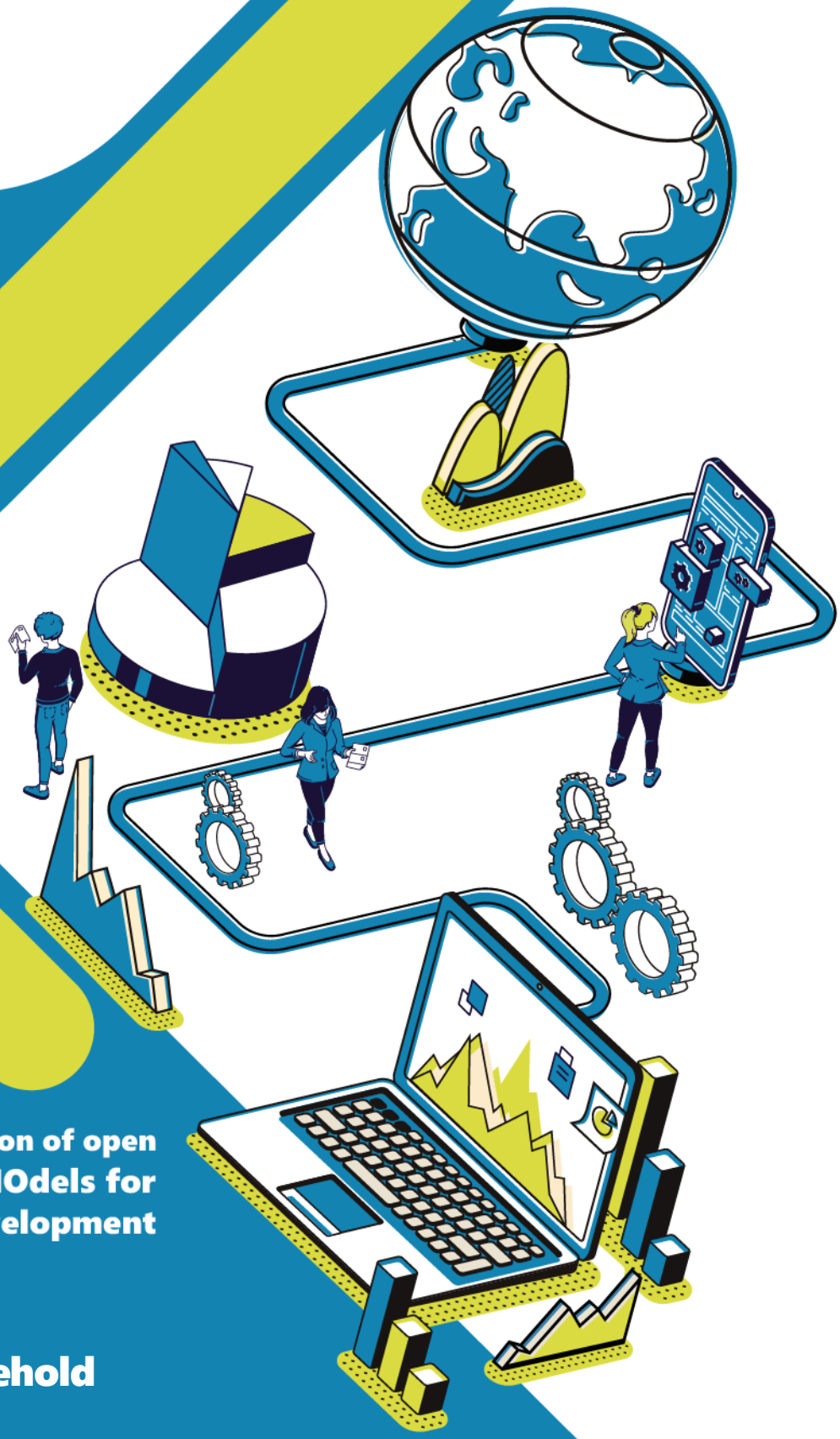
DIAMOND

www.climate-diamond.eu



**Delivering the next generation of open Integrated Assessment MOdels for Net-zero, sustainable Development**

# D4.3 – Enhanced household heterogeneity in IAMs

WP4 – Integrate & Expand. Beyond structural

30/09/2025

Funded by
the European Union

## Disclaimer

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Climate, Infrastructure and Environment Executive Agency (CINEA). Neither the European Union nor the granting authority can be held responsible for them.

## Copyright Message

This report, if not confidential, is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0); a copy is available here: https://creativecommons.org/licenses/by/4.0/. You are free to share (copy and redistribute the material in any medium or format) and adapt (remix, transform, and build upon the material for any purpose, even commercially) under the following terms: (i) attribution (you must give appropriate credit, provide a link to the license, and indicate if changes were made; you may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use); (ii) no additional restrictions (you may not apply legal terms or technological measures that legally restrict others from doing anything the license permits).

| Grant Agreement Number | 101081179 | | Acronym | DIAMOND |
|---|---|---|---|---|
| **Full Title** | Delivering the next generation of open Integrated Assessment MOdels for Net-zero, sustainable Development | | | |
| **Topic** | HORIZON-CL5-2022-D1-02 | | | |
| **Funding scheme** | HORIZON EUROPE, RIA – Research and Innovation Action | | | |
| **Start Date** | December 2022 | **Duration** | 48 Months | |
| **Project URL** | http://www.climate-diamond.eu/ | | | |
| **EU Project Advisor** | Silvia Vaghi | | | |
| **Project Coordinator** | Institute of Communications and Computer Systems - ICCS | | | |
| **Deliverable** | D4.3 – Enhanced household heterogeneity in IAMs | | | |
| **Work Package** | WP4 – Integrate & Expand. Beyond structural | | | |
| **Date of Delivery** | **Contractual** | 30/09/2025 | **Actual** | 30/09/2025 |
| **Nature** | Report | **Dissemination** | Public | |
| **Lead Beneficiary** | CESAR | | | |
| **Responsible Author** | Kurt Kratena | **Email** | kurt.kratena@cesarecon.at | |
| | CESAR | **Phone** | +43 699 12233989 | |
| **Contributors** | Manuel Tomás, Jon Sampedro, Dirk-Jan van de Ven (BC3); Marc Vielle, Sigit Perdana (EPFL); Baptiste Boitier, Pierre Le Mouël, Hayat Mekki (SEURECO) | | | |
| **Reviewer(s)** | Anastasios Karamaneas, Konstantinos Koasidis, Alexandros Nikas (ICCS); Shivika Mittal (CICERO); Sigit Perdana (EPFL); Clàudia Rodés-Bachs (BC3) | | | |
| **Keywords** | Household modelling, statistical matching, income distribution | | | |

## EC Summary Requirements

# 1. Changes with respect to the DoA

Given the detailed goods classification of consumption in NEMESIS-World and the lack thereof in OPEN-PROM, the household heterogeneity dataset will not be used in OPEN-PROM, but in NEMESIS-World instead. This switch ensures the intended diversity in representing household heterogeneity in the project's IAMs, as reflected in the DoA (Task 4.3). The OPEN-PROM team is instead investing these resources in establishing links with the physical/ climate impact models (in Task 4.4 and D4.4), which is in addition to the set of IAMs intended to incorporate probabilistic downscaled climate information.

# 2. Dissemination and uptake

The dataset can be used by different IAMs (as illustrated in this report through the links with GCAM-EU, GEMINI-E3-EU, and NEMESIS-World). Other tasks in the project, as well as researchers in similar fields (especially those working on behavioural change of households) can use the household heterogeneity data in their work/models.

# 3. Short summary of results

Household heterogeneity has gained importance in economic modelling, both in macroeconomics as well as in Integrated Assessment Models (IAMs). In DIAMOND, we constructed a detailed multi-household database for the EU-27. This has been analysed, checked, and built in into three IAMs applied in DIAMOND: GCAM-EU, GEMINI-E3-EU, NEMESIS-World. These models already had some degree of household heterogeneity and of consumption disaggregation into different purposes implemented, with the link to the database significantly enhancing this capacity. The main task in the construction of the database was a statistical matching between the Survey on Income and Living Conditions (EU SILC) and the House Budget Survey (HBS) of Eurostat. Results were analysed and described with respect to their plausibility. The database was then introduced in the three IAMs and partially compared to their existing household heterogeneity. This shows different degrees of extension in the three IAMs. In GCAM, consumer heterogeneity was previously represented in the residential building sector; integration of the new household database in GCAM-Europe combines information on country-level residential energy consumption (Eurostat) with disaggregated residential energy consumption by fuel and income decile from the new database. In GEMINI-E3-EU, household heterogeneity was introduced in its latest version; to integrate the new DIAMOND database into the model, it needed to be reconciled with the published Eurostat HBS data, and the income data already implemented in GEMINI-E3-EU. The current version of NEMESIS-World features a single representative household. a mapping between the project's household database, the distributional national accounts by Eurostat, and the model's household module was thus implemented, yielding a dataset of income and expenditure data for each model income or consumption category, disaggregated by household deciles, which was then integrated.

# 4. Evidence of accomplishment

This report.

Funded by
the European Union

# Preface

DIAMOND will update, upgrade, and fully open six Integrated Assessment Models (IAMs) that are emblematic in scientific and policy processes, improving their sectoral and technological detail, spatiotemporal resolution, and geographic granularity. It will further enhance modelling capacity to assess the feasibility and desirability of Paris-compliant mitigation pathways, their interplay with adaptation, circular economy, and other SDGs, their distributional and equity effects, and their resilience to extremes, as well as robust risk management and investment strategies. This will be done via integration of tools and insights from psychology, finance research, behavioural and labour economics, operational research, and physical science. The project will develop a transdisciplinary scientific approach to legitimise the implementation process and co-create research questions that stretch the frontiers of climate science, as well as establish vibrant communities of practice to transparently open model enhancements and to develop capacities, thereby lowering the entrance barriers to the established IAM community.

| | | |
|---|---|---|
| **ICCS** | INSTITUTE OF COMMUNICATIONS AND COMPUTER SYSTEMS | EL |
| **BC3** | ASOCIACION BC3 BASQUE CENTRE FOR CLIMATE CHANGE - KLIMA ALDAKETA IKERGAI | ES |
| **CESAR** | KRATENA KURT | AT |
| **CICERO** | CICERO SENTER FOR KLIMAFORSKNING | NO |
| **CYI** | THE CYPRUS INSTITUTE | CY |
| **E4SMA** | ENERGY ENGINEERING ECONOMIC ENVIRONMENT SYSTEMS MODELING AND ANALYSIS SRL | IT |
| **HOLISTIC** | HOLISTIC IKE | EL |
| **COMILLAS** | UNIVERSIDAD PONTIFICIA COMILLAS | ES |
| **ISINNOVA** | ISTITUTO DI STUDI PER L'INTEGRAZIONE DEI SISTEMI (I.S.I.S) - SOCIETA'COOPERATIVA | IT |
| **SEURECO** | SEURECO SOCIETE EUROPEENNE D'ECONOMIE SARL | FR |
| **UM** | UNIVERSITEIT MAASTRICHT | NL |
| **ESMIA** | ESMIA CONSULTANTS INC. | CA |
| **USMF** | THE UNIVERSITY OF MARYLAND FOUNDATIION INC | US |
| **UMD** | UNIVERSITY OF MARYLAND | US |
| **EPFL** | ECOLE POLYTECHNIQUE FEDERALE DE LAUSANNE | CH |
| **ETH** | EIDGENOESSISCHE TECHNISCHE HOCHSCHULE ZUERICH | CH |
| **UNIBAS** | UNIVERSITAT BASEL | CH |
| **Imperial** | IMPERIAL COLLEGE OF SCIENCE TECHNOLOGY AND MEDICINE | UK |
| **Oxford** | THE CHANCELLOR, MASTERS AND SCHOLARS OF THE UNIVERSITY OF OXFORD | UK |
| **UCL** | UNIVERSITY COLLEGE LONDON | UK |

# Executive Summary

Household heterogeneity has gained importance in economic modelling, both in macroeconomics as well as in Integrated Assessment Models (IAMs). The main issue motivating a shift from the 'representative consumer' paradigm to different groups or types of households is the need for assessing socioeconomic impacts of decarbonisation, which has become a critical issue in the policy debate.

To this end, the deliverable describes the construction of a detailed multi-household database (the DIAMOND EU household database) for the EU27 countries to support the integration of heterogeneity in selected models of DIAMOND (GCAM-Europe, NEMESIS-World, and GEMINI-E3-EU). First, the multi-household database is checked by analysing the variables and their distribution between household groups and comparing these results with published databases of household income and consumption. Second, integration of the database to the models tis performed by taking the already existing household model structures in these models as a starting point and analysing the resulting changes from the new database.

The new household database is based on EU microdata from the European Union Statistics on Income and Living Conditions (EU-SILC) and the Household Budget Survey (HBS). To accomplish the data fusion of EU-SILC and HBS, we follow an enhanced and updated version of the two-stage non-parametric strategy proposed in Kratena et al. (2021). This approach allows us to merge household disposable income from EU-SILC and total consumption expenditure from HBS, as well as all their underlying components, into a synthetic dataset for the 2015 reference year, covering all EU-27 countries. This database contains 109 variables, of which 26 relate to income and 83 to consumption expenditure, and features income distribution by household income deciles. Once income deciles-based data have been embedded in the various modelling frameworks, the DIAMOND EU household database can be readily extended to include other dimensions such as urban/rural, gender, or household size.

Tailored approaches were used to integrate the database to the three IAMs. In the case of GCAM-Europe, the integration process combines the information on country-level residential energy consumption (Eurostat) with the disaggregated residential energy consumption by fuel and income decile from the DIAMOND EU household database. In the new version of the GEMINI-E3-EU model, household heterogeneity has already been introduced, based on publicly available data from Eurostat, acknowledging that different households experience economic shocks and climate impacts differently. This heterogeneity significantly influences overall outcomes, such as consumption, investment, and welfare. In order to integrate the DIAMOND EU household database, a reconciliation process between the published Eurostat HBS data, as well as with those income and consumption data (from Eurostat), and the currently used data in the GEMINI-E3-EU database took place. That has been achieved by an optimisation procedure, where saving rates by quintile are harmonised with Eurostat data while maintaining the actual aggregate household saving rate in GEMINI-E3-EU. This model version has been used to run preliminary policy scenarios of carbon taxation with different compensation schemes, showing interesting results that motivate further research in the direction of household heterogeneity. In the current version of the NEMESIS-World model, households are represented by a single representative household. In integrating household heterogeneity in NEMESIS-World, a comparison between the DIAMOND EU household database and the distributional national accounts by Eurostat has been carried out. As these two datasets differ, a mapping between the DIAMOND EU household database, the distributional national accounts by Eurostat, and the NEMESIS-World household module was set up. The end result was a dataset of absolute income and expenditure data for each NEMESIS-World income or consumption category, disaggregated by household deciles for both the DIAMOND EU household database and the distributional national accounts by Eurostat. These distributions are then integrated into the NEMESIS-World model.

