S., Onofri, P., Tomasini, S., Vincenzi, L., Barbone, L., Bird, R., Caro, L. V., Valkonen, T., Määttänen, N., Breuer, C., Ebertz, A., Nam, C., Berger, J., Strohner, L., Myck, M., Bach, S., Beznoska, M., Ochmann, R., Badenes, N., Canto, O., Labeaga, J., Trannoy, A., Ruiz, N., Denis, V., Decoster, A. and Verwerft, D. (2011). A retrospective evaluation of elements of the eu vat system: Final report. *Study on behalf of the European Commission (ed.), TAXUD/2010/DE/328, FWC No.TAXUD/2010/CC/104, Institute for Fiscal Studies (Project Leader)*. 12

Decoster, A., Ochmann, R. and Spiritus, K. (2013). *Integrating Indirect Taxation into EUROMOD. Documentation and Results for Germany*. Tech. rep., EUROMOD at the Institute for Social and Economic Research. †

—, Rock, B. D., Swerdt, K. D., Flannery, D., Loughrey, J., O'Donoghue, C. and Verwerft, D. (2007). Aim-ap: Deliverable 3.3 matching tax legislation and commodity aggregation. *Accurate Income Measurement for the Assessment of Public Policies (AIM-AP)*. 5

— and Spiritus, K. (2014). *SInTax: Microsimulation of VAT and Excises in the context of EUROMOD. User manual*. Flemosi Technical Note 16. 1, 13

Hufkens, T., Vanhille, J. and Spiritus, K. (2013). *EUROMOD Country report Belgium (BE) 2009-2012*. Tech. rep., ISER. 2, 9

Sutherland, H. and Figari, F. (2013). Euromod: the european union tax-benefit microsimulation model. *International Journal of Microsimulation*, **6(1)**, 4–26. 1

Wooldridge, J. M. (2003). *Introductory Econometrics - A Modern Approach*. Thomson. 6

A. Appendix – Plots comparing distributions of Durable and Non-Durable Spending

Note: all graphs are based on HBS data (2009) and EUROMOD output (policy 2012, using EU-SILC 2010).

Figure A.1

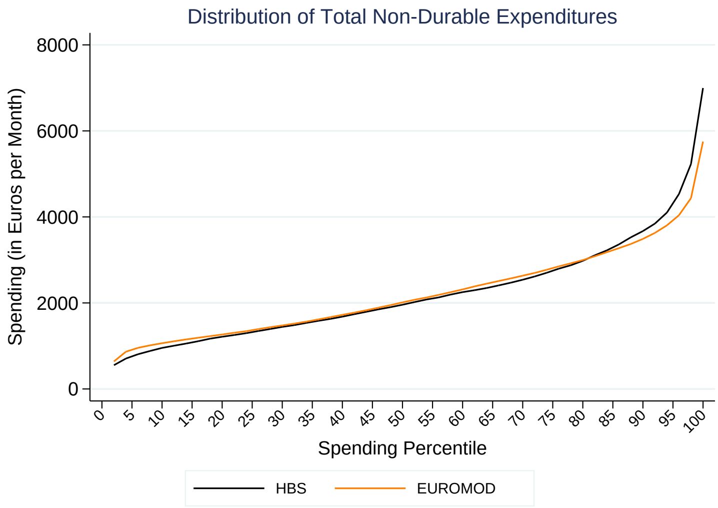
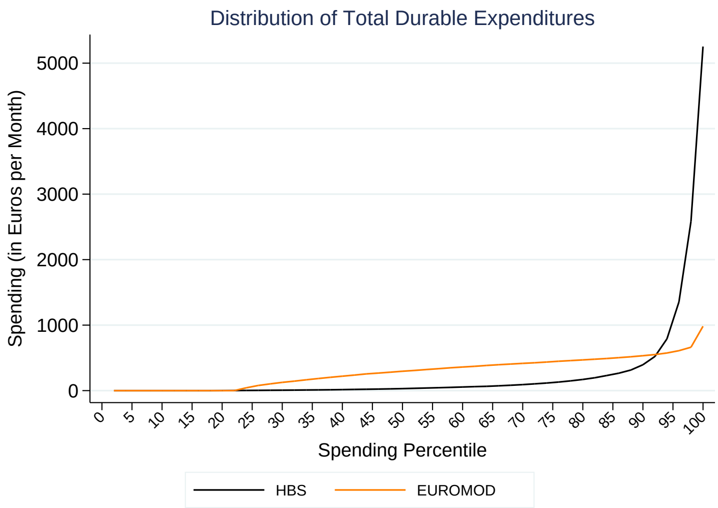


Figure A.2



B. Appendix – Plots comparing distributions of Non-Durable Expenditure Shares

Note: ll graphs are based on HBS data (2009) and EUROMOD output (policy 2012, using EU-SILC 2010).