Funded by the European Union

# Contents

1 Introduction ..... 1
2 The household database in DIAMOND ..... 2
2.1 Using microdata to incorporate structural heterogeneity in modelling frameworks ..... 2
2.2 The DIAMOND EU household database ..... 4
3 Integrating the household dataset in different IAMs ..... 10
3.1 Integration of the household data in GCAM-Europe ..... 10
3.1.1 Representation of households in current GCAM-Europe version (GCAM-Europe-v7.2.0) ..... 10
3.1.2 Improving the representation of household heterogeneity in GCAM-Europe using empirical data 13
3.2 Integration of the household dataset in GEMINI-E3 ..... 15
3.2.1 Representation of households in the GEMINI-E3 standard version ..... 15
3.2.2 Including household heterogeneity in GEMINI-E3-EU ..... 17
3.2.3 Reconciling the GEMINI-E3-EU Input database with the DIAMOND EU household database ..... 18
3.2.4 Running preliminary scenarios ..... 22
3.3 Integration of the household dataset in NEMESIS-World ..... 24
3.3.1 State of the art and objectives ..... 24
3.3.2 Data overview ..... 27
3.3.3 Methodology ..... 32
4 Conclusion and future research ..... 36
ANNEX ..... 37
Bibliography ..... 39

# Table of Figures

Figure 1. HY020 - Total disposable household income in current euros per household ..... 8
Figure 2. EUR_HE00 - Total consumption expenditure in current euros per household ..... 9
Figure 3. Residential floorspace per period, region, and consumer group (left panel) in GCAM-v7.1, and comparison to the previous GCAM core (right panel) (BM2) ..... 11
Figure 4. Residential heating output per period, region, and consumer group in a baseline scenario (EJ) ..... 13
Figure 5. New CES Nesting Structure for Consumption ..... 16
Figure 6. Representation of the Economic Sectors in GEMINI-E3-EU ..... 16
Figure 7. Households' Heterogeneity in GEMINI-E3-EU ..... 17
Figure 8. Aligning Saving Rates of France and Spain ..... 22
Figure 9. $CO_2$ Emissions Change from Households (in percentage (left) and Mt $CO_2$ (right)) ..... 23
Figure 10. Welfare Change of Household Consumption (in percentage) ..... 24
Figure 11. Modelling Framework for the Household Sector within the NEMESIS model ..... 25
Figure 12. Illustration of Households' Representation in Germany before and after the Modelling Improvements for the calibration year ..... 26
Figure 13. Eurostat vs. DIAMOND: Distribution of Household Income Items across Deciles (Belgium, France, Germany, and Spain) ..... 31
Figure 14. Eurostat vs. DIAMOND: Distribution of Household Consumption Categories across Deciles (Belgium,

France, Germany, and Spain)..... 32
Figure 15. Eurostat vs. DIAMOND: Distribution of Household Income Items across Deciles (EU Member States) 37
Figure 16. Eurostat vs. DIAMOND: Distribution of Household Consumption Categories across Deciles (EU Member States)..... 38

# Table of Tables

Table 1. Variables included in the DIAMOND EU household database..... 5
Table 2. Summary of the DIAMOND EU household database variables used in Phase 1 of the enhanced implementation of consumer heterogeneity into GCAM-Europe..... 14
Table 3. Mapping Household Consumption Categories into GEMINI-E3-EU Classification..... 18
Table 4. Mapping Households' Income Items into GEMINI-E3-EU Classification..... 20
Table 5: Summary of Data Reconstruction References..... 28
Table 6: Comparison of distributional national accounts by Eurostat and the DIAMOND EU household database..... 29
Table 7: Mapping between Income Items of Distributional National Accounts by Eurostat or the DIAMOND EU household database and their Equivalent in NEMESIS..... 33
Table 8: Mapping between Consumption Categories of Distributional National Accounts by Eurostat and the DIAMOND EU household database and their Equivalent in NEMESIS..... 33

Funded by the European Union

# 1 Introduction

Integrated Assessment Models have typically dealt with households through single representative agents in the evaluation of climate policies, a process with significant limitations as has been acknowledged in the modelling literature (Emmerling & Tavoni, 2021; Rao et al., 2017). In particular, economic shocks at the aggregate or disaggregated level (e.g., only in one industry or good) could have a uneven impact on different household groups, considering their socioeconomic position. Climate and energy policies will also affect household types in a different way, both in the static perspective of how households are hit by a measure, as well as in the dynamic perspective of how different household types can adjust to these policies through behavioural reactions. Policies might lead to significant social inequalities as well as to different aggregate outcomes, e.g., when certain household types are locked into their consumption patterns, while others are not. These aspects of policies are a top priority in the European climate policy agenda and can only be addressed in models that incorporate relevant aspects of household heterogeneity.

In this deliverable, we present a detailed multi-household database constructed for the EU 27 countries and then built in into three IAMs applied in DIAMOND, namely GCAM-EU, GEMINI-E3-EU and NEMESIS-World. The database and the integration process form the empirical base for extending the household sector in Integrated Assessment Models (IAMs) in DIAMOND.

One crucial variable by which different household types can be identified is disposable income. In this database this is measured by deciles of disposable income. By statistical matching methodologies, consumption data from another statistical source is assigned to the household types by income. Also, in each IAM incorporating the household database, the microeconomic structures of the database (for income and consumption) are adjusted to the macroeconomic totals of these variables. That finally yields a statistically consistent database that can be directly used in each IAM. The integration of household heterogeneity in several IAMs that are used in the DIAMOND project then allows for analysing the social dimension of energy and climate policies. A preliminary simulation exercise clearly reveals the potential of this approach.

This deliverable lays the base for introducing household heterogeneity in IAMs used in DIAMOND. Section 2 describes the household dataset for EU countries that has been constructed and integrated into several IAMs applied in the project. This description comprises an outline of the methodology applied (statistical matching on individual data) as well as a presentation of the data. In Section 3, the methodology of integration of the dataset in three DIAMOND models (GCAM-Europe, GEMINI-E3-EU, and NEMESIS-World) is presented. Finally, Section 4 fleshes out concluding remarks and provides suggestions for immediate future research.

Funded by
the European Union

# 2 The household database in DIAMOND

IAMs have traditionally represented the household sector using macro-aggregates, a process that typically implies the use of a single representative household per region or country. To address the limitations of this approach, researchers have incorporated household-level microdata—most often from income and expenditure surveys—into modelling frameworks to capture heterogeneity in income sources, consumption patterns, and distributional impacts (Budolfson et al., 2021; Landis et al., 2021).

In the case of the models employed in DIAMOND (namely, GCAM-Europe, NEMESIS-World, and GEMINI-E3-EU), these focus on climate change mitigation policies, and as such the analysis primarily focuses on assessing the distributional impacts of policies on EU households. In this section, a detailed multi-household database is constructed for EU countries to support the integration of such heterogeneity in these models. It is based on EU microdata that provide detailed coverage of multiple variables in the variables related to income and consumption. Specifically, the database is created by integrating information from the European Union Statistics on Income and Living Conditions (EU-SILC) and the Household Budget Survey (HBS). In the following sections, the surveys and the methods used for their coherent integration are described, and the resulting database is presented.

## 2.1 Using microdata to incorporate structural heterogeneity in modelling frameworks

Two of the most important EU surveys, namely EU-SILC and HBS, are used to incorporate structural heterogeneity into the models in the DIAMOND project. The microdata used are not publicly available. They were accessed for scientific purposes following Eurostat's application procedure for confidential microdata (Eurostat, 2025b). These surveys are provided by Eurostat as separate surveys, and a brief description of each is provided below:

- EU-SILC - European Union Statistics on Income and Living Conditions (Eurostat, n.d.-a): This survey collects cross-sectional data on household income, including flows such as earnings, transfers, and taxes, as well as indicators of poverty, social exclusion, and living conditions. It also provides detailed information on the socio-demographic characteristics of households and their members. The survey is conducted annually within a harmonised framework that ensures comparability across EU countries.
- HBS - Household Budget Survey (Eurostat, n.d.-b): This survey provides cross-sectional data on household consumption expenditure on goods and services, reported in monetary units for all items and in physical units for selected categories. The consumption data are classified according to the Classification of Individual Consumption According to Purpose (COICOP) structure, and the survey includes socio-demographic characteristics of households and their members. Eurostat has released three harmonised HBS data waves (2010, 2015, 2020), which improve cross-country comparability; however, differences in survey frequency, timing, content, and methodology—stemming from its basis in a voluntary agreement—still lead to some gaps and inconsistencies across EU countries.

The data used here are centred on the reference year 2015. This choice is motivated by considerations primarily related to data availability and the alignment of microdata with the models' base year (see D3.2 for more details). Among the three harmonised HBS reference years (2010, 2015, 2020), the 2015 wave was selected because it covers most EU countries and corresponds most closely to the base year of the DIAMOND models that are to be enhanced with household heterogeneity. Contrary, the 2020 HBS release may be affected by the temporary impact of the COVID-19 pandemic. Surveys conducted close to that shock may not be representative of usual household consumption behaviour and are therefore less suitable for constructing structural heterogeneity that reflects underlying trends. The 2010 wave, on the other hand, may reflect the economic downturn following the 2008

Funded by the European Union

financial crisis, making it less representative of more recent structural consumption patterns. For these reasons, 2015 was chosen as the most appropriate reference year for the HBS microdata. For consistency, we also select the EU-SILC wave that best matches the HBS reference year. Because in most countries the EU-SILC income variables refer to the year preceding the interview, the EU-SILC reference year adopted is the HBS year plus one.

EU household income and consumption data come from separate surveys. Therefore, traditional studies have introduced household heterogeneity into models by using group-level aggregates from each source. For example, EU-SILC (earnings, transfers, taxes) and HBS (Classification of individual consumption by purpose; COICOP consumption) are each aggregated into income deciles, urban/rural groups, or regions, and then merged (Landis et al., 2021). However, both surveys differ in sampling, concepts and measurement quality, and aggregating them separately can therefore lead to inconsistencies (Coli & Tartamella, 2017). To overcome this problem, a statistical matching process needs to be performed. Here, HBS consumption is imputed into EU-SILC (and vice versa) using shared variables such as household size, region, tenure, or urban/rural status. This produces a synthetic dataset with consistent joint distributions between income and consumption at the household level. Reliable aggregates can then be derived for different household types and mapped into IAMs.

To accomplish the data fusion of EU-SILC and HBS, we follow an enhanced and updated version of the two-stage non-parametric strategy proposed in Kratena et al. (2021). This approach allows us to merge household disposable income from EU-SILC and total consumption expenditure from HBS, as well as all their underlying components into a synthetic dataset for the 2015 reference year, covering all EU-27 countries except Austria, for which Eurostat provides no data for that year[1]. The strategy has two main stages. In the first stage, reliable proxies for the target variables are estimated in both surveys using statistical-learning (non-parametric) models. In the second stage, a nearest-neighbour (non-parametric) algorithm based on the Euclidean distance metric is used to match households, drawing on the previously estimated proxies and other common variables after transforming the latter to align with the dimensional space of the estimates.

The strategy involves multiple steps. First, we choose the direction of matching (HBS as donor and EU-SILC as recipient) so that household-level consumption baskets are imputed into EU-SILC. We then define the targets to be transferred from HBS (total consumption and its COICOP components) and harmonise disposable-income concepts across surveys (EU-SILC income is generally more reliable, so we briefly "back-impute" it into HBS to align definitions where needed)[2]. Next, we assemble a set of variables that are common to both sources (namely, household size and type, age and gender of the reference person, population density, tenure status, car ownership, economic status, main income source, and income percentiles) and standardise their definitions, check comparability, and keep only those that behave similarly across surveys. Using HBS as the training source, we build predictive (XGBoost) models for each consumption target with the common variables and income as predictors, then generate high-quality proxies for all targets in both datasets. To strengthen like-with-like pairing, we stratify EU-SILC and HBS into weighted quantiles of predicted total consumption and restrict matches to records within equivalent strata. Within each stratum, we apply nearest-neighbour matching based on a combined similarity measure that draws on both the predicted consumption proxies and the harmonised common variables. Because both surveys use complex sampling, we incorporate weights and design variables in the stratification and matching

[1] To close this gap, Austria's national HBS data are being mapped to Eurostat's harmonised HBS COICOP classification and variable categories. This (unexpected) task is still in progress.

[2] For Italy, the HBS lacks a disposable-income variable. Accordingly, we (i) impute an income ranking (percentiles/deciles) using EU HBS benchmarks, and then (ii) impute a consistent disposable-income distribution from the Italian EU-SILC.

to respect survey designs and improve representativeness. Then, we run several candidate matchings (varying the number of strata and the extent to which donors can be reused) and select the best fused dataset using a composite quality index that compares imputed versus original distributions, as well as key statistics on the "overall" and "by categories" common variables used in the matching. The resulting synthetic integrated EU-SILC–HBS microdata is finally validated to ensure that they preserve realistic joint distributions of income and consumption domains at the household level.

## 2.2 The DIAMOND EU household database

By aggregating the synthetic integrated EU-SILC–HBS 2015 microdata from the previous sub-section by household income deciles[3], we create a tailored EU household database to support the splitting of the representative household in the DIAMOND project's IAMs. Only household income deciles are included in the database, as this is the most relevant aggregation for assessing the distributional impacts of climate-related policies and is therefore the priority direction in terms of introducing household heterogeneity in models for the DIAMOND modelling teams. However, once income deciles-based household heterogeneity is embedded in the various modelling frameworks, the DIAMOND EU household database can be readily extended to include other dimensions such as urban/rural, gender, or household size, which may also matter for climate-policy scenario assessment.

The DIAMOND EU household database contains 109 variables, of which 26 relate to income and 83 to consumption expenditure. More specifically, it includes household-level income variables (totals and components such as earnings, self-employment, pensions and other benefits, taxes, transfers, property income, and imputed rent) and detailed consumption expenditures classified by COICOP (e.g., food, housing and rents, utilities—including electricity, gas, and heat—transport and fuels, health, education, recreation, and other services). A list of all variables can be found in Table 1 (see Eurostat, 2025a).

Data are reported as household averages in current euros by household type (currently, income deciles by country). Because household weights (group sizes) are included in the dataset, these averages can be scaled to totals by multiplying each average by the sum of weights for all households in the group. This, in turn, allows computing each group's share of a given micro-aggregate and then use those shares to split the corresponding macro-aggregate. Figures 1 and 2 present the two main variables included in the database as examples of the type of information provided.

For modelling frameworks that do not exactly match the DIAMOND EU household database reference year (i.e., 2015), it is recommended to apply an appropriate price deflator or inflator to align values and avoid inconsistencies due to price changes. Furthermore, we note that if the reference-year gap spans many years, potential structural changes may also affect the reliability of the shares used to introduce household heterogeneity in the models.

Finally, it is important to note that during the interaction between the teams involved in data generation and model integration, several data gaps/issues were identified for certain variables in specific countries, stemming from the original microdata. At this stage of the research, including the very few missing countries is the priority, so that all EU countries are covered in the DIAMOND EU household database. To generate household heterogeneity in the meantime, these gaps can be bypassed by modellers using shares from partners (similar countries) in the absence of better information. However, over the remainder of the project we will aim to fill these

[3] Income deciles were constructed by ranking households by per-capita disposable income and applying the standard method to form household-weighted income deciles.

gaps by imputing the missing information using a statistical learning model (Chen & Guestrin, 2016).

**Table 1.** Variables included in the DIAMOND EU household database

| Code | Name |
|---|---|
| HY010 | Total household gross income |
| HY020 | Total disposable household income |
| PY010G_hh | Employee cash or near cash income |
| PY020G_hh | Non-cash employee income |
| PY021G_hh | Company car |
| PY050G_hh | Cash profits or losses from self-employment |
| PY080G_hh | Pension from individual private plans |
| PY090G_hh | Unemployment benefits |
| PY100G_hh | Old-age benefits |
| PY110G_hh | Survivor benefits |
| PY120G_hh | Sickness benefits |
| PY130G_hh | Disability benefits |
| PY140G_hh | Education-related allowances |
| HY030G | Imputed rent |
| HY040G | Income from rental of a property or land |
| HY050G | Family/children-related allowances |
| HY060G | Social exclusion not elsewhere classified |
| HY070G | Housing allowances |
| HY080G | Regular interhousehold cash transfer received |
| HY090G | Interest, dividends, profit from capital investments in unincorporated business |
| HY100G | Interest repayments on mortgage |
| HY110G | Income received by people aged under 16 years |
| HY120G | Regular taxes on wealth |
| HY130G | Regular interhousehold cash transfer paid |
| HY140G | Tax on income and social contributions |
| HY170G | Value of goods produced for own consumption |
| EUR_HE00 | Total consumption expenditure |
| EUR_HE011 | Food |
| EUR_HE012 | Non-alcoholic beverages |
| EUR_HE021 | Alcoholic beverages |
| EUR_HE022 | Tobacco |
| EUR_HE023 | Narcotics |
| EUR_HE031 | Clothing |
| EUR_HE032 | Footwear |
| EUR_HE041 | Actual rentals for housing |
| EUR_HE042 | Imputed rentals for housing |
| EUR_HE043 | Maintenance and repair of the dwelling |
| EUR_HE044 | Water supply and miscellaneous services relating to the dwelling |
| EUR_HE0451 | Electricity |
| EUR_HE04510 | Electricity |
| EUR_HE0452 | Gas |
| EUR_HE04521 | Natural gas and town gas |
| EUR_HE04522 | Liquefied hydrocarbons-butane-propane-etc. |
| EUR_HE0453 | Liquid fuels |
| EUR_HE04530 | Liquid fuels |

Funded by
the European Union

| Code | Name |
|---|---|
| EUR_HE0454 | Solid fuels |
| EUR_HE04541 | Coal |
| EUR_HE04549 | Other solid fuels |
| EUR_HE0455 | Heat energy |
| EUR_HE04550 | Heat energy |
| EUR_HE051 | Furniture and furnishings, carpets and other floor coverings |
| EUR_HE052 | Household textiles |
| EUR_HE053 | Household appliances |
| EUR_HE054 | Glassware tableware and household utensils |
| EUR_HE055 | Tools and equipment for house and garden |
| EUR_HE056 | Goods and services for routine household maintenance |
| EUR_HE061 | Medical products, appliances and equipment |
| EUR_HE062 | Out-patient services |
| EUR_HE063 | Hospital services |
| EUR_HE071 | Purchase of vehicles |
| EUR_HE0721 | Spare parts and accessories for personal transport equipment |
| EUR_HE0722 | Fuels and lubricants for personal transport equipment |
| EUR_HE07221 | Diesel |
| EUR_HE07222 | Petrol |
| EUR_HE07223 | Other fuels for personal transport equipment |
| EUR_HE07224 | Lubricants |
| EUR_HE0723 | Maintenance and repair of personal transport equipment |
| EUR_HE0724 | Other services in respect of personal transport equipment |
| EUR_HE0731 | Passenger transport by railway |
| EUR_HE07311 | Passenger transport by train |
| EUR_HE07312 | Passenger transport by underground and tram |
| EUR_HE0732 | Passenger transport by road |
| EUR_HE07321 | Passenger transport by bus and coach |
| EUR_HE07322 | Passenger transport by taxi and hired car with driver |
| EUR_HE0733 | Passenger transport by air |
| EUR_HE07331 | Domestic flights |
| EUR_HE07332 | International flights |
| EUR_HE0734 | Passenger transport by sea and inland waterway |
| EUR_HE07341 | Passenger transport by sea |
| EUR_HE07342 | Passenger transport by inland waterway |
| EUR_HE0735 | Combined passenger transport |
| EUR_HE07350 | Combined passenger transport |
| EUR_HE0736 | Other purchased transport services |
| EUR_HE07361 | Funicular, cable-car and chair-lift transport |
| EUR_HE07362 | Removal and storage services |
| EUR_HE07369 | Other purchased transport services NEC |
| EUR_HE081 | Postal services |
| EUR_HE082 | Telephone and telefax equipment |
| EUR_HE083 | Telephone and telefax services |
| EUR_HE091 | Audio-visual, photographic and information processing equipment |
| EUR_HE092 | Other major durables for recreation and culture |
| EUR_HE093 | Other recreational items and equipment, gardens and pets |
| EUR_HE094 | Recreational and cultural services |

Funded by
the European Union

| Code | Name |
|---|---|
| EUR_HE095 | Newspapers, books and stationery |
| EUR_HE096 | Package holidays |
| EUR_HE101 | Pre-primary and primary education |
| EUR_HE102 | Secondary education |
| EUR_HE103 | Post-secondary nontertiary education |
| EUR_HE104 | Tertiary education |
| EUR_HE105 | Education not definable by level |
| EUR_HE111 | Catering services |
| EUR_HE112 | Accommodation services |
| EUR_HE121 | Personal care |
| EUR_HE122 | Prostitution |
| EUR_HE123 | Personal effects NEC |
| EUR_HE124 | Social protection |
| EUR_HE125 | Insurance |
| EUR_HE126 | Financial services NEC |
| EUR_HE127 | Other services NEC |

Funded by
the European Union

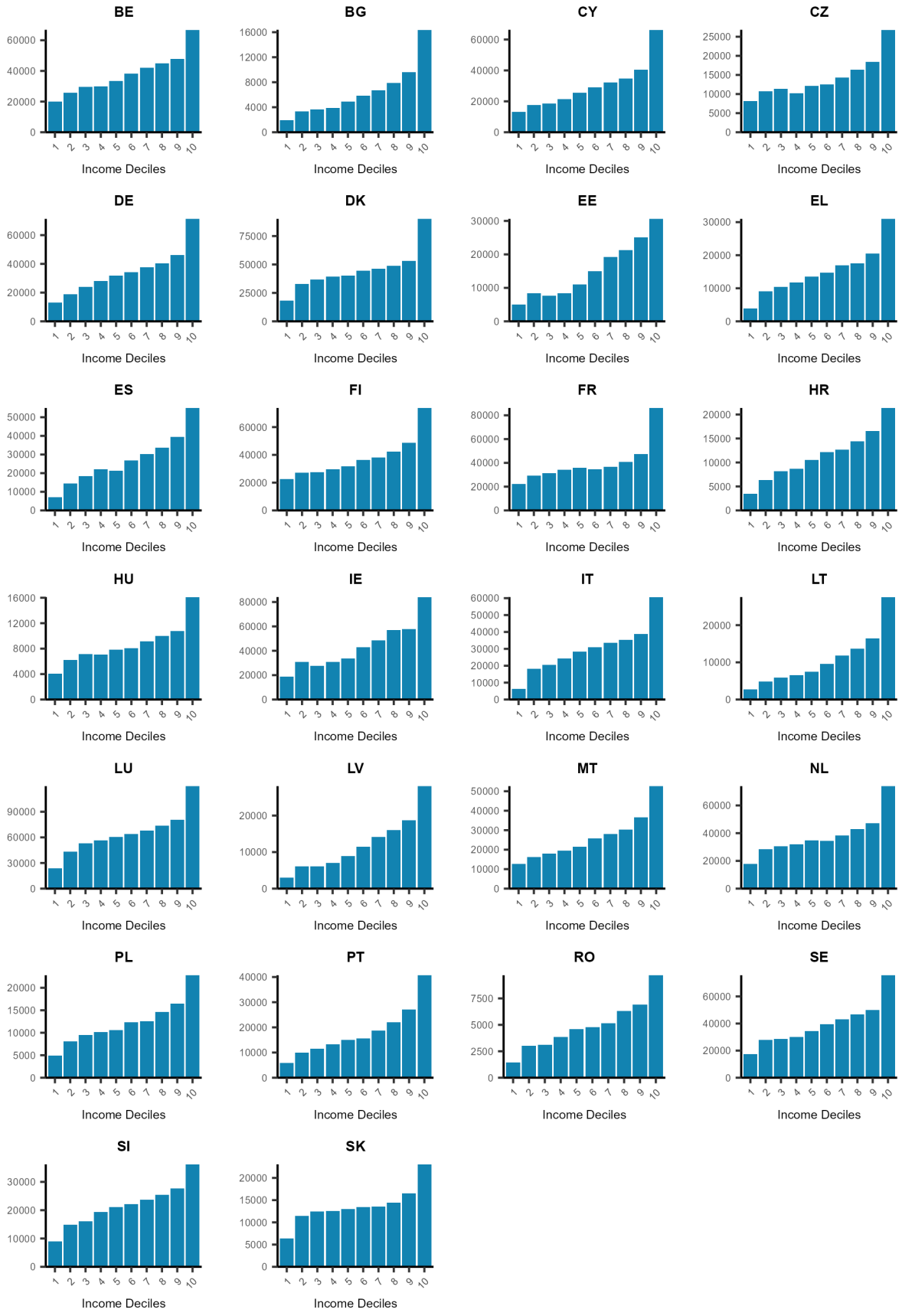



**Figure 1.** HY020 - Total disposable household income in current euros per household

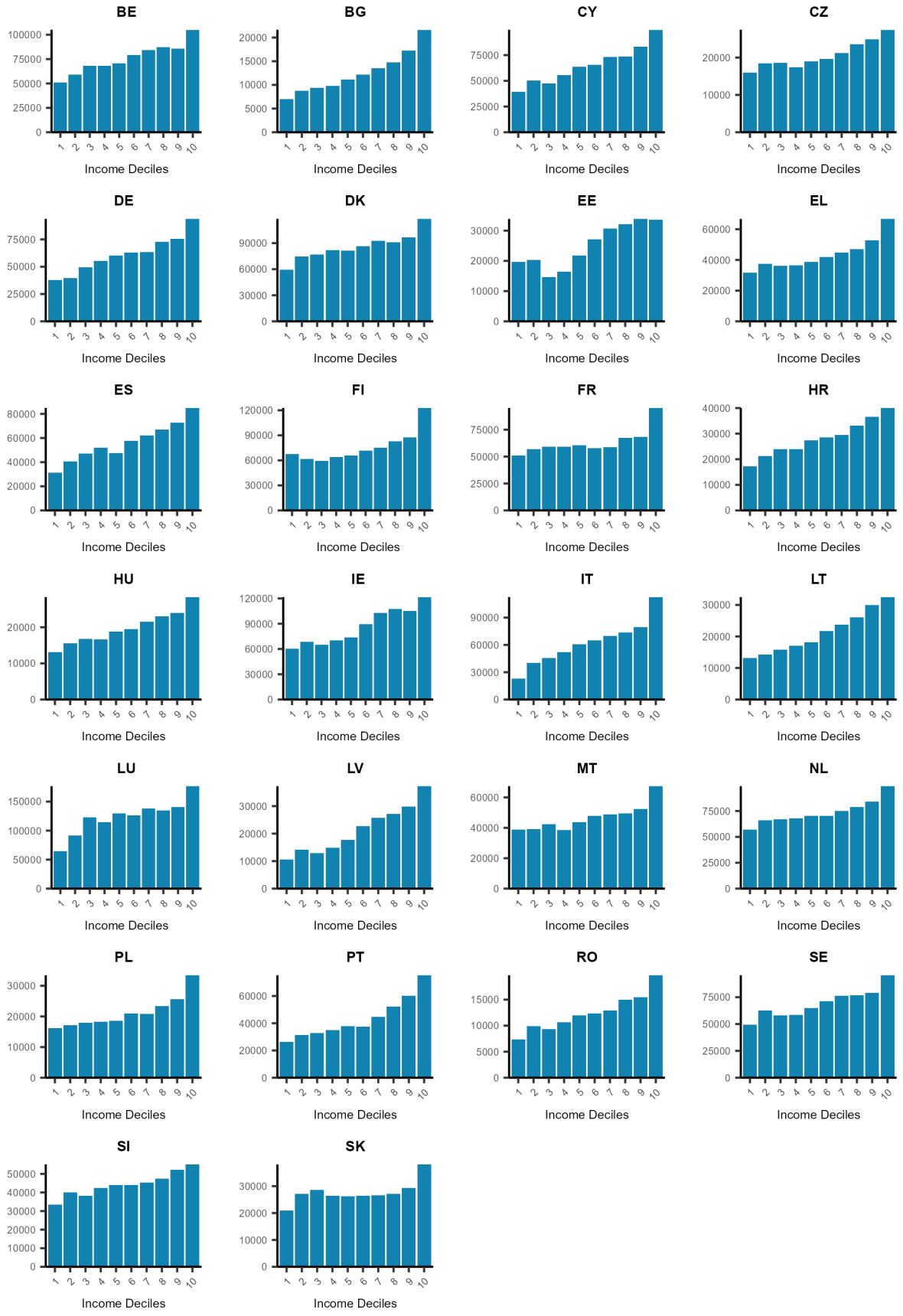



**Figure 2.** EUR_HE00 - Total consumption expenditure in current euros per household

# 3 Integrating the household dataset in different IAMs

## 3.1 Integration of the household data in GCAM-Europe

### 3.1.1 Representation of households in current GCAM-Europe version (GCAM-Europe-v7.2.0)

Historically, residential demand for floorspace and energy services in GCAM was calculated for a single representative consumer, which is still a common practice in integrated assessment modelling (Emmerling & Tavoni, 2021). In GCAM-v7.1 (Bond-Lamberty et al., 2024), consumer heterogeneity has only been introduced in the building sector, allowing residential floorspace demand and energy service use to be projected by income deciles within each GCAM region. GCAM-Europe-v7.2.0 follows the same implementation approach as GCAM-7.1, which is explained in this section and detailed in the "Core Model Proposal (CMP)" report recently published by the GCAM development team[4].

In the global version of GCAM, residential floorspace and energy consumption are calibrated from global energy balances and supplementary global datasets that provide information at national and/or regional, but not subnational, scales. Likewise, while GCAM-Europe is calibrated using Eurostat energy balances (Eurostat, 2025c) that offer detailed country- and service-specific information, data is provided at the national level and do not include subnational detail, such as distinctions among different consumer groups. In this framework, given the lack of harmonised data at subnational scales, the allocation of national floorspace and energy quantities across multiple consumer groups within each region was computed using the corresponding functional forms for floorspace and residential energy, as detailed below. In addition, the main principle for the implementation of multiple consumers was that "consumers within a region demand the same amount of floorspace and per capita energy service when they reach the same income level".

Regarding residential floorspace, the Gompertz-type function used for the estimation of future per capita residential floorspace is by design flexible to the implementation of multiple consumers. First, floorspace demand is estimated for each income decile using the Gompertz function and then re-aggregated to obtain an estimated national total. A "bias-adder" parameter is subsequently calculated to reconcile differences between this estimate and the observed national values. This bias adder is distributed equally across all consumers, ensuring that each follows the same functional form for projecting future residential floorspace. Apart from the calibration adjustment, a lower bound is imposed so that floorspace cannot fall below base-year levels. Without this constraint, floorspace could decrease when per capita GDP in a given decile declines from one period to the next which would otherwise imply an unrealistic relocation or even "destruction" of existing building stock. While this situation is rarely observed with a single representative consumer, it becomes more common if multiple consumers are implemented.

[4] https://github.com/JGCRI/gcam-doc/blob/gh-pages/cmp/362-Multiple_consumers_in_residential_buildings.pdf.

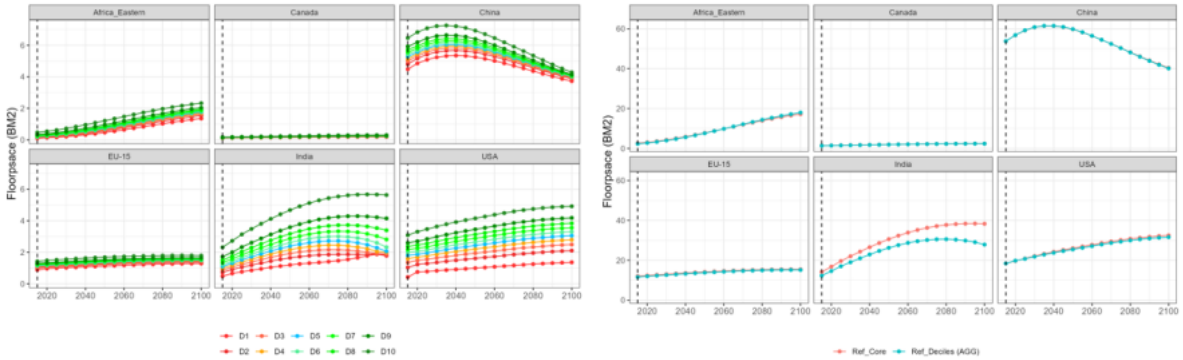



**Figure 3.** Residential floorspace per period, region, and consumer group (left panel) in GCAM-v7.1, and comparison to the previous GCAM core (right panel) (BM2)

Source: CMP #362.

The final functional form for the estimation of residential per capita floorspace (*F*) in period (*t*), region *I* and decile (*i*) is summarised in the following equation:

$$F_{t,r,i} = \left(UnadjSat_r - a * \ln\left(PD_{t,r}\right)\right) * \exp\left(-b * \exp\left(-c * \ln\left(GDPpc_{t,r,i}\right)\right)\right) + k_r$$

- *UnadjSat* is the unadjusted satiation value that represents the maximum per capita floorspace value a consumer demands at their maximum income level,
- *PD* is the population density,
- *k* is a bias correction adder,
- *a*, *b*, and *c* are parameters applied to the population density and per capita income.