Figure B.1

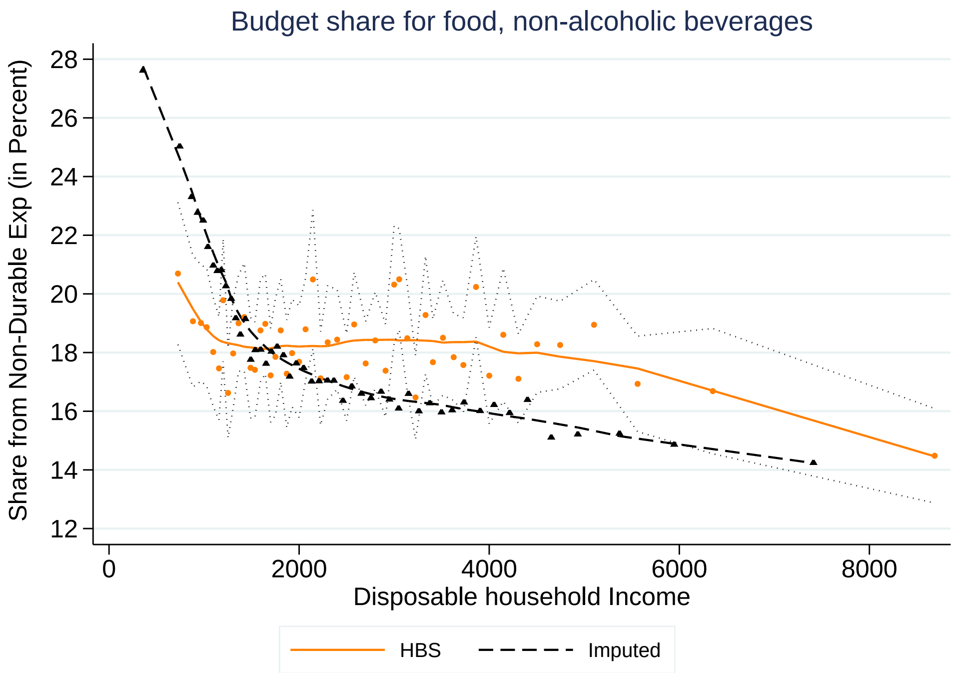


Figure B.2

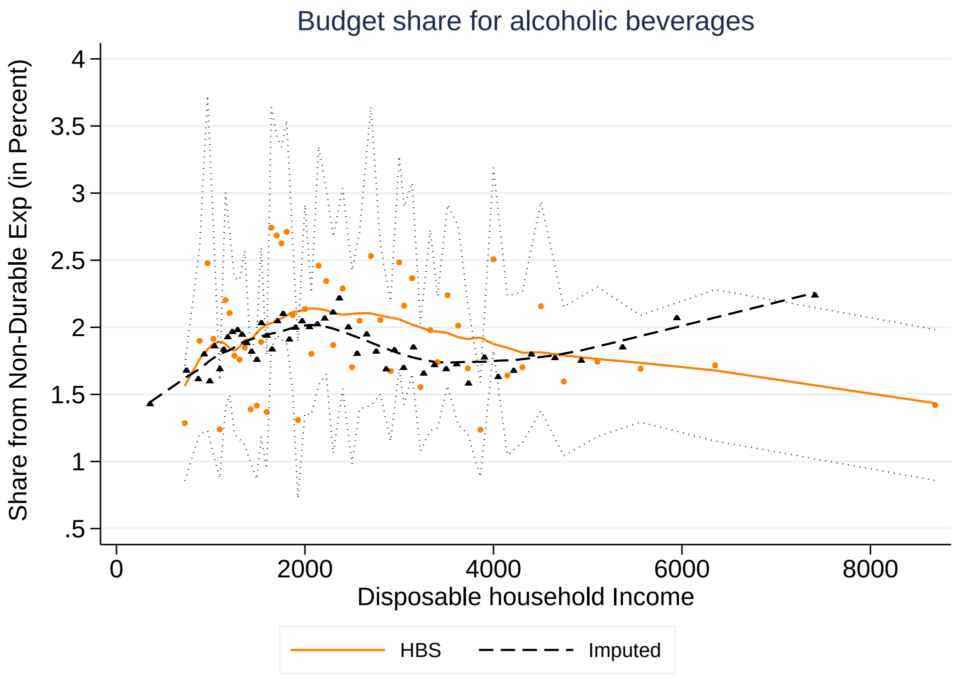


Figure B.3

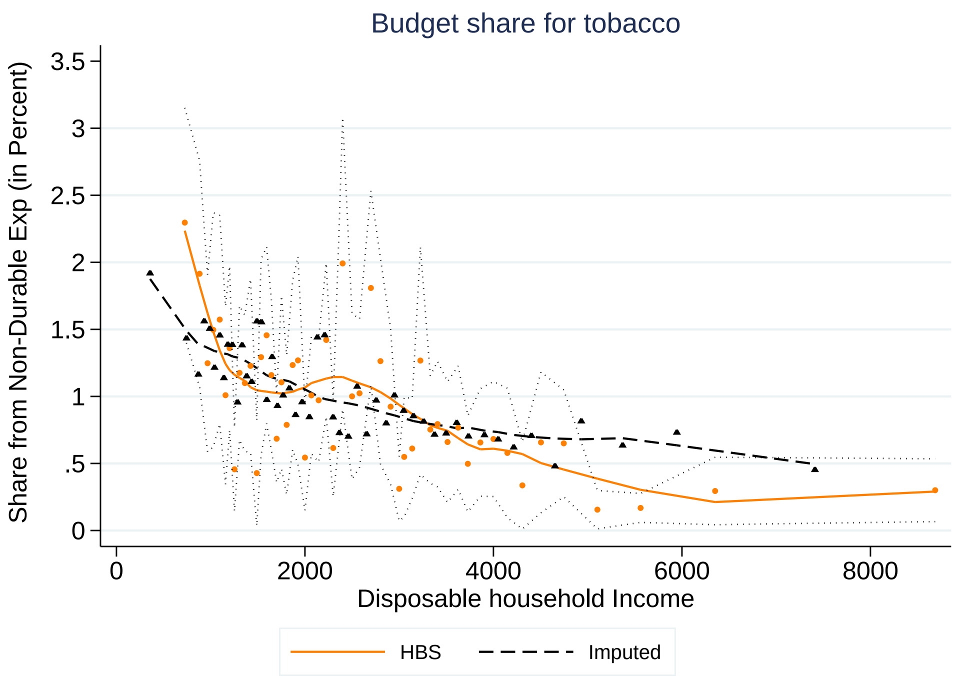


Figure B.4

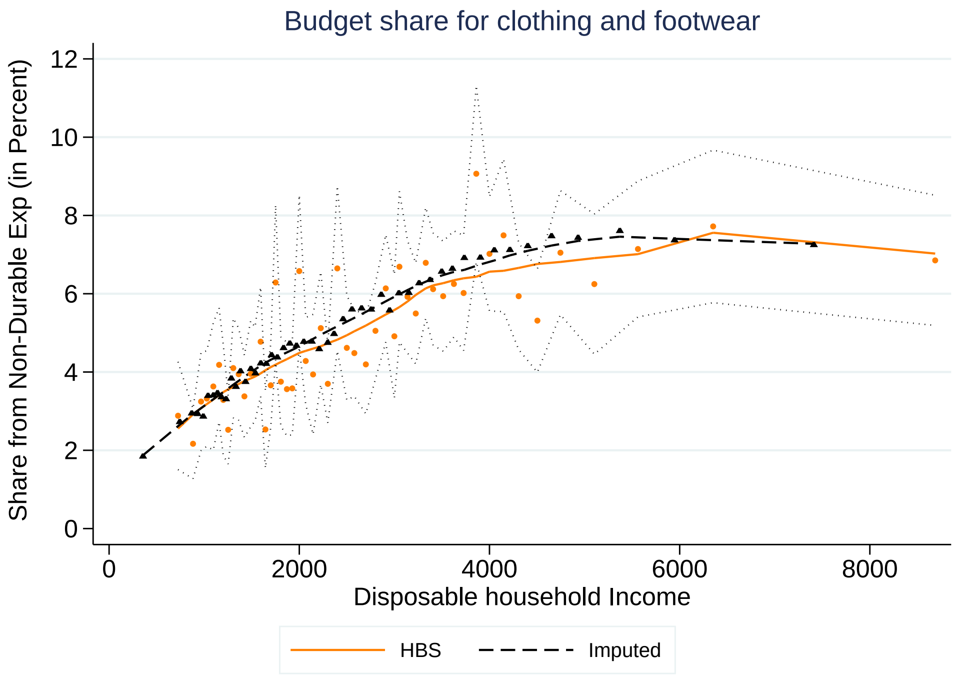


Figure B.5

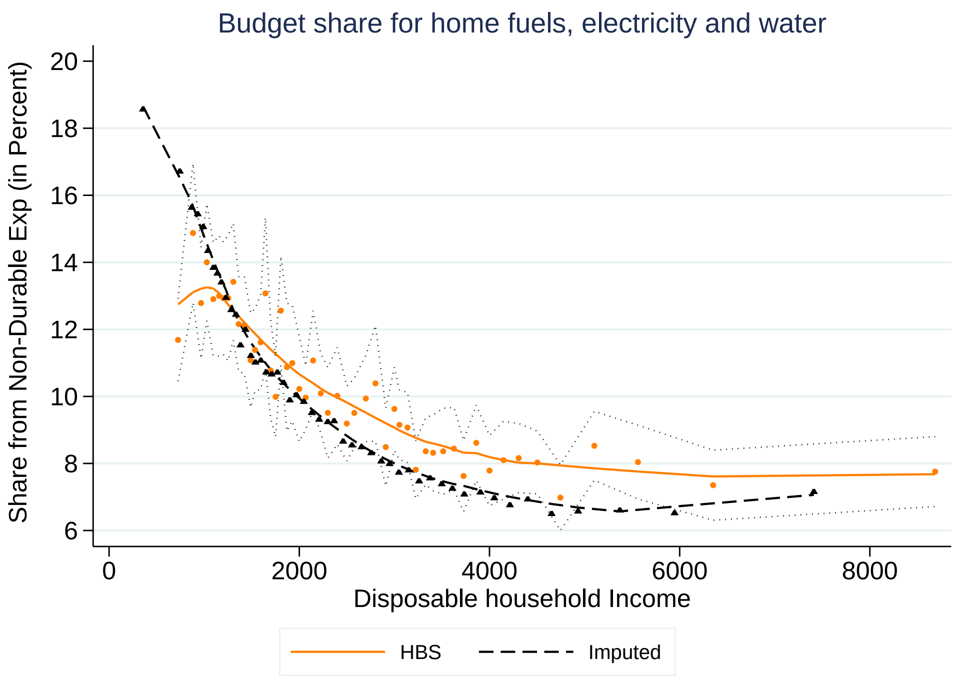


Figure B.6

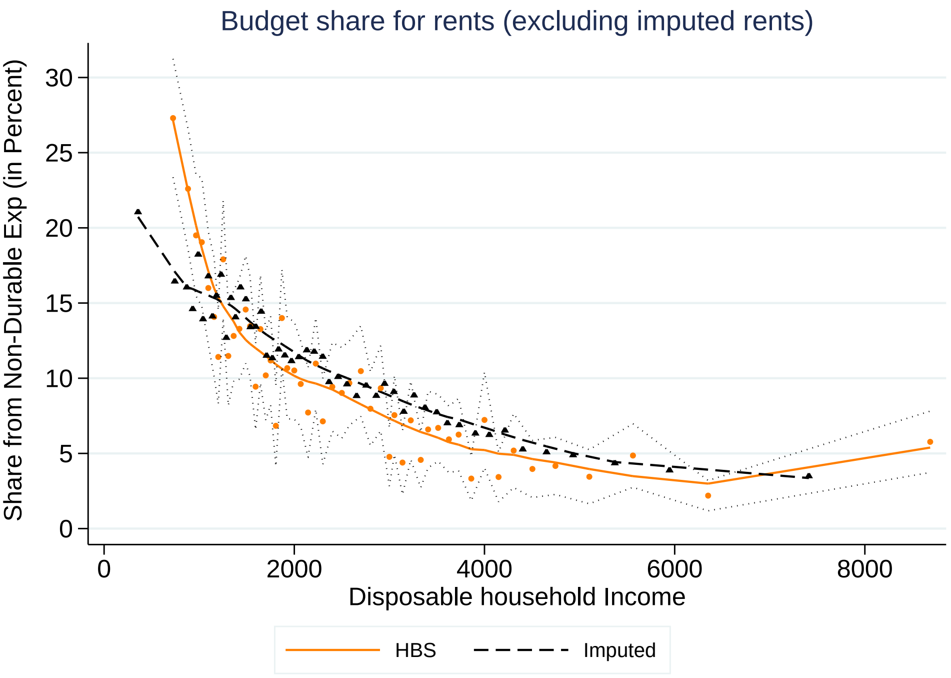


Figure B.7

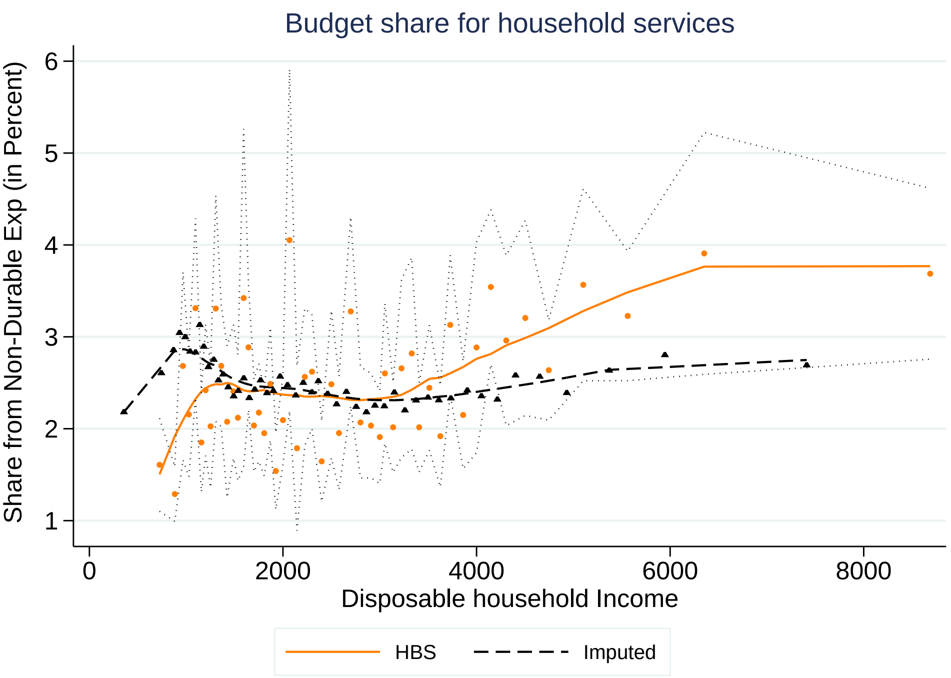


Figure B.8

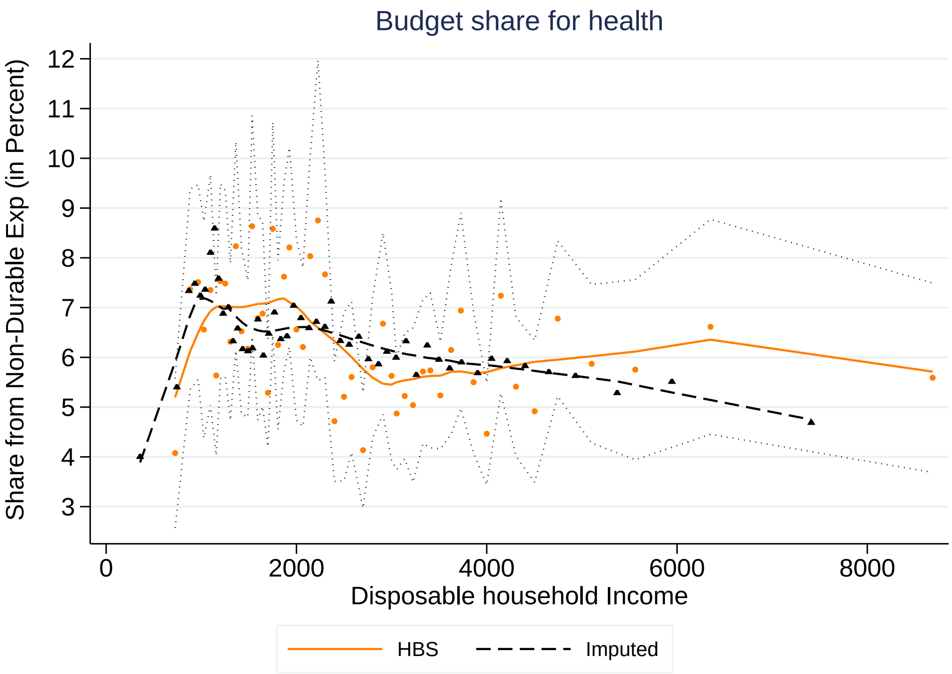


Figure B.9

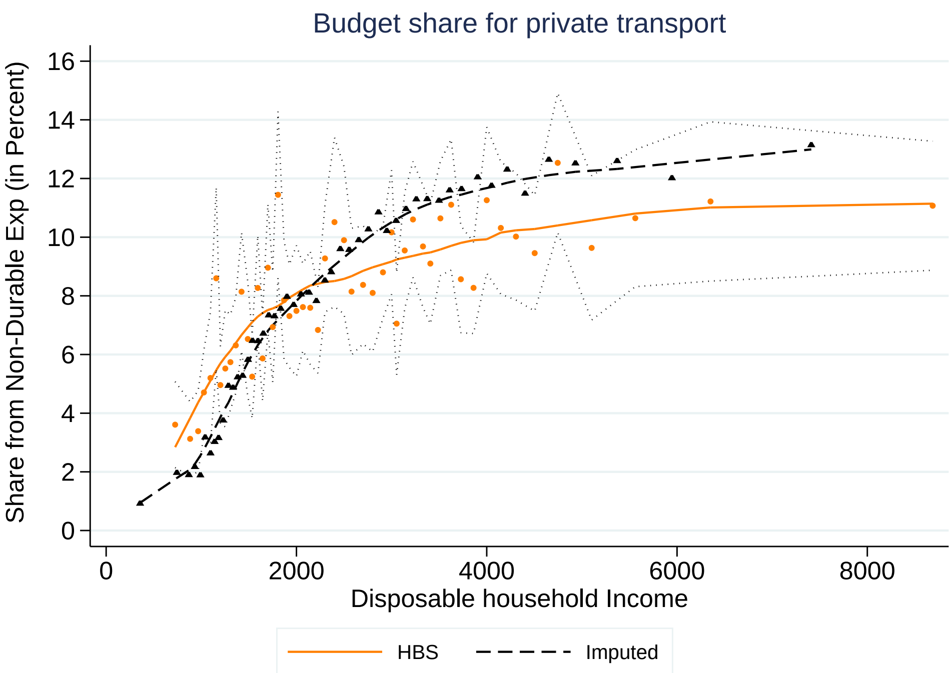


Figure B.10

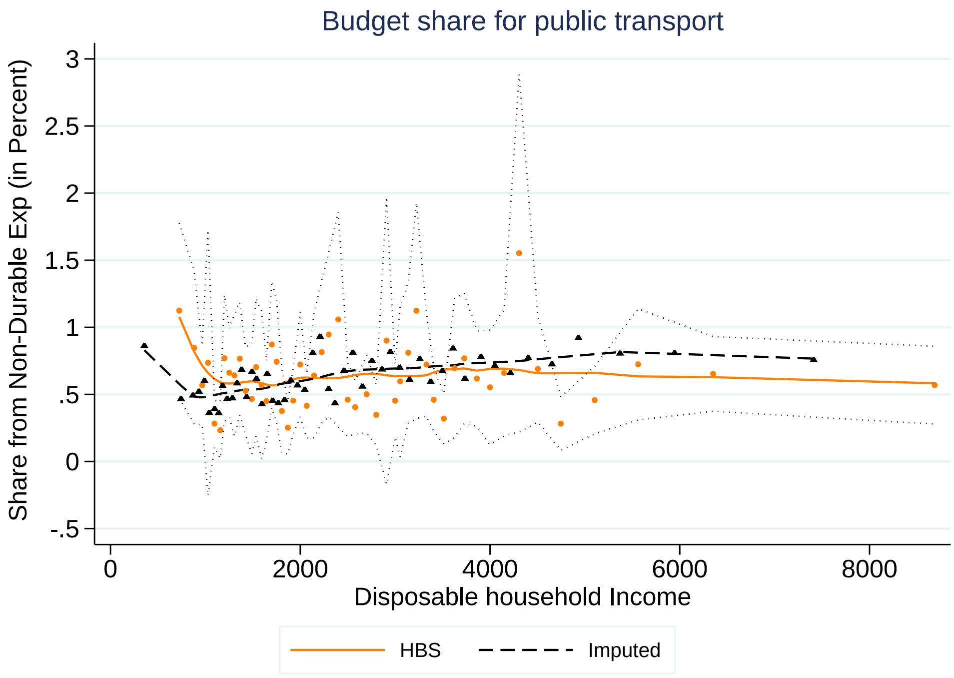


Figure B.11

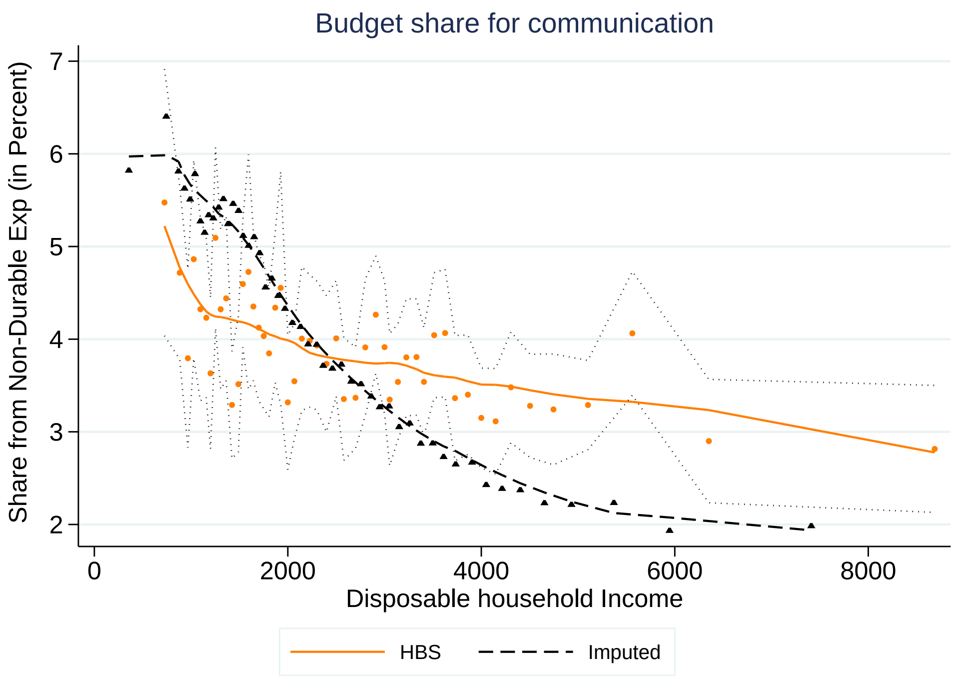


Figure B.12

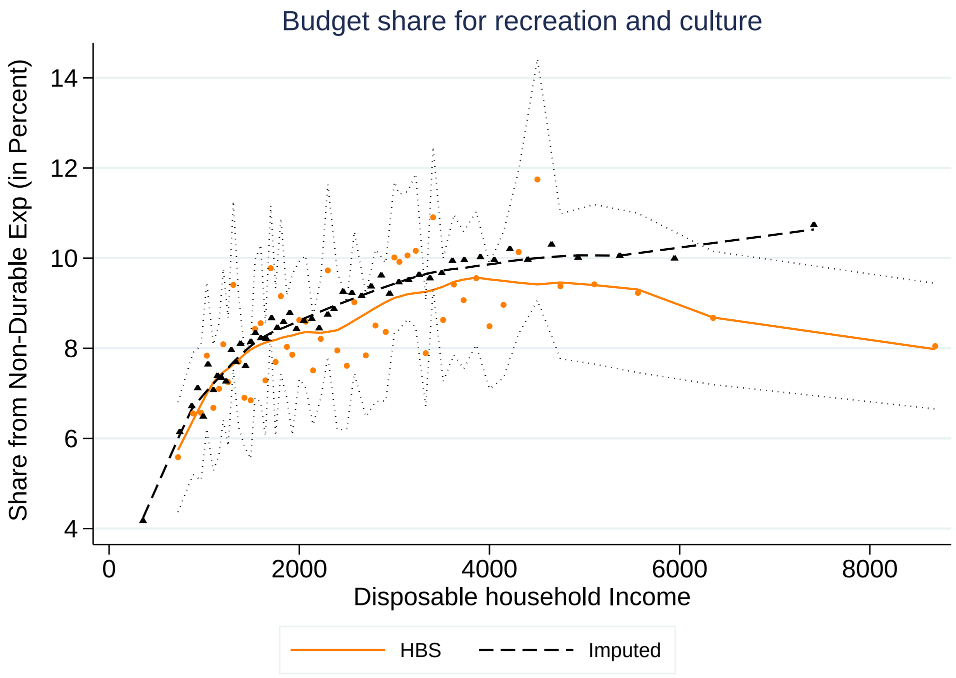


Figure B.13

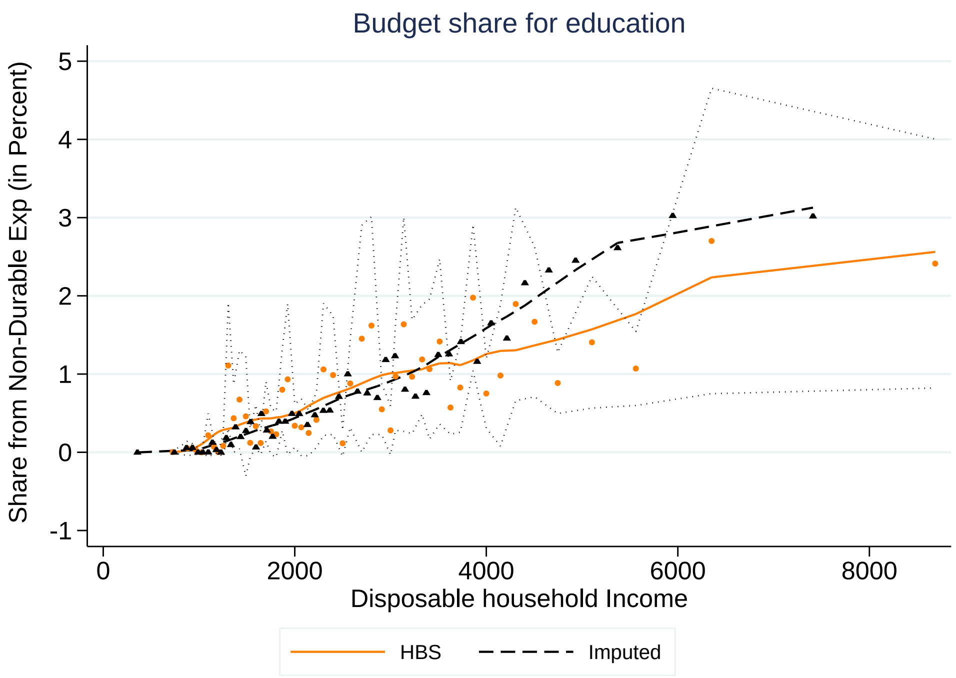


Figure B.14

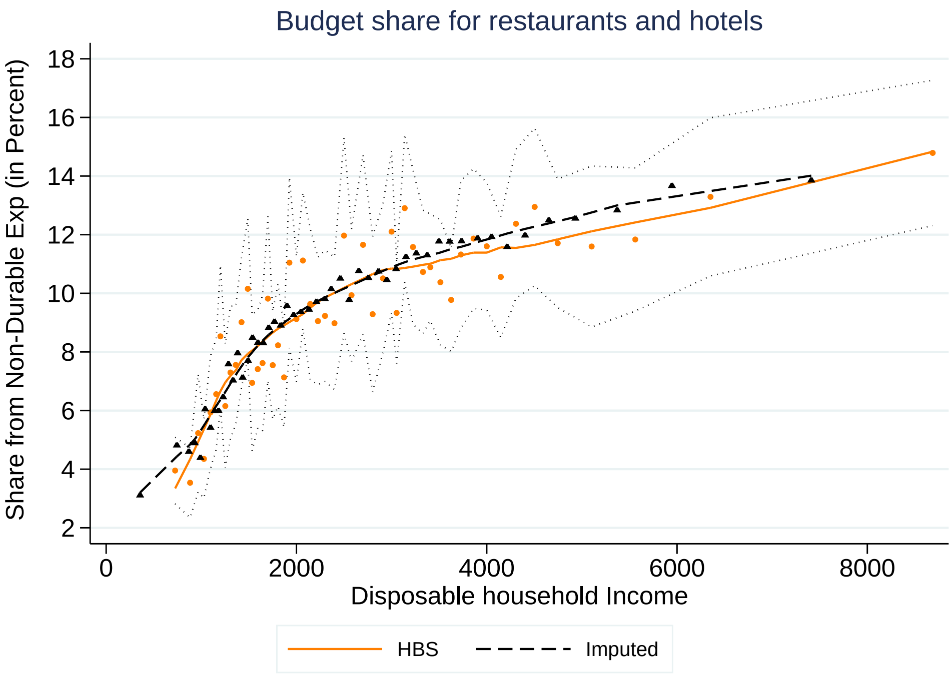
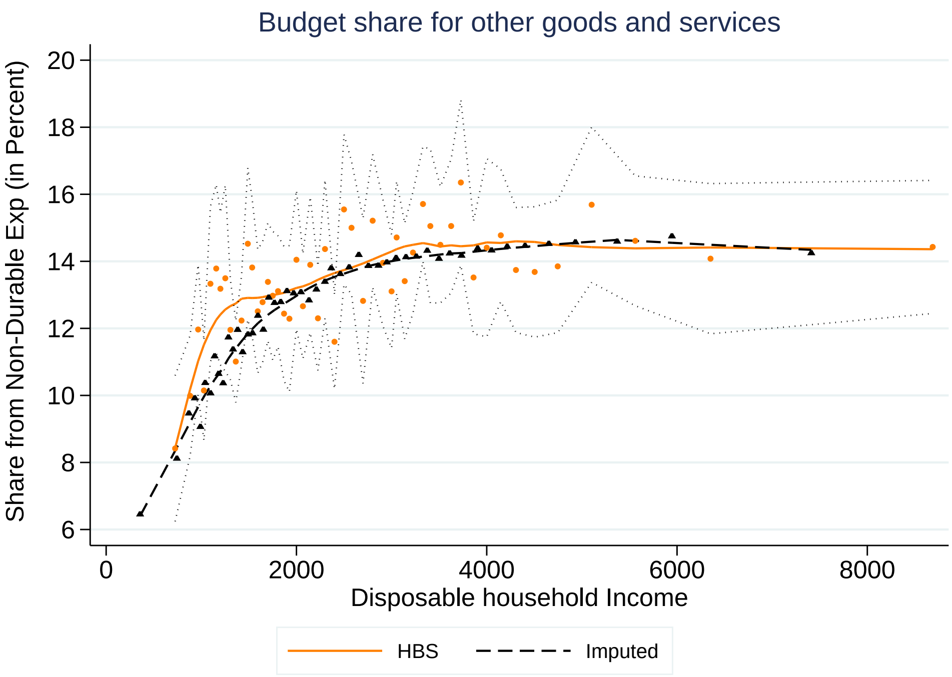


Figure B.15



C. Appendix – Estimation output

The following pages contain the regression output for all estimations. All results are based on Belgian HBS data for 2009. The variables used are described in table C.1.

Table C.1: Variables used in the regressions

Variable	Description
income	Net disposable income
flanders, wallonia	Region dummies
age	Age
male	Gender dummy
hh_persons, hh_children	Number of persons and number of children in the household