For residential energy service demand, considering the lack of harmonised subnational data for residential energy consumption, the allocation of national energy quantities across different consumers (i.e., income deciles) is based on the corresponding demand functions. We first use the demand functions to estimate the income decile-level consumption within regions. The aggregation of the estimated energy consumption across income deciles does not match the observed national (or regional) totals. To address this, a bias-correction adder is calculated as the difference between estimated and observed values. However, in certain cases, this correction can result in negative demand for some deciles, requiring further adjustments to the bias-correction parameter to ensure consistency in both decile-level calibration and forward-looking analysis. These further adjustments are extensively detailed in (Sampedro et al., 2024) and can be summarised in six steps:

1. Calculate e*stimated* decile-level consumption.
2. Calculate total *estimated* regional consumption (sum of all deciles).
3. Calculate *estimated* income decile energy shares of total regional consumption.
4. Compute difference between observed and estimated regional consumption (bias correction)
5. Apply bias correction to each decile.
6. Adjust bias correction to avoid negative demands in some low-income deciles: phase bias correction out over three model periods.

Beyond calibration adjustments, the demand functional forms for estimating future residential energy demand also need to be adjusted for a coherent implementation of the multiple consumers. The demand for modern energy services needs to be adjusted to ensure that, as incomes increase for lower deciles, all else equal, they will demand the same level of services as higher deciles did at the respective income level, with no other tuning parameters affecting the shape of their demand function. Demand for heating (*h*), in period *t*, region *r*, and decile *I* is defined as:

$$h_{t,r,i} = k_{h,r} * \left(HDD_{t,r} * \eta_{r,t,i} * R_r + \lambda_{h,r} * IG_{t,r,i}\right) * \left[1 - \exp\left(-\frac{\ln(2)}{\mu_{h,r}} * \frac{GDPpc_{t,r,i}}{Ph_{t,r}}\right)\right] + a_{h,r}$$

Where:

- *k* represents the satiation level coefficient, which is based on service output per unit of floorspace in final calibration year (2015), exogenous service-specific multipliers, and climate variables (HDD and CDD),
- *HDD* and *CDD* refer to heating and cooling degree days that indicate how many degrees, and for how many days the temperature was below or above a determined level. The values are set exogenously for different model periods (up to 2100) and regions,
- $\eta$ is the shell conductivity that represents building thermal efficiency,
- *R* is the floorspace-to-surface ratio, and is set to 5.5 in all periods and regions,
- *IG* represent internal gains associated with the heat released by non-thermal energy services (e.g., appliances). Internal gains are adjusted for different regions using a scalar parameter ($\lambda$) based on their heating and cooling degree days,
- μ is the "satiation impedance", a shape parameter estimated in final calibration year and kept constant over the whole-time horizon,
- *GDPpc* refers to per capita income,
- *P* represents the average service price, and
- *a* is the bias-adder parameter, which represents the difference between the observed and estimated values.

Note that the two calibration parameters (k and μ) are estimated at region level. The prices used in the estimation of the calibration parameters (P) are equal for all consumer groups due to lack of subnational data. Building shell efficiency ($\eta$) is also differentiated across regions and consumer groups to reflect the tendency of wealthier regions and households to adopt more efficient buildings over time.

For traditional services, namely coal and traditional biomass (TradBio), the function is adjusted to ensure all consumers follow a consistent trajectory. A bias-adder parameter is then calculated and evenly distributed across consumers within each region, as summarised in the following equation:

$$d_{t,r,i,f,s} = \frac{X_{f,s}}{\frac{GDPpc_{t,r,i}}{P_{t,rf,s}} + Y_{f,s}} + a_{r,f,s}$$

Where:

- *s* is the service,
- *f* is fuel (either coal or traditional biomass),
- *P* is average service price,
- *X* and *Y* are calibration parameters estimated from historical data,
- *a* is the bias-adder

Figure 4 illustrates residential heating projections under the baseline scenario for a representative set of regions.

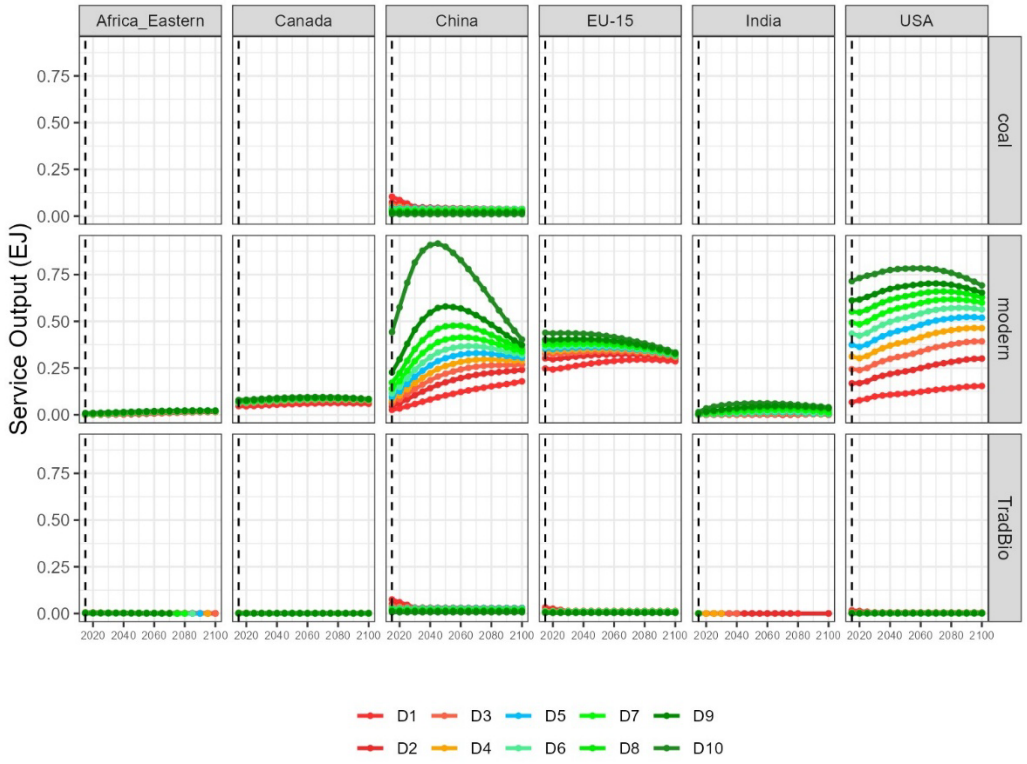



**Figure 4.** Residential heating output per period, region, and consumer group in a baseline scenario (EJ)

In a recently developed model version, GCAM_GRAPHICS (Sampedro, 2025), consumer heterogeneity has been extended beyond the residential sector to additional areas such as food and transportation (Sampedro et al., 2025). This implementation is still preliminary, relying on assumptions analogous to those used in the residential sector rather than on empirical data. Although it represents an important first step toward broadening the representation of multiple consumer groups, it has not yet been incorporated into the core version of the model nor GCAM-Europe.

Building on the foundations established in GCAM_GRAPHICS, and the integration of data from the DIAMOND EU household database (see section 3.1.2), GCAM-Europe adopts an enhanced, data-driven approach to integrate consumer heterogeneity across these sectors in a more systematic and empirically grounded manner.

### 3.1.2 Improving the representation of household heterogeneity in GCAM-Europe using empirical data

One of the key advantages of GCAM-Europe over the global (core) version of the model is its use of detailed country-specific data, which enhances model calibration and leads to more realistic representations and accurate projections. A central element of these model improvements involves replacing global data sources, such as IEA energy balances (IEA, 2022) with Europe-specific sources, such as Eurostat energy balances (Eurostat, 2025c). This regional tailoring allows the model to better capture the structural and behavioural differences within European economies, particularly in the energy sector.

To improve the representation of consumer heterogeneity in the model, we will adopt the same approach by replacing existing data sources and assumptions with Europe-specific sources and empirical evidence. In addition, we will extend the implementation of consumer heterogeneity beyond the residential sector to also include the food and transportation sectors, following a consistent, data-driven methodology. To that end, we will make use of the DIAMOND EU household database (see section 2.2).

In this framework, the first step (*Phase 1*) is to incorporate the information of the database, and replace the current assumptions included in GCAM-Europe to allocate country-level residential energy consumption (provided by Eurostat) across the different within-country consumer groups (currently income deciles). Eurostat (2025d) provides country-level disaggregated residential energy consumption by fuel for six different sectors, namely space cooling, space heating, water heating, cooking, lighting and electrical appliances, and "other end uses". Given the lack of more detailed subnational data, a complementary dataset, such as the DIAMOND EU household database, is required to allocate those quantities across the different deciles within each member state. This database provides income-decile level energy consumption information at monetary values, which can be directly transformed to physical values using average service prices. Since in GCAM consumer demand is modelled on a per-capita basis, we use the number of persons in each household group (currently income decile) to convert the database values from household averages to per-capita terms.

In addition, the database does not provide service-specific information. Therefore, the allocation of national quantities across deciles is based on the disaggregated residential energy consumption by fuel provided by the DIAMOND EU household database and then scaled to make the values match with the total service-and fuel-specific values provided by Eurostat. The categories in the database that will be used for the downscaling in Phase 1 are summarised in the following table:

**Table 2.** Summary of the DIAMOND EU household database variables used in Phase 1 of the enhanced implementation of consumer heterogeneity into GCAM-Europe

| Code | Name |
|---|---|
| EUR_HE0451 | Electricity |
| EUR_HE04510 | |
| EUR_HE0452 | Gas |
| EUR_HE04521 | Natural gas and town gas |
| EUR_HE04522 | Liquefied hydrocarbons-butane-propane-etc. |
| EUR_HE0453 | Liquid fuels |
| EUR_HE04530 | |
| EUR_HE0454 | Solid fuels |
| EUR_HE04541 | Coal |
| EUR_HE04549 | Other solid fuels |
| EUR_HE0455 | Heat energy |
| EUR_HE04550 | |

In a second step (Phase 2), we use the information of the DIAMOND EU household database to implement consumer heterogeneity in additional end-use sectors in GCAM-Europe, namely food, transportation, and potentially, municipal water demand. In terms of the food sector, the food demand by country and type (staples versus non-staples) is calibrated in GCAM-Europe based on FAO Food Balance Sheets[5]. In order to allocate food across different consumer groups within each country, we draw on the information on food available in the database ("EUR_HE011, Food") and complement it with additional assumptions that allow us to differentiate between staple and non-staple commodities. For the disaggregation of the transportation sector, we use the categories "EUR_HE071 - EUR_HE07369" in the database to allocate passenger transport demand, by country and mode, across different consumer groups within each country. Likewise, category EUR_HE044 is used to

[5] https://www.fao.org/faostat/en/#data/FBS

disaggregate the demand of municipal water.

Finally, an implementation pathway to be developed during (or potentially beyond) the course of the project is the introduction of a flexible framework for consumer heterogeneity. This would allow users to choose among different forms of consumer disaggregation, such as income groups, gender, urban–rural distinctions, or geographical subgroups (e.g., NUTS regions), and their intersections. Such flexibility will enable GCAM-Europe users to tailor the model structure to their specific research questions and policy needs. For instance, incorporating subnational geographical units like NUTS regions would support regional policymakers in the design and evaluation of targeted policies, ensuring that the model outcomes are both more granular and more relevant for decision-making. However, in addition to computational challenges, this flexible implementation also faces important data-related constraints. In particular, all resulting consumer groups must comply with Eurostat's representativeness criteria, which means that further validation procedures will be required to ensure adherence to legal requirements. Also, the model would require current data and future estimates on population and income in each disaggregated subgroup, ideally from official sources aligned with other data sources used in GCAM-Europe, such as Eurostat. This, in turn would make it necessary to take into account changes in the size of groups due to changes in the number as well as in the socioeconomic position of households.

## 3.2 Integration of the household dataset in GEMINI-E3

### 3.2.1 Representation of households in the GEMINI-E3 standard version

GEMINI-E3-EU, like most Computable General Equilibrium models applied to energy and climate policies, represents only one representative consumer (i.e., household) in its standard version. Household behaviour consists of three interdependent decisions: (1) labour supply, (2) savings, and (3) consumption of the various goods and services. In GEMINI-E3-EU, both labour supply and the rate of savings are assumed to be exogenous.

Demand for different commodities has prices of consumption and income (more precisely "spent" income, i.e., income after savings) as arguments and is derived from nested Constant Elasticity of Substitution (CES) utility functions. At the top level of the consumption function, households allocate their spending across three main categories: housing, transport, and other goods and services. Energy use is further divided into two components: transportation and housing. Within transportation, demand is distinguished between purchased services and own-use transport. The nested CES structure is represented in Figure 5.

Household income includes labour remuneration, capital and fixed factor remunerations (since firms are, by definition, owned by households in GEMINI-E3-EU), as well as various transfers between households and the two other economic agents represented in GEMINI-E3-EU, namely the government and the Rest of the World (ROW). These transfers include direct taxes and social benefits, as well as other financial transfers that are represented in aggregate form.

The transfers between these three economic sectors are represented in Figure 6. To evaluate these transfers, we have compiled several international economic databases.

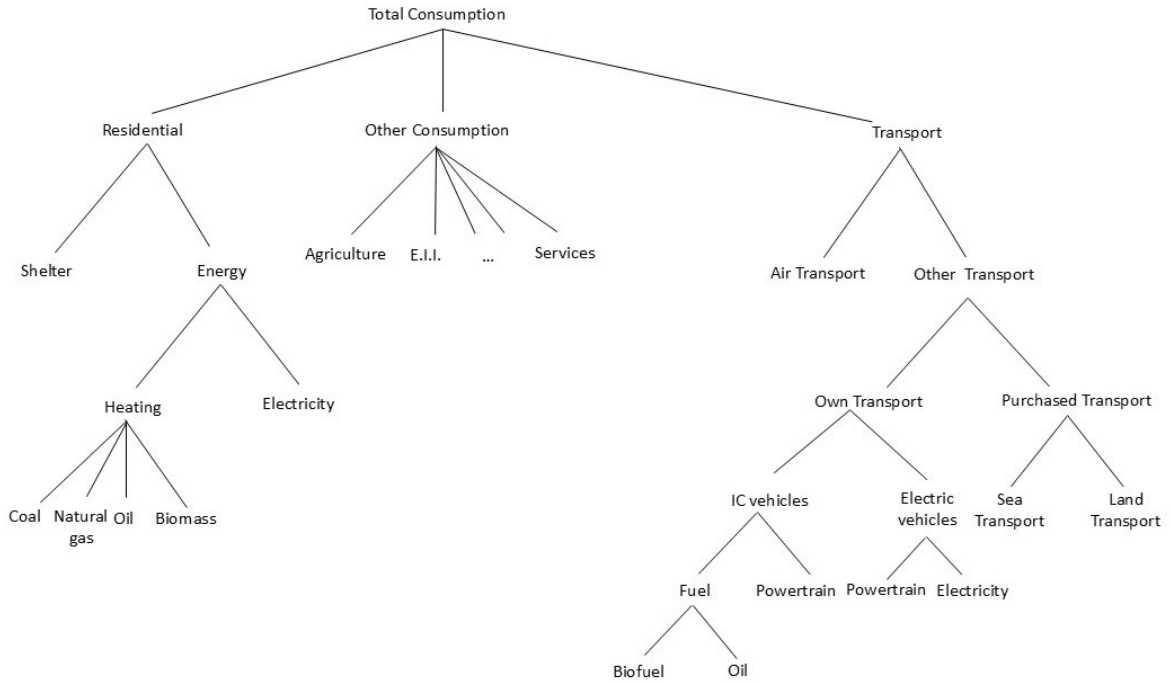



**Figure 5.** New CES Nesting Structure for Consumption

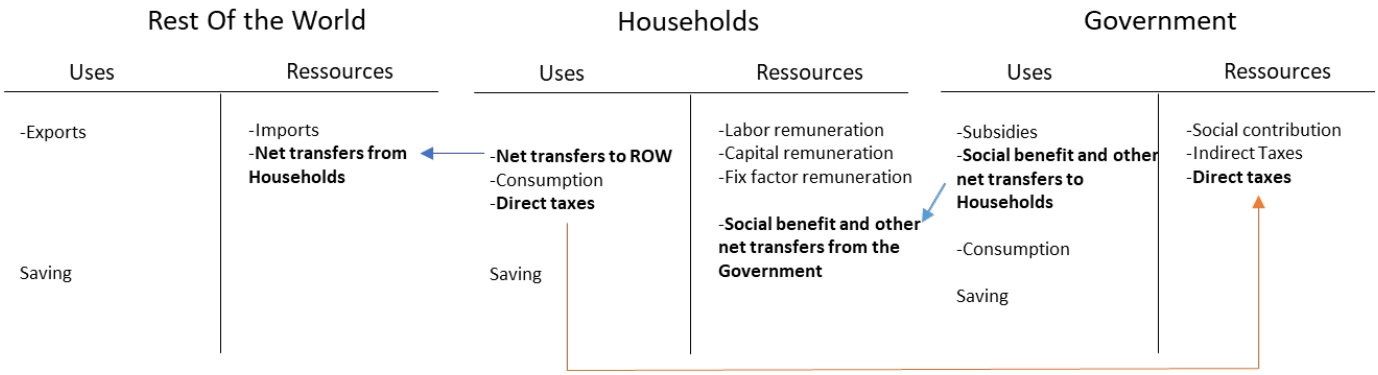



**Figure 6.** Representation of the Economic Sectors in GEMINI-E3-EU

### 3.2.2 Including household heterogeneity in GEMINI-E3-EU

Recognising the importance of equity in climate and energy transitions, GEMINI-E3-EU incorporates household heterogeneity by arranging the population into income-based groups or representative households (Figure 7). Each group is characterised by distinct income levels, consumption preferences, expenditure patterns, and elasticities of substitution. The model thereby captures differential responses to changes in relative prices, tax structures, and environmental policy instruments.