hh_actives	Number of persons in the household active on the labor market
secondary_educ, high_educ	Education dummies
self_employed, employed, unemployed, pensioner	Labor market status dummies

A prefix *log* in the estimation results indicates the log of a variable. A letter x indicates that an interaction term was constructed. A post-fix *2* or *3* indicates that powers were taken of the variables. Where a prefix *r_* is present, this means that residuals were used of a regression of higher order variables on all lower orders, in order to avoid multicollinearity of consecutive powers.

Table C.2: Estimation: Total non-durable expenditures

	(1) logtotexpnondur	
logincome	0.795***	(0.155)
r_logincome2	0.155	(0.142)
r_logincome3	-0.00218	(0.0437)
flanders	0.00223	(0.0500)
wallonia	-0.0518	(0.0500)
age	-0.00193	(0.00188)
r_age2	-0.0000180	(0.0000332)
r_age3	-0.00000400*	(0.00000185)
male	-0.0608	(0.0391)
hh_persons	0.127***	(0.0294)
hh_children	-0.129**	(0.0449)
hh_actives	0.141	(0.0753)
secondary_educ	0.0339	(0.0714)
high_educ	0.142*	(0.0714)
self_employed	0.0533	(0.188)
employed	0.0997	(0.129)
unemployed	0.229	(0.140)
pensioner	0.294*	(0.116)
agexincome	0.00000130	(0.00000108)