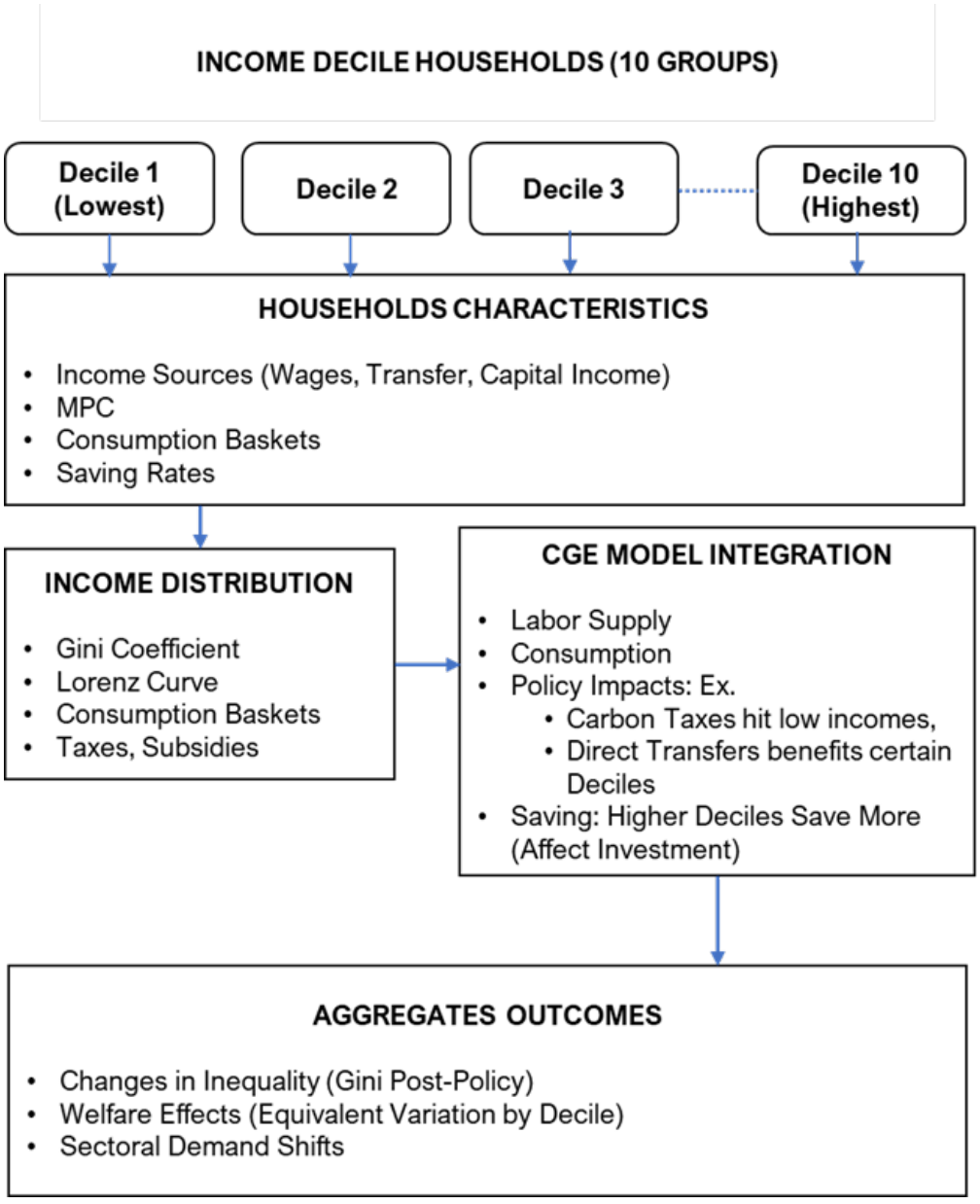



**Figure 7.** Households' Heterogeneity in GEMINI-E3-EU

In the new version of the GEMINI-E3-EU model, the incorporation of household heterogeneity is a key focus, acknowledging that different households experience economic shocks and climate impacts differently. This variation significantly influences overall outcomes, such as consumption, investment, and welfare. By integrating household diversity, the model aims to develop targeted adaptation and mitigation strategies that better respond to the unique circumstances of various household groups.

The model emphasises the EU's diverse economic and social landscapes, leveraging high-quality datasets, in turn allowing for a nuanced analysis of policy impacts across different income deciles, reflecting the economic stratification within the society.

### 3.2.3 Reconciling the GEMINI-E3-EU Input database with the DIAMOND EU household database

#### 3.2.3.1 Mapping household consumption and income

Prior to estimating disaggregation coefficients for the income quintiles, we conducted a comprehensive mapping of household consumption categories and income items between the GEMINI-E3-EU model and the DIAMOND EU household database (for details see section 2.2). This crucial step ensured alignment and consistency between the two datasets, allowing for accurate disaggregation of macroeconomic aggregates into household-level data. The mapping process involved creating detailed correspondence tables that linked the aggregate macroeconomic indicators in GEMINI-E3-EU with the more granular household expenditure categories and income items in the DIAMOND EU household database, facilitating a seamless integration. Tables 3 and 4 below detail the map.

**Table 3.** Mapping Household Consumption Categories into GEMINI-E3-EU Classification

| DIAMOND EU household database | EXPLANATION | GEMINI-E3-EU CLASSIFICATION |
|---|---|---|
| EUR_HE00 | Total consumption expenditure | Total Consumption |
| EUR_HE011 | Food | S07 |
| EUR_HE012 | Non-alcoholic beverages | S07 |
| EUR_HE021 | Alcoholic beverages | S07 |
| EUR_HE022 | Tobacco | S07 |
| EUR_HE023 | Narcotics | S07 |
| EUR_HE031 | Clothing | S13 |
| EUR_HE032 | Footwear | S13 |
| EUR_HE041 | Actual rentals for housing | S13 |
| EUR_HE042 | Imputed rentals for housing | S13 |
| EUR_HE043 | Maintenance and repair of the dwelling | S13 |
| EUR_HE044 | Water supply and miscellaneous services relating to the dwelling | S13 |
| EUR_HE04510 | Electricity | S05 |
| EUR_HE04521 | Natural gas and town gas | S03 |
| EUR_HE04522 | Liquefied hydrocarbons-butane-propane-etc. | S04 |
| EUR_HE04530 | Liquid fuels | S04 |
| EUR_HE04541 | Coal | S01 |
| EUR_HE04549 | Other solid fuels | S01 |
| EUR_HE04550 | Heat energy | S05 |
| EUR_HE051 | Furniture and furnishings, carpets and other floor coverings | S09 |
| EUR_HE052 | Household textiles | S09 |
| EUR_HE053 | Household appliances | S09 |

| DIAMOND EU household database | EXPLANATION | GEMINI-E3-EU CLASSIFICATION |
|---|---|---|
| EUR_HE054 | Glassware tableware and household utensils | S09 |
| EUR_HE055 | Tools and equipment for house and garden | S09 |
| EUR_HE056 | Goods and services for routine household maintenance | S13 |
| EUR_HE061 | Medical products, appliances and equipment | S08 |
| EUR_HE062 | Out-patient services | S13 |
| EUR_HE063 | Hospital services | S13 |
| EUR_HE071 | Purchase of vehicles | S09 |
| EUR_HE0721 | Spare parts and accessories for personal transport equipment | S08 |
| EUR_HE07221 | Diesel | S04 |
| EUR_HE07222 | Petrol | S04 |
| EUR_HE07223 | Other fuels for personal transport equipment | S04 |
| EUR_HE07224 | Lubricants | S04 |
| EUR_HE0723 | Maintenance and repair of personal transport equipment | S13 |
| EUR_HE0724 | Other services in respect of personal transport equipment | S13 |
| EUR_HE07311 | Passenger transport by train | S10 |
| EUR_HE07312 | Passenger transport by underground and tram | S10 |
| EUR_HE07321 | Passenger transport by bus and coach | S10 |
| EUR_HE07322 | Passenger transport by taxi and hired car with driver | S10 |
| EUR_HE07331 | Domestic flights | S12 |
| EUR_HE07332 | International flights | S12 |
| EUR_HE07341 | Passenger transport by sea | S11 |
| EUR_HE07342 | Passenger transport by inland waterway | S11 |
| EUR_HE07350 | Combined passenger transport | S10 |
| EUR_HE07361 | Funicular, cable-car and chair-lift transport | S10 |
| EUR_HE07362 | Removal and storage services | S13 |
| EUR_HE07369 | Other purchased transport services NEC | S13 |
| EUR_HE081 | Postal services | S13 |
| EUR_HE082 | Telephone and telefax equipment | S13 |
| EUR_HE083 | Telephone and telefax services | S13 |
| EUR_HE091 | Audio-visual, photographic and information processing equipment | S13 |
| EUR_HE092 | Other major durables for recreation and culture | S13 |
| EUR_HE093 | Other recreational items and equipment, gardens and pets | S13 |
| EUR_HE094 | Recreational and cultural services | S13 |
| EUR_HE095 | Newspapers, books and stationery | S13 |
| EUR_HE096 | Package holidays | S13 |
| EUR_HE101 | Pre-primary and primary education | S13 |
| EUR_HE102 | Secondary education | S13 |
| EUR_HE103 | Post-secondary nontertiary education | S13 |
| EUR_HE104 | Tertiary education | S13 |
| EUR_HE105 | Education not definable by level | S13 |
| EUR_HE111 | Catering services | S13 |
| EUR_HE112 | Accommodation services | S13 |
| EUR_HE121 | Personal care | S13 |
| EUR_HE122 | Prostitution | S13 |
| EUR_HE123 | Personal effects NEC | S13 |

Funded by the European Union

| DIAMOND EU household database | EXPLANATION | GEMINI-E3-EU CLASSIFICATION |
|---|---|---|
| EUR_HE124 | Social protection | S13 |
| EUR_HE125 | Insurance | S13 |
| EUR_HE126 | Financial services NEC | S13 |
| EUR_HE127 | Other services NEC | S13 |

**Table 4.** Mapping Households' Income Items into GEMINI-E3-EU Classification

| DIAMOND EU household database | EXPLANATION | GEMINI-E3-EU CLASSIFICATION |
|---|---|---|
| HY010 | Total household gross income | |
| HY020 | Total disposable household income | Disposable Income |
| PY010G_hh | Employee cash or near cash income | Income from Labour Remuneration |
| PY020G_hh | Non-cash employee income | |
| PY021G_hh | Company car | |
| PY050G_hh | Cash profits or losses from self-employment | Income from Labour Remuneration |
| PY080G_hh | Pension from individual private plans | |
| PY090G_hh | Unemployment benefits | Transfers from government to Households |
| PY100G_hh | Old-age benefits | |
| PY110G_hh | Survivor benefits | |
| PY120G_hh | Sickness benefits | Transfers from government to Households |
| PY130G_hh | Disability benefits | Transfers from government to Households |
| PY140G_hh | Education-related allowances | Transfers from government to Households |
| HY030G | Imputed rent | |
| HY040G | Income from rental of a property or land | Income from Capital Remuneration |
| HY050G | Family/children-related allowances | |
| HY060G | Social exclusion not elsewhere classified | Transfers from government to Households |
| HY070G | Housing allowances | Transfers from government to Households |
| HY080G | Regular interhousehold cash transfer received | Transfer Between Households |
| HY090G | Interest, dividends, profit from capital investments in unincorporated business | Income from Capital Remuneration |
| HY100G | Interest repayments on mortgage | |
| HY110G | Income received by people aged under 16 years | |
| HY120G | Regular taxes on wealth | Direct Taxes |
| HY130G | Regular interhousehold cash transfer paid | Transfer Between Households |
| HY140G | Tax on income and social contributions | Direct Taxes |
| HY170G | Value of goods produced for own consumption | |

Funded by the European Union

### 3.2.3.2 Aligning saving rates by decile with Eurostat data

To reconcile the DIAMOND EU household database and the GEMINI-E3-EU database, we follow a process aiming to align saving rates by quintile with Eurostat data while maintaining the aggregate household saving rate from the GEMINI-E3-EU database. Our approach involves replicating the profile of differences between quintile-level saving rates and the aggregate saving rate observed in Eurostat data. To achieve this, we use government-to-household transfers at the decile level as a control variable and solve an optimisation problem to ensure consistency across datasets, by following these equations:

$$\min_{Transfer_G_H_D} Objective$$

$$Objective = \sum_{Q} \left[ \left( Saving\ rate_Q - \overline{Saving\ rate} \right) - \left( Saving\ rate_Q^{Eurostat} - \overline{Saving\ rate^{Eurostat}} \right) \right]^2$$

$$Saving_D = Disposable\ income_D - \widehat{Consumption_D} \quad (1)$$

$$Disposable\ income_D = \widehat{Disposable\ income_D} - \widehat{Transfer_G_H_D} + Transfer_G_H_D \quad (2)$$

$$Saving\ rate_Q = \frac{Saving_D + Saving_{D+1}}{Disposable\ income_D + Disposable\ income_{D+1}},\ D\ and\ D+1 \in Q \quad (3)$$

$$\overline{Saving\ rate} = \frac{\sum_D Saving_D}{\sum_D Disposable\ income_D} \quad (4)$$

$$Saving\ rate_{D+1} \geq Saving\ rate_D \quad (5)$$

$$\sum_D Transfer_G_H_D = Transfer_G_H^{Gemini} \quad (6)$$

This optimisation method improves alignment of regional aggregate saving rates in GEMINI-E3-EU with Eurostat at the quantile level and with the DIAMOND EU household database at the decile level, as illustrated in Figure 8 for France and Spain.