malexincome	0.0000306	(0.0000224)
hhactivesxincome	-0.0000709*	(0.0000307)
secondaryeducxincome	0.0000249	(0.0000556)
higheducxincome	0.00000986	(0.0000524)
flandersxincome	-0.0000232	(0.0000288)
walloniaxincome	0.00000171	(0.0000280)
self_employedxincome	-0.0000760	(0.0000941)
employedxincome	-0.0000387	(0.0000797)
unemployedxincome	-0.000150	(0.000137)
pensionerxincome	-0.0000850	(0.0000846)
hh_personsxincome	-0.0000196	(0.0000124)
hh_childrenxincome	0.0000464*	(0.0000183)
agexincome2	-8.34e-11	(1.34e-10)
malexincome2	-1.75e-09	(2.68e-09)
hhactivesxincome2	4.95e-09	(2.70e-09)
secondaryeducxincome2	-7.03e-09	(9.34e-09)
higheducxincome2	-6.67e-09	(8.69e-09)
flandersxincome2	8.35e-10	(3.29e-09)
walloniaxincome2	-1.41e-09	(2.98e-09)
self_employedxincome2	6.32e-09	(8.74e-09)
employedxincome2	1.37e-09	(7.77e-09)
unemployedxincome2	2.37e-08	(2.56e-08)
pensionerxincome2	-5.36e-09	(1.02e-08)
hh_personsxincome2	1.62e-09	(1.08e-09)
hh_childrenxincome2	-2.99e-09	(1.64e-09)
Constant	1.174	(1.158)
Observations	3595	
R^2	0.533	
Adjusted R^2	0.528	

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Income and expenditure variables in Euros per month. Data weighted by population weights. All amounts are downrated to 2009.