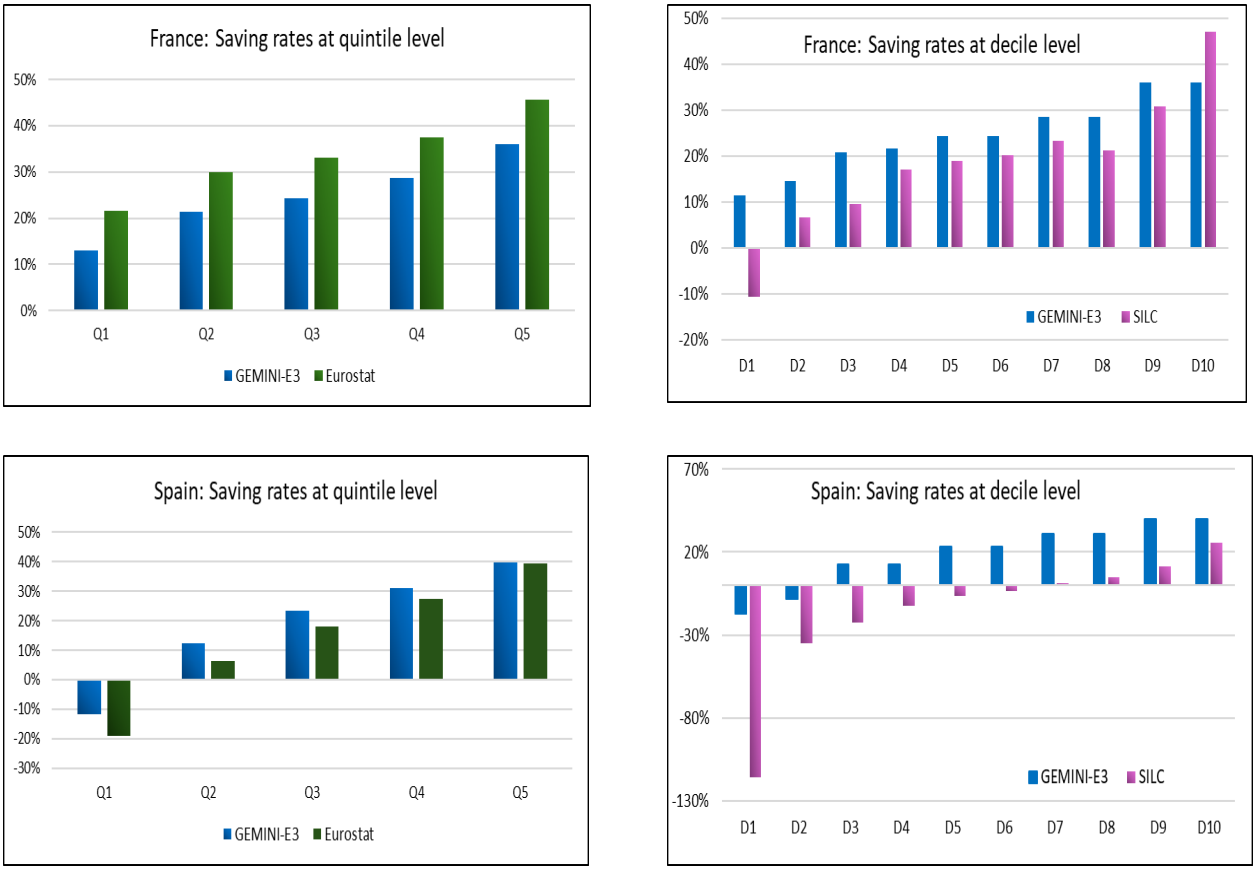



**Figure 8.** Aligning Saving Rates of France and Spain

### 3.2.4 Running preliminary scenarios

#### 3.2.4.1 Scenario design

During the previous period, a beta version of the GEMINI-E3-EU model that includes household heterogeneity has been built, aiming to test the equations as well as to check the solving property of the model. For this purpose, we use an initial version of the DIAMOND EU Households database early on in the process to develop a scenario in which a carbon tax is gradually introduced in the European Union and reaches US$ 100 per ton of $CO_2$ in 2050. Although not a final scenario exercise, this test run serves to validate the usefulness of the approach used in the project to improve household heterogeneity in IAMs.

In the standard version of GEMINI-E3-EU, where different household types are not distinguished, the government budget is balanced based on the government's deficit or surplus as determined by the reference scenario. The balancing variable used is direct taxes paid by households ($FGRS_DT_H(reg)$), which follows a lump-sum transfer rule. Households' direct tax payments to the government are calculated based on their revenues, which serve as the tax base multiplied by the tax rate ($TXFGRS_H(reg)$):

$$FGRS_DT_H(reg) = TXFGRS_H(reg) \times (FSEFA_H_K(reg) + FSEFA_H_L(reg) + FSEFA_H_F(reg) + FSESE_H_G(reg))$$

When distinguishing between different household types, a scenario should be designed that adapts redistribution rules. This is crucial in determining the effects of fiscal policies at the decile level. The direct tax calculation at the decile level is now given by:

$$FGRS_DT_H(d,reg) = TXFGRS_H(d,reg) \times (A_FSEFA_H_K(d,reg) \times FSEFA_H_K(reg) + A_FSEFA_H_L(d,reg) \times FSEFA_H_L(reg) + A_FSEFA_H_F(d,reg) \times FSEFA_H_F(reg) + FSESE_H_G(d,reg))$$

We have tested two redistribution rules under this EU carbon tax scenario:

- **Proportional Adjustment:** The tax rate is modified by the same multiplier factor $(K_TXFGRS_H(d,reg))$ for each decile, such that: $K_TXFGRS_H(d,reg) \times TXFGRS_H(d,reg)$, called "**rule 1**",
- **Fixed Absolute Adjustment:** A same fixed absolute amount ($A_TXFGRS_H(d,reg)$), which can be positive or negative, is added at each decile level, called "**rule 2**". The equation then becomes:

$$FGRS_DT_H(d,reg) = TXFGRS_H(d,reg) \times (A_FSEFA_H_K(d,reg) \times FSEFA_H_K(reg) + A_FSEFA_H_L(d,reg) \times FSEFA_H_L(reg) + A_FSEFA_H_F(d,reg) + FSESE_H_G(d,reg)) + A_TXFGRS_H(d,reg)$$

These approaches allow for greater flexibility in modeling tax redistribution and assessing the impact of fiscal policies at different income levels.

### 3.2.4.2 Presentation of results

To scrutinise the results, we use France as a test case. Figure 9 shows the impact of the $CO_2$ tax on $CO_2$ emissions from households by decile in 2050. In percentage terms, the variations in $CO_2$ emissions are comparable by decile, ranging from 7.3% to 9.1%. There is no significant difference according to the redistribution rule (i.e., rule 1 versus rule 2). In absolute terms, the results are more mixed. Low-income households have lower emissions reductions than wealthy households, due to the higher emissions of the latter. Thus, the reduction in emissions for decile 1 is 0.5 Mt of $CO_2$, while that for decile 10 is 0.8 Mt of $CO_2$ in the case of rule 1.

We calculate the impact of both scenarios on welfare by decile. Both scenarios are regressive, meaning that poor households are more affected than rich households. However, rule 1 accentuates regressivity, while rule 2 tends to reduce the gap between poor and rich households. However, neither of these two rules is satisfactory from the point of view of fairness, so other rules will have to be considered.

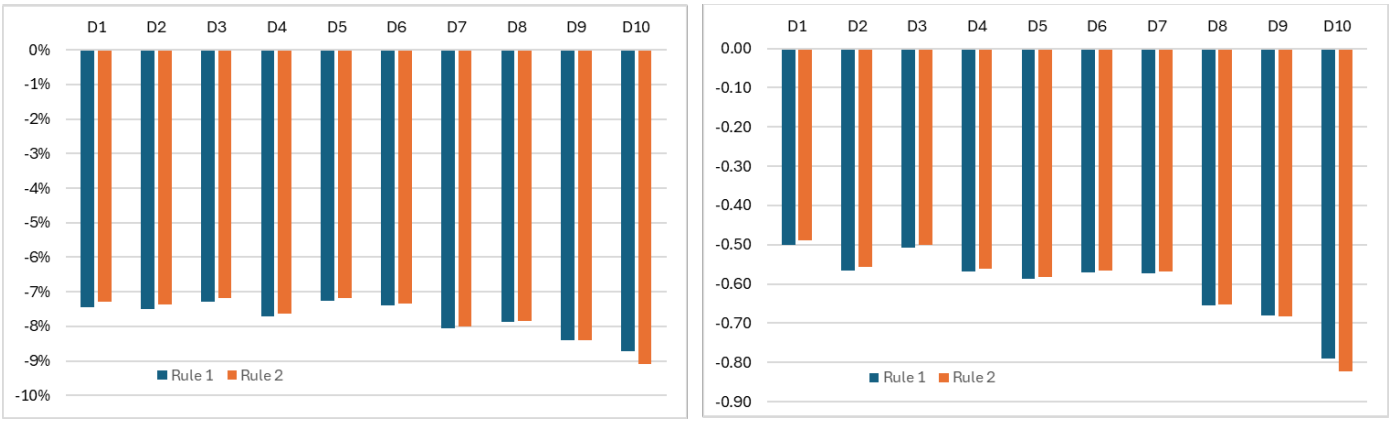



**Figure 9.** $CO_2$ Emissions Change from Households (in percentage (left) and Mt $CO_2$ (right))

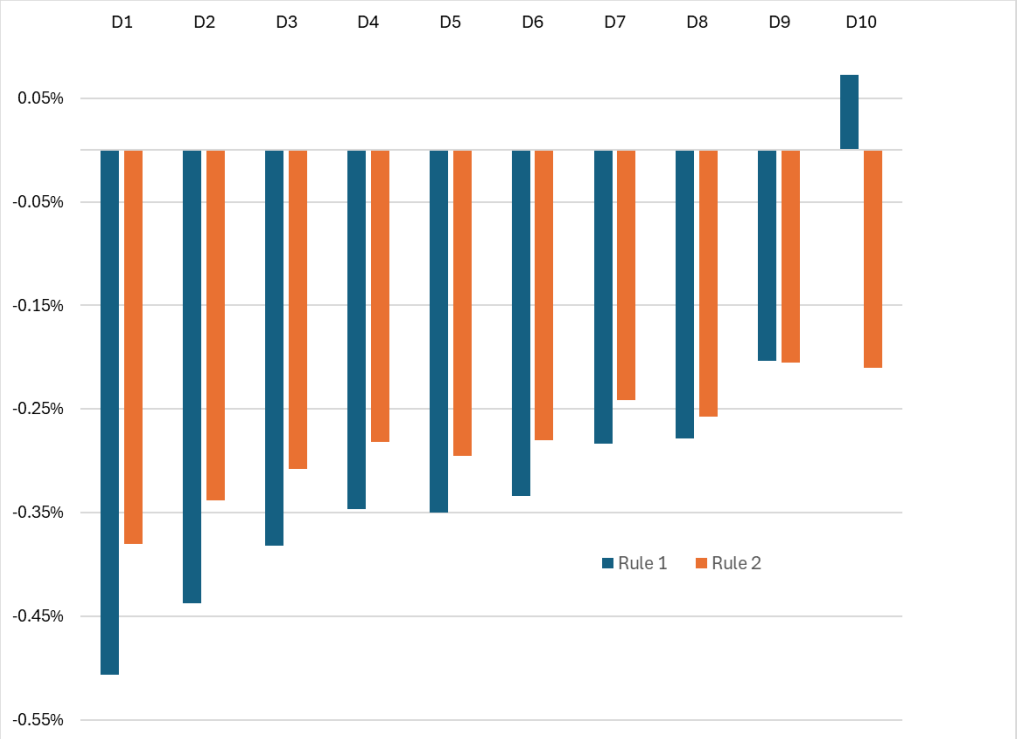



**Figure 10.** Welfare Change of Household Consumption (in percentage)

## 3.3 Integration of the household dataset in NEMESIS-World

### 3.3.1 State of the art and objectives

The NEMESIS model presented here (EU version)[6] is developed, managed, and enhanced by SEURECO. The model is designed to capture sectoral interdependencies and long-term macroeconomic dynamics across all EU Member States (with the World version expanding the regional granularity). It combines a detailed multi-sectoral structure, covering 30 industries, with behavioural equations for key economic agents, including households, firms, and the government. The model incorporates dynamic adjustment processes, allowing for realistic short- and long-term policy simulations. Among its various components, the household module plays a central role in capturing consumption and income dynamics. In the current version, households are represented by a single representative household. The current development phase aims to overcome this assumption by introducing heterogeneity based on income. The consumption decisions of the representative household depend on disposable income, relative prices, and other socioeconomic factors.

Total household consumption is determined as a function of real disposable income, interest rate, unemployment rate (used as a proxy for economic uncertainty), and population structure[7]. The model adopts a permanent income hypothesis, meaning that actual consumption progressively adjusts to a "desired" level. A long-term equation is estimated econometrically – via a cointegration function (Engle & Granger, 1987) – complemented by a short-term equation to capture short-run adjustments. Total household consumption is then allocated across different

[6] The NEMESIS-World model is following a similar structure with the EU version. Since the database established in this report is focused on the EU, integration of the household heterogeneity database in the NEMESIS model implies the use of the EU version, with enhancements made inherently being part of the "World" version as well.

[7] See https://www.erasme-team.eu/en/the-nemesis-model/ for details.

consumption purposes, according to COICOP, such as food, transport, and health. This distribution is modelled using a modified version of the Almost Ideal Demand System (AIDS), which models the budget share of each function (Bracke & Meyermans, 1997). These budget shares depend on the household income level (income effect), relative prices between categories (substitution effect), as well as structural trends (preferences, demographic changes, etc.). Income and price elasticities are used to differentiate behaviour across types of goods (e.g., normal, inferior, or luxury goods). Finally, consumption by purpose (category) is converted into product demand to align with the classification of economic activities on the production side of the model.

Figure 11 provides an overview of the structure of the modelling of households' income and consumption in NEMESIS and the interactions between income and expenditure.