Table C.3: Estimation: Total non-durable expenditures (Probit)

	(1) positive_dur	
positive_dur		
logincome	1.839	(1.068)
r_logincome2	0.273	(0.995)
r_logincome3	-0.0984	(0.267)
flanders	0.493*	(0.246)
wallonia	0.123	(0.235)
age	0.0158	(0.00965)
r_age2	-0.000114	(0.000133)
r_age3	-0.0000206**	(0.00000717)
male	-0.587**	(0.190)
hh_persons	0.281	(0.165)
hh_children	-0.405	(0.241)
hh_actives	0.192	(0.459)
secondary_educ	0.434	(0.403)
high_educ	0.498	(0.421)
self_employed	-1.032	(1.019)
employed	0.194	(0.679)
unemployed	0.00842	(0.593)
pensioner	-0.141	(0.573)
agexincome	-0.0000114	(0.00000724)
malexincome	0.000257	(0.000139)
hhactivesxincome	-0.0000455	(0.000224)
secondaryeducxincome	-0.000225	(0.000400)
higheducxincome	-0.000251	(0.000402)
flandersxincome	-0.000279	(0.000178)
walloniaxincome	-0.0000194	(0.000163)
self_employedxincome	0.000373	(0.000631)
employedxincome	-0.0000115	(0.000536)
unemployedxincome	0.000169	(0.000645)
pensionerxincome	0.000540	(0.000559)
hh_personsxincome	-0.0000718	(0.0000934)
hh_childrenxincome	0.000247	(0.000130)
agexincome2	1.81e-09	(1.20e-09)
malexincome2	-2.94e-08	(2.18e-08)
hhactivesxincome2	2.55e-09	(2.46e-08)
secondaryeducxincome2	2.98e-10	(8.73e-08)
higheducxincome2	1.26e-08	(8.62e-08)
flandersxincome2	4.44e-08	(2.63e-08)
walloniaxincome2	1.08e-08	(2.24e-08)
self_employedxincome2	-6.77e-08	(0.000000116)
employedxincome2	-2.01e-08	(0.000000113)
unemployedxincome2	-7.14e-08	(0.000000149)
pensionerxincome2	-0.000000137	(0.000000123)
hh_personsxincome2	-8.96e-10	(1.17e-08)
hh_childrenxincome2	-2.30e-08	(1.59e-08)
Constant	-13.53	(7.894)
Observations	3595	

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Income and expenditure variables in Euros per month. Data weighted by population weights. All amounts are downrated to 2009.

Table C.4: Estimation: Total durable expenditures

	(1)	
	logtotexpdur	
logincome	0.499	(0.831)
r_logincome2	-0.112	(0.759)
r_logincome3	-0.166	(0.232)
flanders	-0.0360	(0.288)
wallonia	0.118	(0.290)
age	-0.0154	(0.0107)
r_age2	-0.000152	(0.000184)
r_age3	-0.00000413	(0.0000100)
male	0.208	(0.224)
hh_persons	0.157	(0.161)
hh_children	-0.305	(0.240)
hh_actives	0.104	(0.396)
secondary_educ	-0.706	(0.418)
high_educ	-0.422	(0.415)
self_employed	1.166	(1.082)
employed	0.910	(0.748)
unemployed	-0.293	(0.895)
pensioner	1.283	(0.709)
agexincome	0.00000338	(0.00000583)
malexincome	-0.0000356	(0.000121)
hhactivesxincome	-0.000126	(0.000158)
secondaryeducxincome	0.000449	(0.000305)
higheducxincome	0.000340	(0.000288)
flandersxincome	0.000161	(0.000156)
walloniaxincome	0.000124	(0.000152)
self_employedxincome	-0.000480	(0.000513)
employedxincome	-0.000320	(0.000435)
unemployedxincome	0.000667	(0.000812)
pensionerxincome	-0.000596	(0.000492)
hh_personsxincome	-0.0000338	(0.0000653)
hh_childrenxincome	0.000116	(0.0000955)
agexincome2	2.43e-10	(6.95e-10)
malexincome2	3.67e-09	(1.38e-08)
hhactivesxincome2	1.27e-08	(1.37e-08)
secondaryeducxincome2	-6.90e-08	(4.89e-08)
higheducxincome2	-6.26e-08	(4.57e-08)
flandersxincome2	-1.60e-08	(1.71e-08)
walloniaxincome2	-7.97e-09	(1.55e-08)
self_employedxincome2	4.65e-08	(4.59e-08)
employedxincome2	3.93e-08	(4.08e-08)
unemployedxincome2	-0.000000139	(0.000000141)
pensionerxincome2	8.13e-08	(6.39e-08)
hh_personsxincome2	7.02e-10	(5.53e-09)
hh_childrenxincome2	-4.78e-09	(8.35e-09)
Constant	-0.195	(6.263)
Observations	2850	
R^2	0.088	
Adjusted R^2	0.073	

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Income and expenditure variables in Euros per month. Data weighted by population weights. All amounts are downrated to 2009.

Table C.5: Estimation: Food, non alcoholic beverages

	(1) share_agg_1	
logremainingexp	-16.60***	(2.115)
r_logremainingexp2	-0.610	(0.897)
flanders	-2.057*	(0.969)
wallonia	-2.548*	(1.031)
age	0.0356	(0.0381)
r_age2	0.000685	(0.000656)
r_age3	-0.0000800*	(0.0000368)
male	1.176	(0.775)
hh_persons	5.726***	(0.583)
hh_children	-3.648***	(0.896)
hh_actives	3.769**	(1.276)
secondary_educ	-1.098	(1.372)
high_educ	-0.772	(1.465)
self_employed	-10.77***	(3.225)
employed	-11.77***	(2.376)
unemployed	-4.256*	(2.116)
pensioner	-5.909**	(2.246)
agexremainexp	0.0000576	(0.0000322)
malexremainexp	-0.000527	(0.000661)
secondaryeducxremainexp	0.000408	(0.00156)
higheducxremainexp	0.000104	(0.00159)
flandersxremainexp	0.00104	(0.000795)
walloniaxremainexp	0.00177*	(0.000873)
hhactivesxremainexp	-0.00147	(0.000883)
self_employedxremainexp	0.00575*	(0.00272)
employedxremainexp	0.00666**	(0.00235)
unemployedxremainexp	-0.0000617	(0.00245)
pensionerxremainexp	0.00375	(0.00244)
hh_personsxremainexp	-0.00113**	(0.000387)
hh_childrenxremainexp	0.00132*	(0.000574)
hh_personsxremainexp2	8.95e-08	(5.42e-08)
hh_childrenxremainexp2	-0.000000109	(8.16e-08)
agexremainexp2	-8.66e-09	(5.95e-09)
malexremainexp2	9.66e-08	(0.000000120)
hhactivesxremainexp2	0.000000101	(0.000000137)
secondaryeducxremainexp2	0.000000211	(0.000000380)
higheducxremainexp2	0.000000225	(0.000000378)
flandersxremainexp2	-0.000000157	(0.000000130)
walloniaxremainexp2	-0.000000225	(0.000000149)
self_employedxremainexp2	-0.000000364	(0.000000509)
employedxremainexp2	-0.000000496	(0.000000474)
unemployedxremainexp2	0.000000412	(0.000000493)
pensionerxremainexp2	-0.000000220	(0.000000499)
Constant	129.6***	(15.03)
Observations	3595	
R^2	0.375	
Adjusted R^2	0.368	

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Income and expenditure variables in Euros per month. Data weighted by population weights. All amounts are downrated to 2009.

Table C.6: Estimation: Alcoholic beverages

	(1)	
	share_agg_2	
logremainingexp	0.664	(0.977)
r_logremainingexp2	0.311	(0.414)
flanders	0.857	(0.448)
wallonia	1.006*	(0.476)
age	-0.00476	(0.0176)
r_age2	-0.000782**	(0.000303)
r_age3	-0.0000369*	(0.0000170)
male	1.544***	(0.358)
hh_persons	-0.387	(0.269)
hh_children	-0.253	(0.414)
hh_actives	0.811	(0.590)
secondary_educ	0.612	(0.634)
high_educ	0.926	(0.677)
self_employed	-1.081	(1.490)
employed	-0.0874	(1.098)
unemployed	1.134	(0.978)
pensioner	1.816	(1.038)
agexremainexp	0.0000290	(0.0000149)
malexremainexp	-0.000587	(0.000306)
secondaryeducxremainexp	-0.000221	(0.000720)
higheducxremainexp	-0.000281	(0.000733)
flandersxremainexp	-0.000699	(0.000367)
walloniaxremainexp	-0.000356	(0.000403)
hhactivesxremainexp	-0.000600	(0.000408)
self_employedxremainexp	0.000273	(0.00126)
employedxremainexp	-0.000185	(0.00109)
unemployedxremainexp	-0.00170	(0.00113)
pensionerxremainexp	-0.00145	(0.00113)
hh_personsxremainexp	0.000103	(0.000179)
hh_childrenxremainexp	0.000158	(0.000265)
hh_personsxremainexp2	-1.05e-08	(2.51e-08)
hh_childrenxremainexp2	-1.91e-08	(3.77e-08)
agexremainexp2	-4.58e-09	(2.75e-09)
malexremainexp2	6.06e-08	(5.54e-08)
hhactivesxremainexp2	9.84e-08	(6.31e-08)
secondaryeducxremainexp2	-1.10e-09	(0.000000175)
higheducxremainexp2	9.87e-09	(0.000000174)
flandersxremainexp2	8.37e-08	(6.02e-08)
walloniaxremainexp2	2.24e-08	(6.89e-08)
self_employedxremainexp2	-4.05e-08	(0.000000235)
employedxremainexp2	2.79e-08	(0.000000219)
unemployedxremainexp2	0.000000240	(0.000000228)
pensionerxremainexp2	0.000000229	(0.000000231)
Constant	-4.263	(6.944)
Observations	3595	
R^2	0.068	
Adjusted R^2	0.057	

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Income and expenditure variables in Euros per month. Data weighted by population weights. All amounts are downrated to 2009.

Table C.7: Estimation: Tobacco (Probit)

	(1) positive_agg_3	
positive_agg_3		
logtotexpnondur	0.121*	(0.0578)
r_logtotexpnondur2	-0.0379	(0.0640)
flanders	-0.137*	(0.0652)
wallonia	0.0637	(0.0674)
age	-0.0119***	(0.00284)
r_age2	-0.000428***	(0.000129)
r_age3	-0.00000231	(0.00000667)
male	0.0705	(0.0482)
hh_persons	-0.0322	(0.0332)
hh_children	-0.0877	(0.0475)
hh_actives	0.0126	(0.0574)
secondary_educ	-0.137	(0.0838)
high_educ	-0.420***	(0.0873)
self_employed	-0.676**	(0.208)
employed	-0.272	(0.170)
unemployed	-0.0296	(0.169)
pensioner	-0.277	(0.167)
Constant	-0.367	(0.436)
Observations	3595	

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Income and expenditure variables in Euros per month. Data weighted by population weights. All amounts are downrated to 2009.

Table C.8: Estimation: Tobacco

	(1) share_agg_3	
logtotexpnondur	-1.804***	(0.318)
r_logtotexpnondur2	0.624	(0.337)
flanders	0.232	(0.350)
wallonia	0.426	(0.353)
age	-0.00420	(0.0152)
r_age2	-0.00197**	(0.000683)
r_age3	0.00000248	(0.0000300)
male	-0.0295	(0.251)
hh_persons	-0.149	(0.184)
hh_children	-0.225	(0.258)
hh_actives	-0.234	(0.300)
secondary_educ	0.778	(0.417)
high_educ	0.210	(0.444)
self_employed	1.568	(1.102)
employed	1.191	(0.815)
unemployed	0.428	(0.785)
pensioner	1.591*	(0.806)
Constant	16.46***	(2.322)
Observations	926	
R^2	0.115	
Adjusted R^2	0.098	

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Income and expenditure variables in Euros per month. Data weighted by population weights. All amounts are downrated to 2009.

Table C.9: Estimation: Clothing and footwear

	(1) share_agg_4	
logremainingexp	0.226	(2.032)
r_logremainingexp2	-0.566	(0.861)
flanders	-1.699	(0.931)
wallonia	-1.703	(0.991)
age	-0.0811*	(0.0366)
r_age2	0.000296	(0.000630)
r_age3	0.00000926	(0.0000353)
male	-2.122**	(0.744)
hh_persons	0.183	(0.560)
hh_children	0.178	(0.861)
hh_actives	-1.876	(1.226)
secondary_educ	-1.093	(1.318)
high_educ	-1.608	(1.407)
self_employed	0.410	(3.098)
employed	1.446	(2.283)
unemployed	-0.0499	(2.033)
pensioner	0.417	(2.158)
agexremainexp	0.0000193	(0.0000309)
malexremainexp	0.000128	(0.000635)
secondaryeducxremainexp	0.000107	(0.00150)
higheducxremainexp	0.000375	(0.00152)
flandersxremainexp	0.00176*	(0.000764)
walloniaxremainexp	0.00132	(0.000838)
hhactivesxremainexp	0.000869	(0.000848)
self_employedxremainexp	0.000874	(0.00261)
employedxremainexp	-0.000150	(0.00226)
unemployedxremainexp	-0.000454	(0.00236)
pensionerxremainexp	0.000111	(0.00234)
hh_personsxremainexp	-0.000374	(0.000371)
hh_childrenxremainexp	0.000542	(0.000551)
hh_personsxremainexp2	6.31e-08	(5.21e-08)
hh_childrenxremainexp2	-0.000000104	(7.84e-08)
agexremainexp2	-2.90e-09	(5.71e-09)
malexremainexp2	2.95e-08	(0.000000115)
hhactivesxremainexp2	-2.54e-08	(0.000000131)
secondaryeducxremainexp2	6.73e-08	(0.000000365)
higheducxremainexp2	6.24e-08	(0.000000363)
flandersxremainexp2	-0.000000297*	(0.000000125)
walloniaxremainexp2	-0.000000266	(0.000000143)
self_employedxremainexp2	-0.000000214	(0.000000489)
employedxremainexp2	-9.84e-08	(0.000000455)
unemployedxremainexp2	0.000000162	(0.000000474)
pensionerxremainexp2	-0.000000148	(0.000000479)
Constant	8.329	(14.44)
Observations	3595	
R^2	0.096	
Adjusted R^2	0.085	

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Income and expenditure variables in Euros per month. Data weighted by population weights. All amounts are downrated to 2009.

Table C.10: Estimation: Home fuels and electricity

	(1) share_agg_5	
logremainingexp	-7.592***	(1.841)
r_logremainingexp2	1.041	(0.781)
flanders	1.116	(0.844)
wallonia	4.241***	(0.898)
age	0.0967**	(0.0332)
r_age2	0.000128	(0.000571)
r_age3	-0.0000525	(0.0000320)
male	-0.705	(0.674)
hh_persons	2.312***	(0.507)
hh_children	-0.311	(0.780)
hh_actives	-1.305	(1.111)
secondary_educ	-0.383	(1.195)
high_educ	-1.957	(1.275)
self_employed	-1.564	(2.808)
employed	-2.213	(2.069)
unemployed	-2.275	(1.842)
pensioner	-4.855*	(1.956)
agexremainexp	-0.0000219	(0.0000280)
malexremainexp	0.000586	(0.000576)
secondaryeducxremainexp	0.000213	(0.00136)
higheducxremainexp	0.000966	(0.00138)
flandersxremainexp	0.0000234	(0.000692)
walloniaxremainexp	-0.000930	(0.000760)
hhactivesxremainexp	0.000459	(0.000769)
self_employedxremainexp	0.000863	(0.00237)
employedxremainexp	0.00125	(0.00205)
unemployedxremainexp	0.00118	(0.00214)
pensionerxremainexp	0.00257	(0.00212)
hh_personsxremainexp	-0.000601	(0.000337)
hh_childrenxremainexp	-0.000194	(0.000500)
hh_personsxremainexp2	3.38e-08	(4.72e-08)
hh_childrenxremainexp2	4.17e-08	(7.10e-08)
agexremainexp2	1.80e-09	(5.18e-09)
malexremainexp2	-0.000000120	(0.000000104)
hhactivesxremainexp2	-4.85e-09	(0.000000119)
secondaryeducxremainexp2	3.91e-08	(0.000000330)
higheducxremainexp2	1.66e-08	(0.000000329)
flandersxremainexp2	-7.50e-09	(0.000000113)
walloniaxremainexp2	3.05e-08	(0.000000130)
self_employedxremainexp2	-1.76e-08	(0.000000444)
employedxremainexp2	-0.000000104	(0.000000413)
unemployedxremainexp2	-0.000000109	(0.000000429)
pensionerxremainexp2	-0.000000235	(0.000000435)
Constant	61.51***	(13.09)
Observations	3595	
R^2	0.332	
Adjusted R^2	0.324	

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Income and expenditure variables in Euros per month. Data weighted by population weights. All amounts are downrated to 2009.

Table C.11: Estimation: Rents (Probit)

	(1)	
	positive_agg_6	
positive_agg_6		
logtotexpnondur	0.420***	(0.0549)
r_logtotexpnondur2	-0.163**	(0.0597)
flanders	-0.491***	(0.0667)
wallonia	-0.532***	(0.0695)
age	-0.0145***	(0.00270)
r_age2	0.000154	(0.000124)
r_age3	-0.0000174*	(0.00000716)
male	0.0280	(0.0465)
hh_persons	-0.0832**	(0.0315)
hh_children	-0.0588	(0.0446)
hh_actives	-0.124*	(0.0547)
secondary_educ	-0.308***	(0.0862)
high_educ	-0.449***	(0.0887)
self_employed	-0.591**	(0.212)
employed	-0.377*	(0.188)
unemployed	-0.0454	(0.194)
pensioner	-0.552**	(0.182)
Constant	-0.559	(0.423)
Observations	3595	

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Income and expenditure variables in Euros per month. Data weighted by population weights. All amounts are downrated to 2009.

Table C.12: Estimation: Rents

	(1) share_agg_6	
logtotexpnondur	-6.257***	(0.680)
r_logtotexpnondur2	1.737*	(0.795)
flanders	-7.027***	(0.696)
wallonia	-7.397***	(0.740)
age	-0.0849**	(0.0325)
r_age2	0.00607***	(0.00151)
r_age3	0.0000506	(0.0000850)
male	-0.626	(0.550)
hh_persons	-1.609***	(0.378)
hh_children	0.209	(0.543)
hh_actives	-2.203**	(0.681)
secondary_educ	-2.508**	(0.959)
high_educ	-4.663***	(0.996)
self_employed	-1.611	(2.271)
employed	-2.597	(1.884)
unemployed	0.0418	(1.858)
pensioner	-6.781***	(1.869)
Constant	86.46***	(5.035)
Observations	2321	
R^2	0.265	
Adjusted R^2	0.260	

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Income and expenditure variables in Euros per month. Data weighted by population weights. All amounts are downrated to 2009.

Table C.13: Estimation: Household services

	(1) share_agg_7	
logremainingexp	-1.571	(1.096)
r_logremainingexp2	-0.626	(0.465)
flanders	-0.141	(0.502)
wallonia	-0.446	(0.534)
age	0.0391*	(0.0197)
r_age2	0.000860*	(0.000340)
r_age3	0.00000841	(0.0000191)
male	-0.174	(0.401)
hh_persons	-0.107	(0.302)
hh_children	0.00330	(0.464)
hh_actives	0.685	(0.661)
secondary_educ	0.811	(0.711)
high_educ	0.00369	(0.759)
self_employed	-4.941**	(1.671)
employed	-2.737*	(1.231)
unemployed	-1.656	(1.096)
pensioner	-2.162	(1.164)
agexremainexp	0.00000263	(0.0000167)
malexremainexp	0.0000141	(0.000343)
secondaryeducxremainexp	-0.000838	(0.000807)
higheducxremainexp	-0.000230	(0.000822)
flandersxremainexp	0.000221	(0.000412)
walloniaxremainexp	0.000191	(0.000452)
hhactivesxremainexp	-0.000493	(0.000458)
self_employedxremainexp	0.00402**	(0.00141)
employedxremainexp	0.00246*	(0.00122)
unemployedxremainexp	0.00113	(0.00127)
pensionerxremainexp	0.00185	(0.00126)
hh_personsxremainexp	-0.00000156	(0.000200)
hh_childrenxremainexp	0.000204	(0.000297)
hh_personsxremainexp2	-5.37e-09	(2.81e-08)
hh_childrenxremainexp2	-1.35e-08	(4.23e-08)
agexremainexp2	1.88e-09	(3.08e-09)
malexremainexp2	1.63e-08	(6.21e-08)
hhactivesxremainexp2	8.70e-08	(7.08e-08)
secondaryeducxremainexp2	0.000000204	(0.000000197)
higheducxremainexp2	0.000000150	(0.000000196)
flandersxremainexp2	-3.64e-08	(6.75e-08)
walloniaxremainexp2	2.23e-09	(7.73e-08)
self_employedxremainexp2	-0.000000619*	(0.000000264)
employedxremainexp2	-0.000000472	(0.000000246)
unemployedxremainexp2	-0.000000275	(0.000000255)
pensionerxremainexp2	-0.000000417	(0.000000259)
Constant	11.80	(7.787)
Observations	3595	
R^2	0.042	
Adjusted R^2	0.031	

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Income and expenditure variables in Euros per month. Data weighted by population weights. All amounts are downrated to 2009.