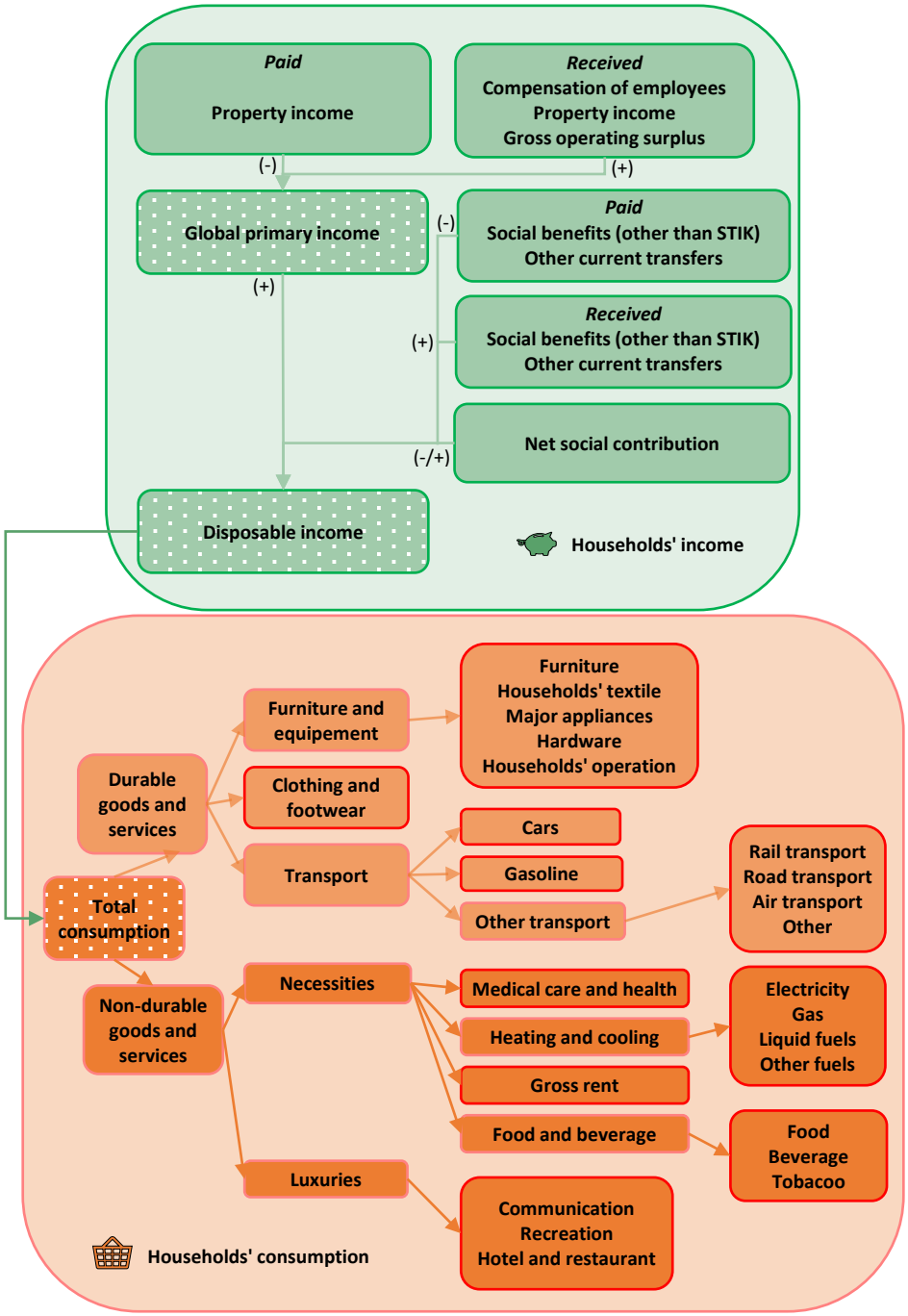



**Figure 11.** Modelling Framework for the Household Sector within the NEMESIS model

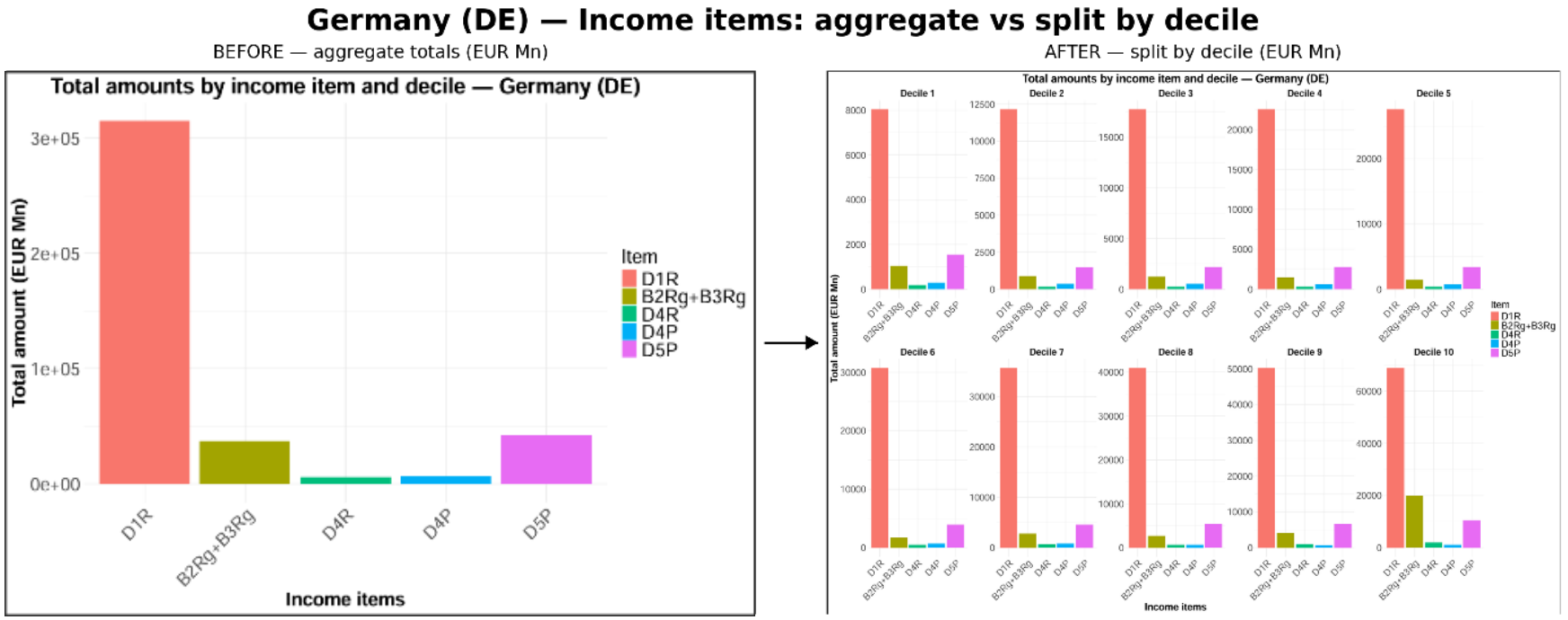



**Figure 12.** Illustration of Households' Representation in Germany before and after the Modelling Improvements for the calibration year

The different income items correspond to the National Accounts Indicators (Eurostat) for the components of household income. D1R: Compensation of employees, B2Rg+B3Rg: Gross operating surplus and mixed income, D4R and D4P: property income received and paid respectively, D5P: Taxes on income and wealth.

The objective of this work is the improvement and the expansion of the household sector to account for heterogeneity. In practice, this heterogeneity may involve both economic factors (mainly, income) and sociological dimensions (e.g., location, gender, age, household composition). Based on the DIAMOND EU household database (see details in section 2.2), we can capture income-based heterogeneity in household income and consumption by shifting from a single representative household to ten distinct income groups (or deciles), each with its own set of income and consumption functions. Accordingly, all equations within the household module must be disaggregated into ten versions to represent these ten household groups.

The model's nomenclature allows for incorporation of additional heterogeneity dimensions, once relevant data become available.

### 3.3.2 Data overview

The DIAMOND EU household database provides detailed household data at the decile level, which offer a key resource towards improving the household representation within the NEMESIS model. Building this dataset is a complex process, requiring the collection and matching of data from various sources (see Section 2). This process involves methodological choices that, while robust, introduce a degree of uncertainty. To address this uncertainty, we replicated our work using an alternative database produced by Eurostat (2024a), which follows a different data processing methodology. This allows for an assessment of the uncertainty, and it ensures the findings are more reliable and transparent.

As part of the European project on income, consumption, and wealth inequalities (ICW), Eurostat conducted a centralised methodological exercise to produce household income and consumption distributions aligned with national accounts (Eurostat, *Distribution of income and consumption for the Household sector*, 2024). This work combines macroeconomic data from sectoral accounts with microdata from EU-SILC and HBS surveys. Since survey data are often incomplete or only partially representative, various adjustment methods were applied to ensure consistency with the national accounts. The first method, referred to as the proportional allocation, distributes the gap uniformly across all households, under the assumption that sampling errors or underreporting affect the entire population in a similar manner. The second method, known as the Pareto approach, specifically adjusts the upper tail of the distribution by imposing a Pareto law beyond the 90th percentile, in order to better capture very high incomes that are often underreported; the remaining gap is then allocated proportionally. The third method relies on decile-based reallocation, assuming either that the gap originates primarily from high-income households ("to the top") or, alternatively, from low-income households ("to the bottom"). Finally, the combined approach applies different methods for different variables: for instance, the Pareto adjustment for capital income, ascending or descending allocation for certain transfers or taxes, and proportional allocation for consumption.

The resulting database provides information on both household income and expenditure by decile, similar to the DIAMOND EU household database. It includes data on wages and salaries (D1R), gross operating surplus (B2Rg), and mixed income (B3Rg). It also covers property income, both received and paid (D4R, D4P), taxes on income and wealth (D5P), as well as social benefits other than social transfers in kind, received and paid (D62R, D62P). In addition, it contains information on net social contributions (D61N) and other current transfers, received and paid (D7R, D7P). Consumption is broken down using the standard COICOP 1999 classification (12 categories). Households are divided into income deciles. As with the database of DIAMOND, other heterogeneity factors are not currently accounted for.

To address missing values, we complemented the distributional national accounts by Eurostat with alternative sources, either from Eurostat itself or from national statistical institutes (e.g., the Office for National Statistics for the United Kingdom). In such cases, only minor adjustments were required, such as aligning items with Eurostat's

classification and harmonising measurement units. When data were not available by household deciles, we applied an allocation procedure based on the distribution of a neighbouring country. For example, French data, available only by quintiles, were disaggregated into deciles using Belgian data, thereby avoiding a simple equal split and allowing for the attribution of differentiated weights to each decile.

**Table 5:** Summary of Data Reconstruction References

| Data reconstruction with alternatives sources | | Data reconstruction with neighbouring approach | |
|---|---|---|---|
| **Country** | **Source** | **Country** | **Neighbour** |
| France (income, consumption data) | Eurostat (2024b) | France (income, consumption data) | Belgium |
| Czechia (income, consumption data) | Eurostat (2024b) | Czechia (income, consumption data) | Slovakia |
| United Kingdom (income, consumption data) | ONS (2023, 2024) | Iceland (consumption data) | Denmark |
| Iceland (consumption data) | Statistics Iceland (2023) | Switzerland (consumption data) | Germany |
| Switzerland (consumption data) | OFS (2024) | Luxembourg (consumption data) | Germany |
| Luxembourg (consumption data) | Eurostat (2025e) | Norway (consumption data) | Sweden |
| Norway (consumption data) | Eurostat (2025e) | | |

### Key differences between both datasets

The DIAMOND EU household database and the distributional national accounts by Eurostat differ in several aspects:

- Methodological approach and data sources,
- Data granularity and precision,
- Coverage of countries and consumption categories and income items,
- Level of alignment with the NEMESIS-World consumption categories and income items.

These differences are detailed in Table 6, supporting our idea of using both datasets to address robustness and sensitivity of model outcomes, while the more detailed granularity of the DIAMOND EU household database presented further underpins the novelty of the dataset created.

Funded by the European Union

**Table 6:** Comparison of distributional national accounts by Eurostat and the DIAMOND EU household database

| | EUROSTAT | | DIAMOND | |
|---|---|---|---|---|
| **Methodology and sources** | Merging information from EU-SILC and national accounts with consolidation methods (Eurostat, 2024a) | | Merging information from EU-SILC (recipient) and HBS (donor) using a statistical matching strategy | |
| **Countries not covered (before data processing)** | **Consumption data** | **Income data** | **Consumption data** | **Income data** |
| | FR, CZ, LU, IS, UK, CH, NO | FR, CZ, IS, UK | AT, NO, CH, UK | AT, NO, CH, UK |
| **Individual variables not covered for some countries (before data processing)** | **Eurostat variable code** | **Country** | **EU-SILC variable code** | **Country** |
| | D62P - Social benefits, other than STiK (paid) | AT, BG, MT, EE, FR, UK | EUR_HE071 (Car) | BG (decile 1), RO (decile 7) |
| | D62R - Social benefits, other than STiK (received) | MT, FR | EUR_HE07221 (Diesel) | DE |
| | B5g - Balance of primary incomes, gross | MT | EUR_HE0722 (Other fuels for personal transport) | CY, DE, DK, ES, FI, FR, LU, SE |
| | B6g - Disposable income, gross | MT | EUR_HE07222 (Petrol) | DE, ES, FR, SE |
| | D4P - Property income (paid) | MT, UK | EUR_HE0731 (Passenger transport by train)<br>EUR_HE07312 (Passenger transport by underground and tram) | CY, DE, DK, LT (decile1), MT |
| | D4R - Property income (received) | MT | EUR_HE07321 (Passenger transport by bus and coach)<br>EUR_HE07322 (Passenger transport by taxi and hired car with driver) | DE |
| | D5P - Current taxes on income and wealth | MT | EUR_HE07331 (Domestic flights)<br>EUR_HE07332 (International flights) | HU, BG (decile 5), PL (deciles 2, 7), RO (decile 4); SE (deciles 1, 4) |
| | B2Rg+B3Rg - Operating surplus and mixed income, gross | CH, UK | EUR_HE07341 (Passenger transport by sea)<br>EUR_HE07342 (Passenger transport by inland waterway)<br>EUR_HE07350 (Combined passenger transport)<br>EUR_HE07361 (Funicular, cable-car and chair-lift transport)<br>EUR_HE07362 (Removal and storage services)<br>EUR_HE07369 (Other purchased transport services NEC) | DE, LT |
| | D61N - Net social contribution | CZ, BE, EE, FR, UK | EUR_HE04521 (Natural gas and town gas)<br>EUR_HE04522 (Liquefied hydrocarbons-butane-propane, etc) | SE |
| | D7P - Other current transfers (paid) | FR, MT | EUR_HE04530 (Liquid fuels) | EE (decile 2), LT, HU, MT, NL (decile1), RO (decile1) |

| | EUROSTAT | | DIAMOND | |
|---|---|---|---|---|
| | D7R - Other current transfers (received) | FR, MT | EUR_HE04541 (Coal) | DE, DK, GR, FI, IT (deciles 3, 6), LV, MT, NL, SE |
| | | | EUR_HE04549 (Other solid fuels) | DE, ES, FR, LU, LV, SE |
| | | | EUR_HE0455 (Heat energy) | BE, CY, ES, IE, LU, MT, NL |
| | | | HY100G - Interest repayments on mortgage | LU |
| **Data granularity** | 12 consumption categories from Eurostat vs 27 in NEMESIS | | 71 consumption categories from the database of DIAMOND vs 27 in NEMESIS (allowing a disaggregation of energy) | |
| | Eurostat income variables vs 15 in NEMESIS | | 26 income variable the database of DIAMOND vs 15 in NEMESIS | |

Figures 13 and 14 show the share of each household decile for each income item and consumption category, based on distributional national accounts by Eurostat (left) and the dataset of DIAMOND (right). In these figures, Belgium, France, Germany, and Spain are used as examples to illustrate the differences across decile depending on the data source considered. For results covering all countries, see Figure 15 and Figure 16 in the Appendix.

We can observe that the distribution across deciles is relatively similar between the two datasets, particularly for the middle deciles. The different methodologies in data processing and adjusting the data to national accounts totals, though, lead to pronounced differences at the extreme values of income distribution, i.e., for deciles 1 and 10. For example, in the case of property income received (D4R), the share represented by decile 10 is generally larger in the Eurostat version of the data. Similarly, *social benefits other than social transfers in kind (STIKs) paid* (D62P) is often missing from the income side in the database of DIAMOND, and the same applies to the *liquid fuels* category on the consumption side.

Funded by
the European Union

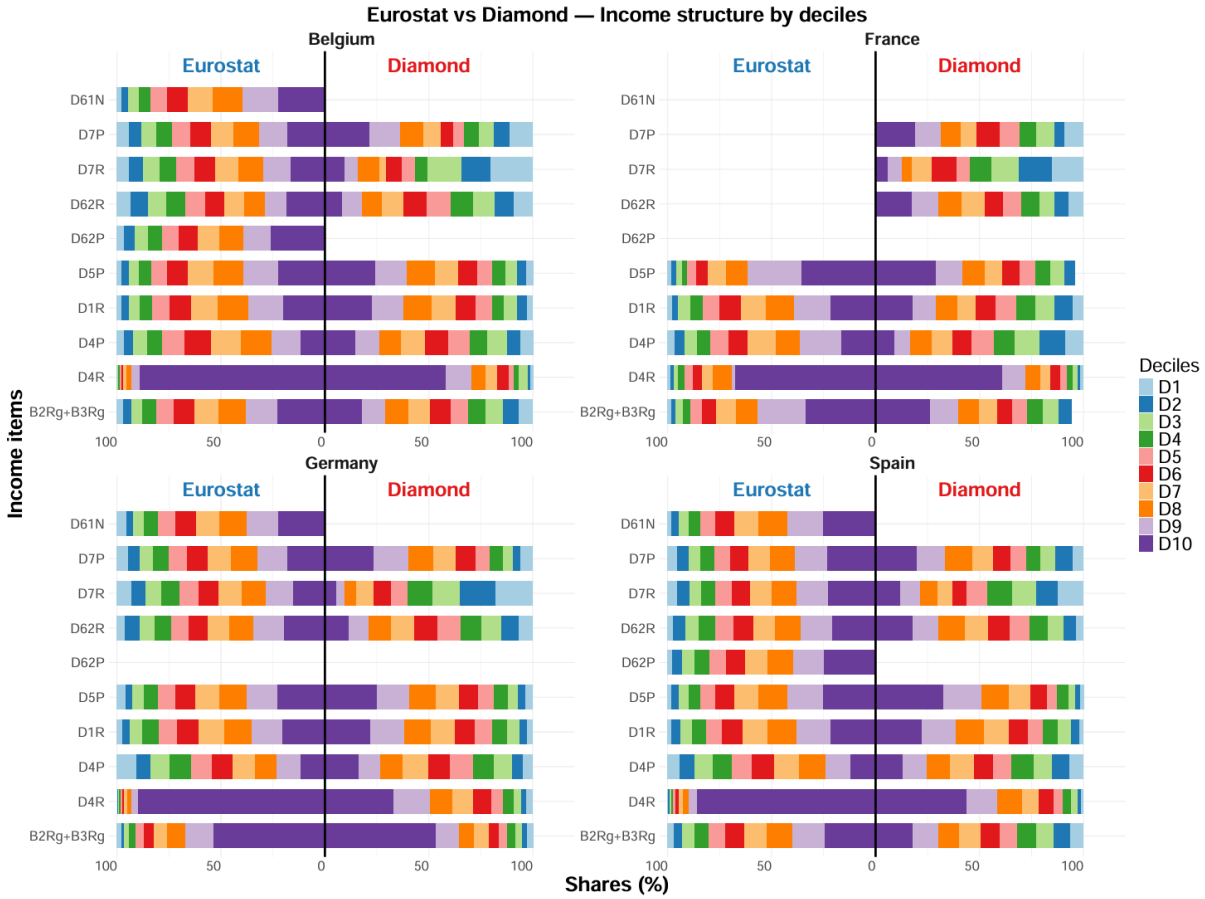



**Figure 13.** Eurostat vs. DIAMOND: Distribution of Household Income Items across Deciles (Belgium, France, Germany, and Spain)