Table C.14: Estimation: Health

	(1) share_agg_8	
logremainingexp	1.245	(2.416)
r_logremainingexp2	-0.771	(1.025)
flanders	-2.226*	(1.107)
wallonia	-1.321	(1.178)
age	0.0991*	(0.0435)
r_age2	0.00221**	(0.000749)
r_age3	0.0000339	(0.0000420)
male	-2.038*	(0.885)
hh_persons	-0.790	(0.666)
hh_children	-0.625	(1.024)
hh_actives	1.345	(1.458)
secondary_educ	-0.526	(1.568)
high_educ	-2.435	(1.673)
self_employed	-1.837	(3.685)
employed	-1.161	(2.715)
unemployed	0.567	(2.418)
pensioner	-0.114	(2.566)
agexremainexp	-0.0000200	(0.0000368)
malexremainexp	0.000166	(0.000756)
secondaryeducxremainexp	-0.00286	(0.00178)
higheducxremainexp	-0.00120	(0.00181)
flandersxremainexp	0.00124	(0.000909)
walloniaxremainexp	0.000720	(0.000997)
hhactivesxremainexp	-0.00109	(0.00101)
self_employedxremainexp	0.00182	(0.00311)
employedxremainexp	0.00246	(0.00269)
unemployedxremainexp	0.00203	(0.00280)
pensionerxremainexp	0.00226	(0.00279)
hh_personsxremainexp	0.000565	(0.000442)
hh_childrenxremainexp	0.000302	(0.000656)
hh_personsxremainexp2	-7.42e-08	(6.20e-08)
hh_childrenxremainexp2	-3.72e-08	(9.32e-08)
agexremainexp2	3.75e-09	(6.80e-09)
malexremainexp2	6.80e-08	(0.000000137)
hhactivesxremainexp2	6.91e-08	(0.000000156)
secondaryeducxremainexp2	0.000000590	(0.000000434)
higheducxremainexp2	0.000000234	(0.000000431)
flandersxremainexp2	-0.000000137	(0.000000149)
walloniaxremainexp2	-0.000000119	(0.000000170)
self_employedxremainexp2	-0.000000272	(0.000000582)
employedxremainexp2	-0.000000434	(0.000000542)
unemployedxremainexp2	-0.000000448	(0.000000563)
pensionerxremainexp2	-0.000000599	(0.000000570)
Constant	-2.611	(17.17)
Observations	3595	
R^2	0.070	
Adjusted R^2	0.059	

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Income and expenditure variables in Euros per month. Data weighted by population weights. All amounts are downrated to 2009.

Table C.15: Estimation: Private transport

	(1) share_agg_9	
logremainingexp	13.73***	(2.768)
r_logremainingexp2	2.843*	(1.173)
flanders	2.905*	(1.268)
wallonia	1.478	(1.349)
age	0.0391	(0.0499)
r_age2	-0.00101	(0.000858)
r_age3	0.0000729	(0.0000481)
male	2.303*	(1.013)
hh_persons	0.0460	(0.762)
hh_children	0.185	(1.172)
hh_actives	-3.742*	(1.670)
secondary_educ	-0.407	(1.795)
high_educ	-1.121	(1.917)
self_employed	9.007*	(4.221)
employed	3.786	(3.109)
unemployed	-1.907	(2.769)
pensioner	-2.773	(2.939)
agexremainexp	-0.000128**	(0.0000421)
malexremainexp	0.0000652	(0.000866)
secondaryeducxremainexp	0.000850	(0.00204)
higheducxremainexp	0.00131	(0.00208)
flandersxremainexp	-0.00109	(0.00104)
walloniaxremainexp	0.000206	(0.00114)
hhactivesxremainexp	0.00232*	(0.00116)
self_employedxremainexp	-0.00212	(0.00356)
employedxremainexp	0.000275	(0.00308)
unemployedxremainexp	0.00542	(0.00321)
pensionerxremainexp	0.00411	(0.00319)
hh_personsxremainexp	-0.00126*	(0.000506)
hh_childrenxremainexp	-0.000374	(0.000751)
hh_personsxremainexp2	0.000000207**	(7.10e-08)
hh_childrenxremainexp2	9.68e-08	(0.000000107)
agexremainexp2	1.35e-08	(7.78e-09)
malexremainexp2	-7.62e-08	(0.000000157)
hhactivesxremainexp2	-0.000000256	(0.000000179)
secondaryeducxremainexp2	-0.000000333	(0.000000497)
higheducxremainexp2	-0.000000308	(0.000000494)
flandersxremainexp2	0.000000112	(0.000000171)
walloniaxremainexp2	-9.04e-08	(0.000000195)
self_employedxremainexp2	-0.000000214	(0.000000667)
employedxremainexp2	-0.000000529	(0.000000620)
unemployedxremainexp2	-0.00000123	(0.000000645)
pensionerxremainexp2	-0.00000102	(0.000000653)
Constant	-85.25***	(19.67)
Observations	3595	
R^2	0.140	
Adjusted R^2	0.129	

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Income and expenditure variables in Euros per month. Data weighted by population weights. All amounts are downrated to 2009.

Table C.16: Estimation: Public Transport (Probit)

	(1) positive_agg_10	
positive_agg_10		
logtotexpnondur	0.329***	(0.0562)
r_logtotexpnondur2	0.101	(0.0631)
flanders	-0.627***	(0.0615)
wallonia	-0.556***	(0.0648)
age	-0.00688*	(0.00277)
r_age2	-0.000200	(0.000129)
r_age3	-0.00000859	(0.00000741)
male	-0.104*	(0.0471)
hh_persons	0.163***	(0.0318)
hh_children	-0.281***	(0.0456)
hh_actives	-0.0966	(0.0551)
secondary_educ	-0.00561	(0.0887)
high_educ	0.215*	(0.0902)
self_employed	-0.134	(0.200)
employed	0.0897	(0.171)
unemployed	0.143	(0.173)
pensioner	-0.104	(0.168)
Constant	-2.387***	(0.427)
Observations	3595	

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Income and expenditure variables in Euros per month. Data weighted by population weights. All amounts are downrated to 2009.

Table C.17: Estimation: Public Transport

	(1)	
	share_agg_10	
logtotexpnondur	-1.220***	(0.280)
r_logtotexpnondur2	0.188	(0.305)
flanders	-0.551*	(0.272)
wallonia	-0.0272	(0.293)
age	-0.0289*	(0.0140)
r_age2	-0.00122	(0.000670)
r_age3	0.0000530	(0.0000366)
male	0.0752	(0.227)
hh_persons	-0.303*	(0.135)
hh_children	-0.145	(0.200)
hh_actives	0.239	(0.253)
secondary_educ	0.364	(0.476)
high_educ	0.382	(0.476)
self_employed	-0.714	(0.979)
employed	0.224	(0.839)
unemployed	-0.408	(0.857)
pensioner	-0.182	(0.844)
Constant	13.38***	(2.149)
Observations	1151	
R^2	0.075	
Adjusted R^2	0.061	

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Income and expenditure variables in Euros per month. Data weighted by population weights. All amounts are downrated to 2009.

Table C.18: Estimation: Communication

	(1) share_agg_11	
logremainingexp	-5.496***	(1.207)
r_logremainingexp2	-2.152***	(0.512)
flanders	-1.914***	(0.553)
wallonia	-1.500*	(0.589)
age	-0.0803***	(0.0217)
r_age2	-0.000665	(0.000374)
r_age3	0.0000315	(0.0000210)
male	0.138	(0.442)
hh_persons	-0.0919	(0.332)
hh_children	0.787	(0.511)
hh_actives	-0.503	(0.728)
secondary_educ	-0.0125	(0.783)
high_educ	-1.060	(0.836)
self_employed	0.626	(1.841)
employed	0.400	(1.356)
unemployed	-1.061	(1.208)
pensioner	-0.462	(1.282)
agexremainexp	0.0000428*	(0.0000184)
malexremainexp	-0.000210	(0.000378)
secondaryeducxremainexp	0.000276	(0.000889)
higheducxremainexp	0.000953	(0.000905)
flandersxremainexp	0.000636	(0.000454)
walloniaxremainexp	0.000648	(0.000498)
hhactivesxremainexp	0.000453	(0.000504)
self_employedxremainexp	-0.00237	(0.00155)
employedxremainexp	-0.00231	(0.00134)
unemployedxremainexp	-0.000253	(0.00140)
pensionerxremainexp	-0.00193	(0.00139)
hh_personsxremainexp	0.000244	(0.000221)
hh_childrenxremainexp	-0.000661*	(0.000328)
hh_personsxremainexp2	-3.17e-08	(3.10e-08)
hh_childrenxremainexp2	8.70e-08	(4.66e-08)
agexremainexp2	-5.27e-09	(3.40e-09)
malexremainexp2	4.12e-08	(6.84e-08)
hhactivesxremainexp2	-6.91e-08	(7.80e-08)
secondaryeducxremainexp2	-5.26e-08	(0.000000217)
higheducxremainexp2	-0.000000143	(0.000000216)
flandersxremainexp2	-7.35e-08	(7.44e-08)
walloniaxremainexp2	-8.20e-08	(8.51e-08)
self_employedxremainexp2	0.000000555	(0.000000291)
employedxremainexp2	0.000000561*	(0.000000271)
unemployedxremainexp2	0.000000266	(0.000000281)
pensionerxremainexp2	0.000000525	(0.000000285)
Constant	47.93***	(8.577)
Observations	3595	
R^2	0.127	
Adjusted R^2	0.116	

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Income and expenditure variables in Euros per month. Data weighted by population weights. All amounts are downrated to 2009.

Table C.19: Estimation: Recreation and culture

	(1) share_agg_12	
logremainingexp	3.034	(2.261)
r_logremainingexp2	-0.874	(0.959)
flanders	-0.707	(1.036)
wallonia	0.515	(1.103)
age	-0.0366	(0.0407)
r_age2	-0.00213**	(0.000701)
r_age3	0.00000245	(0.0000393)
male	-1.165	(0.828)
hh_persons	-2.483***	(0.623)
hh_children	1.770	(0.958)
hh_actives	1.465	(1.365)
secondary_educ	-0.574	(1.467)
high_educ	0.812	(1.566)
self_employed	-4.330	(3.449)
employed	0.596	(2.541)
unemployed	4.328	(2.262)
pensioner	5.844*	(2.402)
agexremainexp	-0.0000352	(0.0000344)
malexremainexp	0.00105	(0.000707)
secondaryeducxremainexp	0.00140	(0.00167)
higheducxremainexp	0.000925	(0.00170)
flandersxremainexp	0.000157	(0.000850)
walloniaxremainexp	-0.00117	(0.000933)
hhactivesxremainexp	-0.00186*	(0.000944)
self_employedxremainexp	0.00202	(0.00291)
employedxremainexp	-0.0000495	(0.00252)
unemployedxremainexp	-0.00245	(0.00262)
pensionerxremainexp	-0.00447	(0.00261)
hh_personsxremainexp	0.00114**	(0.000413)
hh_childrenxremainexp	-0.000780	(0.000614)
hh_personsxremainexp2	-5.82e-08	(5.80e-08)
hh_childrenxremainexp2	-2.70e-08	(8.72e-08)
agexremainexp2	4.94e-09	(6.36e-09)
malexremainexp2	-0.000000176	(0.000000128)
hhactivesxremainexp2	0.000000305*	(0.000000146)
secondaryeducxremainexp2	-0.000000270	(0.000000406)
higheducxremainexp2	-0.000000344	(0.000000404)
flandersxremainexp2	-8.58e-08	(0.000000139)
walloniaxremainexp2	0.000000275	(0.000000159)
self_employedxremainexp2	-0.000000394	(0.000000545)
employedxremainexp2	-9.21e-08	(0.000000507)
unemployedxremainexp2	0.000000128	(0.000000527)
pensionerxremainexp2	0.000000579	(0.000000534)
Constant	-7.137	(16.07)
Observations	3595	
R^2	0.064	
Adjusted R^2	0.052	

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Income and expenditure variables in Euros per month. Data weighted by population weights. All amounts are downrated to 2009.

Table C.20: Estimation: Education (Probit)

	(1) positive_agg_13	
positive_agg_13		
logtotexpnondur	0.729***	(0.108)
r_logtotexpnondur2	-0.188	(0.116)
flanders	0.339***	(0.102)
wallonia	0.198	(0.107)
age	-0.0198**	(0.00689)
r_age2	-0.00138***	(0.000373)
r_age3	0.0000392*	(0.0000165)
male	-0.259***	(0.0717)
hh_persons	0.466***	(0.0400)
hh_children	0.169**	(0.0531)
hh_actives	-0.188**	(0.0706)
secondary_educ	0.0477	(0.157)
high_educ	0.141	(0.157)
self_employed	0.304	(0.360)
employed	0.498	(0.332)
unemployed	0.745*	(0.338)
pensioner	0.0207	(0.371)
Constant	-7.617***	(0.875)
Observations	3595	

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Income and expenditure variables in Euros per month. Data weighted by population weights. All amounts are downrated to 2009.

Table C.21: Estimation: Education

	(1) share_agg_13	
logtotexpnondur	0.0566	(1.034)
r_logtotexpnondur2	1.521	(0.874)
flanders	-2.362**	(0.805)
wallonia	-1.677*	(0.853)
age	-0.129	(0.0900)
r_age2	-0.0103*	(0.00491)
r_age3	-0.000283	(0.000207)
male	-0.788	(0.560)
hh_persons	1.257***	(0.309)
hh_children	-1.421***	(0.357)
hh_actives	-2.260***	(0.545)
secondary_educ	-0.813	(1.264)
high_educ	-0.427	(1.281)
self_employed	2.538	(3.115)
employed	3.511	(2.931)
unemployed	3.144	(2.898)
pensioner	1.839	(3.703)
Constant	8.873	(8.594)
Observations	551	
R^2	0.154	
Adjusted R^2	0.127	

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Income and expenditure variables in Euros per month. Data weighted by population weights. All amounts are downrated to 2009.

Table C.22: Estimation: Restaurants

	(1) share_agg_14	
logremainingexp	3.190	(3.008)
r_logremainingexp2	-0.564	(1.275)
flanders	1.900	(1.378)
wallonia	-1.145	(1.467)
age	-0.107*	(0.0542)
r_age2	0.000340	(0.000933)
r_age3	0.0000572	(0.0000523)
male	4.984***	(1.102)
hh_persons	-3.024***	(0.828)
hh_children	-0.392	(1.274)
hh_actives	-1.121	(1.815)
secondary_educ	-1.254	(1.951)
high_educ	1.986	(2.083)
self_employed	5.834	(4.587)
employed	2.867	(3.380)
unemployed	0.609	(3.010)
pensioner	0.330	(3.195)
agexremainexp	0.00000773	(0.0000458)
malexremainexp	-0.00277**	(0.000941)
secondaryeducxremainexp	0.00310	(0.00222)
higheducxremainexp	0.000936	(0.00226)
flandersxremainexp	-0.00161	(0.00113)
walloniaxremainexp	-0.00165	(0.00124)
hhactivesxremainexp	0.000882	(0.00126)
self_employedxremainexp	-0.00330	(0.00387)
employedxremainexp	-0.00214	(0.00335)
unemployedxremainexp	-0.00148	(0.00349)
pensionerxremainexp	0.00142	(0.00347)
hh_personsxremainexp	0.000675	(0.000550)
hh_childrenxremainexp	0.000723	(0.000816)
hh_personsxremainexp2	-0.000000115	(7.71e-08)
hh_childrenxremainexp2	-0.000000113	(0.000000116)
agexremainexp2	8.27e-09	(8.46e-09)
malexremainexp2	0.000000346*	(0.000000170)
hhactivesxremainexp2	-0.000000189	(0.000000194)
secondaryeducxremainexp2	-0.000000760	(0.000000540)
higheducxremainexp2	-0.000000526	(0.000000537)
flandersxremainexp2	0.000000426*	(0.000000185)
walloniaxremainexp2	0.000000366	(0.000000212)
self_employedxremainexp2	0.000000618	(0.000000725)
employedxremainexp2	0.000000646	(0.000000674)
unemployedxremainexp2	0.000000159	(0.000000701)
pensionerxremainexp2	6.27e-08	(0.000000710)
Constant	-8.448	(21.38)
Observations	3595	
R^2	0.170	
Adjusted R^2	0.160	

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Income and expenditure variables in Euros per month. Data weighted by population weights. All amounts are downrated to 2009.

Table C.23: Estimation: Other goods and services

	(1) share_agg_15	
logremainingexp	9.180***	(2.692)
r_logremainingexp2	1.967	(1.141)
flanders	1.966	(1.233)
wallonia	1.423	(1.312)
age	-0.000379	(0.0485)
r_age2	0.0000676	(0.000835)
r_age3	-0.0000463	(0.0000468)
male	-3.942***	(0.986)
hh_persons	-1.384	(0.741)
hh_children	2.306*	(1.140)
hh_actives	0.472	(1.624)
secondary_educ	3.925*	(1.746)
high_educ	5.225**	(1.864)
self_employed	8.648*	(4.105)
employed	8.875**	(3.024)
unemployed	4.568	(2.693)
pensioner	7.869**	(2.859)
agexremainexp	0.0000458	(0.0000410)
malexremainexp	0.00208*	(0.000842)
secondaryeducxremainexp	-0.00243	(0.00198)
higheducxremainexp	-0.00386	(0.00202)
flandersxremainexp	-0.00168	(0.00101)
walloniaxremainexp	-0.000750	(0.00111)
hhactivesxremainexp	0.000518	(0.00112)
self_employedxremainexp	-0.00781*	(0.00346)
employedxremainexp	-0.00826**	(0.00299)
unemployedxremainexp	-0.00336	(0.00312)
pensionerxremainexp	-0.00822**	(0.00310)
hh_personsxremainexp	0.000636	(0.000492)
hh_childrenxremainexp	-0.00124	(0.000730)
hh_personsxremainexp2	-9.85e-08	(6.90e-08)
hh_childrenxremainexp2	0.000000197	(0.000000104)
agexremainexp2	-1.27e-08	(7.57e-09)
malexremainexp2	-0.000000285	(0.000000152)
hhactivesxremainexp2	-0.000000117	(0.000000174)
secondaryeducxremainexp2	0.000000307	(0.000000483)
higheducxremainexp2	0.000000623	(0.000000481)
flandersxremainexp2	0.000000172	(0.000000166)
walloniaxremainexp2	8.54e-08	(0.000000190)
self_employedxremainexp2	0.000000962	(0.000000648)
employedxremainexp2	0.000000991	(0.000000603)
unemployedxremainexp2	0.000000690	(0.000000627)
pensionerxremainexp2	0.00000125*	(0.000000635)
Constant	-51.45**	(19.13)
Observations	3595	
R^2	0.043	
Adjusted R^2	0.031	

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Income and expenditure variables in Euros per month. Data weighted by population weights. All amounts are downrated to 2009.