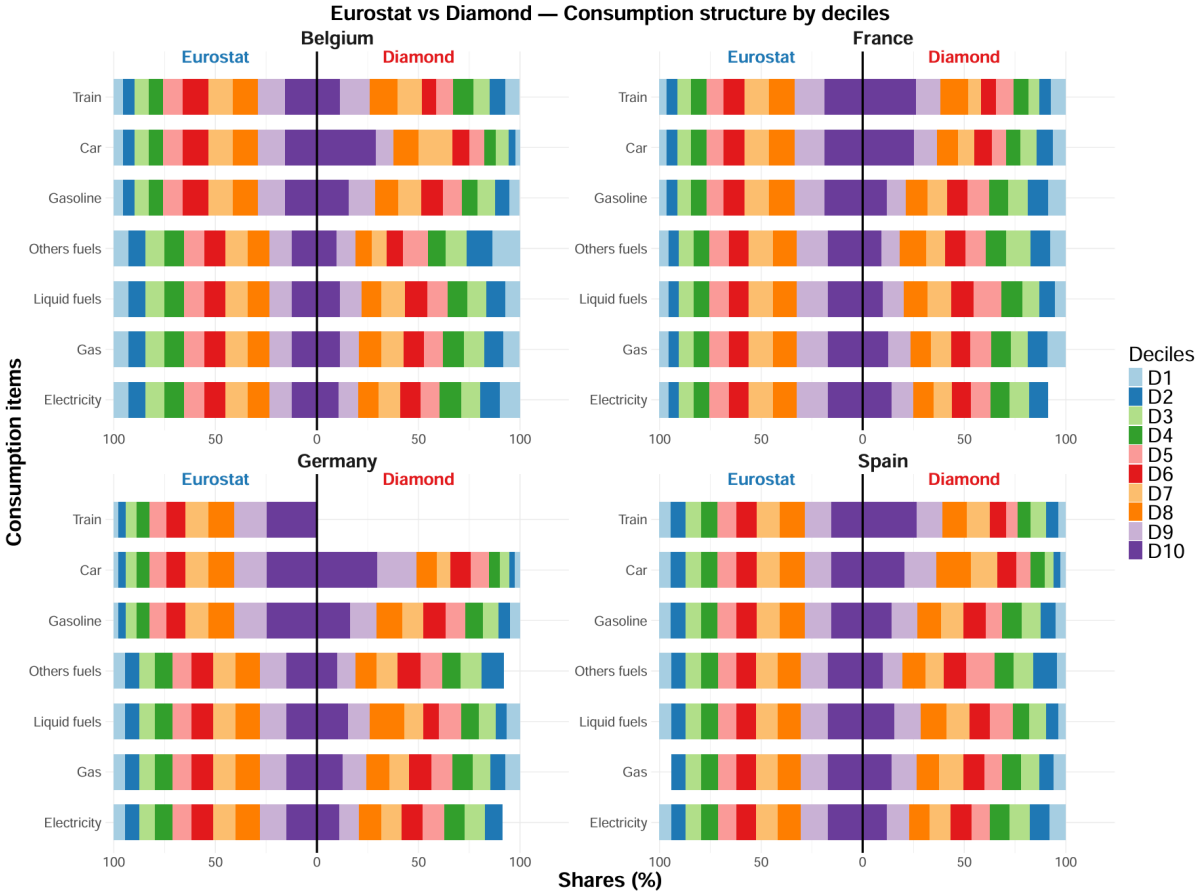



**Figure 14.** Eurostat vs. DIAMOND: Distribution of Household Consumption Categories across Deciles (Belgium, France, Germany, and Spain)

### 3.3.3 Methodology

Since the income and consumption categories used in the distributional national accounts by Eurostat and the database of DIAMOND do not have the same granularity as in the NEMESIS household module, the first step of the methodology consists in establishing a mapping. In both datasets, certain variables may be more aggregated or, conversely, more disaggregated than those used in the NEMESIS model, thereby including or excluding different components. For example, the variable representing employees' compensation in the NEMESIS model, is split into two distinct categories in the DIAMOND database: *Employee cash or near-cash income* and *Non-cash employee income*.

The mapping between both datasets and NEMESIS categories is presented in Table 7 for income and Table 8 for consumption purposes.

**Table 7:** Mapping between Income Items of Distributional National Accounts by Eurostat or the DIAMOND EU household database and their Equivalent in NEMESIS

| NEMESIS income items | Eurostat income items corresponding | DIAMOND income items corresponding |
|---|---|---|
| **Gross operating surplus** | B2Rg+B3Rg (Operating surplus and mixed income, gross) | PY050G_hh (Cash profits or losses from self-employment) + HY030G (Imputed rent) + HY040G (Income from rental of a property or land) + HY170G (Value of goods produced for own consumption) |
| **Income property, received** | D4R (Property income, received) | HY090G (Interest, dividends, profit from capital investments in unincorporated business) |
| **Income property paid** | D4R (Property income, paid) | HY100G (Interest repayments on mortgage) |
| **Compensation of employees, received** | D1R (Compensation of employees) | PY010G_hh (Employee cash or near cash income) + PY020G_hh (Non-cash employee income) |
| **Direct taxes, paid** | D5P (Current taxes on income and wealth) | HY120G (Regular taxes on wealth) + α*HY140G (Tax on income and social contributions) |
| **Other social benefits, received** | D62P (Social benefits, other than STiK, received) | PY090G_hh (Unemployment benefits) + PY100G_hh (Old-age benefits) + PY110G_hh (Survivor benefits) + PY120G_hh (Sickness benefits) + PY130G_hh (Disability benefits) + PY140G_hh (Education-related allowances) + HY050G (Family/children-related allowances) + HY060G (Social exclusion not elsewhere classified) + HY070G (Housing allowances) |
| **Other social benefits, paid** | D62P (Social benefits, other than STiK, paid) | NA |
| **Current transfers, received** | D7R (Other current transfers, received) | HY080G (Regular interhousehold cash transfer received) |
| **Current transfers, paid** | D7R (Other current transfers, paid) | HY130G (Regular interhousehold cash transfer paid) |
| **Net social contribution, received – Net social contribution** | D61N (Net social contributions) | β * HY140G (Tax on income and social contributions) |
| **Disposable income** | B6g (Disposable income, gross) | HY020 (Total disposable household income) |
| **Global primary income** | B5g (Balance of primary incomes, gross) | PY010G_hh (Employee cash or near cash income) + PY020G_hh (Non-cash employee income) + PY050G_hh (Cash profits or losses from self-employment) +HY030G (Imputed rent) + HY040G (Income from rental of a property or land) + HY170G (Value of goods produced for own consumption) + HY090G (Interest, dividends, profit from capital investments in unincorporated business) + HY100G (Interest repayments on mortgage) |

**Table 8:** Mapping between Consumption Categories of Distributional National Accounts by Eurostat and the DIAMOND EU household database and their Equivalent in NEMESIS

| NEMESIS consumption items | Eurostat consumption items corresponding | DIAMOND consumption items corresponding |
|---|---|---|
| **Total consumption** | CP00 (Final domestic consumption expenditure) | EUR_HE00 (Total consumption expenditure) |
| **Food (01)** | α*CP01 (Food and non-alcoholic beverages) | EUR_HE011 (Food) |
| **Beverages (02)** | β*CP01 (Food and non-alcoholic beverages) | EUR_HE012 (Non-alcoholic beverages + EUR_HE021 Alcoholic beverages) |
| **Tobacco (03)** | CP02 (Alcoholic beverages, tobacco and narcotics) | EUR_HE022 (Tobacco) +EUR_HE023 (Narcotics) |
| **Clothing and footwear (04)** | CP03 (Clothing and footwear) | EUR_HE031 (Clothing) + EUR_HE032 (Footwear) |
| **Gross rent (05)** | α*CP04 (Housing, water, electricity, gas and other fuels) | EUR_HE041 (Actual rentals for housing) + EUR_HE042 (Imputed rentals for housing) |
| **Electricity (06)** | β*CP04 (Housing, water, electricity, gas and other fuels) | EUR_HE04510 (Electricity) |
| **Gas 07** | δ*CP04 (Housing, water, electricity, gas and other fuels) | EUR_HE04521 (Natural gas and town gas) + EUR_HE04522 (Liquefied hydrocarbons-butane-propane-etc.) = EUR_HE0452 (Gas) |
| **Liquid fuels (08)** | λ*CP04 (Housing, water, electricity, gas and other fuels) | EUR_HE04530 (Liquid fuels) |
| **Other fuels (09)** | μ*CP04 (Housing, water, electricity, gas and other fuels) | EUR_HE04541 (Coal) + EUR_HE04549 (Other solid fuels) + EUR_HE0455 (Heat energy) = EUR_HE0454 (Solid fuels) + EUR_HE0455 (Heat energy) |

Funded by the European Union

| NEMESIS consumption items | Eurostat consumption items corresponding | DIAMOND consumption items corresponding |
|---|---|---|
| **Furnitures, etc (10)** | α*CP05 (Furnishings, household equipment and routine household maintenance) | EUR_HE051 (Furniture and furnishings, carpets and other floor coverings) |
| **Household textiles (11)** | β*CP05 (Furnishings, household equipment and routine household maintenance) | EUR_HE052 (Household textiles) |
| **Major appliances (12)** | δ*CP05 (Furnishings, household equipment and routine household maintenance) | EUR_HE053 (Household appliances) |
| **Hardware (13)** | λ*CP05 (Furnishings, household equipment and routine household maintenance) | EUR_HE054 (Glassware tableware and household utensils) + EUR_HE055 (Tools and equipment for house and garden) + EUR_HE056 (Goods and services for routine household maintenance) |
| **Household operation (14)** | ν*CP04 (Housing, water, electricity, gas and other fuels) | EUR_HE043 (Maintenance and repair of the dwelling) |
| **Domestic services (15)** | σ*CP04 (Housing, water, electricity, gas and other fuels) | EUR_HE044 (Water supply and miscellaneous services relating to the dwelling) |
| **Medical care and education (16)** | CP06 (Health) | EUR_HE061 (Medical products, appliances, and equipment) + EUR_HE062 (Out-patient services) + EUR_HE063 (Hospital services) + EUR_HE101 (Pre-primary and primary education) + EUR_HE102 (Secondary education) + EUR_HE103 (Post-secondary nontertiary education) + EUR_HE104 (Tertiary education) + EUR_HE105 (Education not definable by level) |
| **Cars (17)** | α*CP07 (Transport) | EUR_HE071 (Purchase of vehicles) |
| **Gasoline (18)** | β*CP07 (Transport) | EUR_HE0721 (Spare parts and accessories for personal transport equipment) + EUR_HE07221 (Diesel) + EUR_HE07222 (Petrol) + EUR_HE07223 (Other fuels for personal transport equipment) + EUR_HE07224 (Lubricants) + EUR_HE0723 (Maintenance and repair of personal transport equipment) + EUR_HE0724 (Other services in respect of personal transport equipment) |
| **Rail transport (19)** | δ*CP07 (Transport) | EUR_HE07311 (Passenger transport by train) + EUR_HE07312 (Passenger transport by underground and tram) = EUR_HE0731 (Passenger transport by railway) |
| **Road transport (20)** | λ*CP07 (Transport) | EUR_HE07321 (Passenger transport by bus and coach) + EUR_HE07322 (Passenger transport by taxi and hired car with driver) = EUR_HE0732 (Passenger transport by road) |
| **Air transport (21)** | μ*CP07 (Transport) | EUR_HE07331 (Domestic flights) + EUR_HE07332 (International flights) = EUR_HE0733 (Passenger transport by air) |
| **Other transports (22)** | ν*CP07 (Transport) | EUR_HE07341 (Passenger transport by sea) + EUR_HE07342 (Passenger transport by inland waterway) + EUR_HE07350 (Combined passenger transport) + EUR_HE07361 (Funicular, cable-car and chair-lift transport) + EUR_HE07362 (Removal and storage services) + EUR_HE07369 (Other purchased transport services NEC) = EUR_HE0734 (Passenger transport by sea and inland waterway) + EUR_HE0735 (Combined passenger transport) + EUR_HE0736 (Other purchased transport services) |
| **Communication (23)** | CP08 (Communications) | EUR_HE081 (Postal services) + EUR_HE082 (Telephone and telefax equipment) + EUR_HE083 (Telephone and telefax services) |
| **Equipment and accessories (24)** | α*CP09 (Recreation and culture) | EUR_HE091 (Audio-visual, photographic and information processing equipment) + EUR_HE092 (Other major durables for recreation and culture) + EUR_HE093 (Other recreational items and equipment, gardens and pets) |
| **Recreation (25)** | β*CP09 (Recreation and culture) | EUR_HE094 (Recreational and cultural services) + EUR_HE095 (Newspapers, books and stationery) + EUR_HE096 (Package holidays) |
| **Hotel and restaurant (26)** | CP11 (Restaurants and hotels) | EUR_HE111 (Catering services) + EUR_HE112 (Accommodation services) |
| **Miscellaneous goods and services (27)** | CP12 (Miscellaneous goods and services) | EUR_HE121 (Personal care) + EUR_HE122 (Prostitution) + EUR_HE123 (Personal effects NEC) + EUR_HE124 (Social protection) + EUR_HE125 (Insurance) + EUR_HE126 (Financial services NEC) + EUR_HE127 (Other services NEC) |

Once the mapping of all categories is complete, we have the income and expenditure data (in million euros) for each NEMESIS income or consumption category, disaggregated by household deciles for both the DIAMOND EU household database and the distributional national accounts by Eurostat. We then compute the share of each

household decile for every income and consumption category. This allows us to observe how each type of income and expenditure is distributed across household deciles for each data source (see Figures 13 and 14). This approach also resolves any inconsistencies between data levels across the datasets and within the NEMESIS model when introducing household income deciles. These distributions are then integrated into the NEMESIS model. We begin with the total (aggregate) values of the income and consumption categories within the model and apply the previously calculated shares. This process allows us to disaggregate each total income and consumption variable – along with its associated equations – into ten variants, each corresponding to a specific household decile.

For example, the wage variable, compensation of employees – named $CMPEMREH$ in NEMESIS – is disaggregated using the shares assigned to each decile according to DIAMOND or Eurostat data:

- **Aggregated $CMPEMREH$ × share of decile 01 (*DIAMOND*) → $CMPEMREH$ for decile 01**
- **…**
- **Aggregated $CMPEMREH$ × share of decile 10 (*DIAMOND*) → $CMPEMREH$ for decile 10**

The same method is applied to each variable and equation in the household module. This results in an important expansion of the model – of about 50 000 equations.

Funded by
the European Union

# 4 Conclusion and future research

This deliverable describes the construction of the DIAMOND EU household database covering detailed microeconomic structures of income and consumption for the EU27. These microstructure data have then been integrated to a different level of detail following tailored approaches into GEMINI-E3-EU, NEMESIS-World (through the NEMESIS EU version) and GCAM-Europe. The integration exercises comprise the adjustment to income and consumption totals from other statistical sources as well as the comparison with published data of income and consumption structures from Eurostat. In the case of GEMINI-E3 the integrated model has also been used for a preliminary model simulation, emphasising the importance of household heterogeneity in energy and climate policy analysis.

The DIAMOND EU household database will still be improved (with respect to filling data gaps), including outstanding work on the integration in the models, but can already been used for analysis of modelling household behaviour.

Funded by
the European Union

# ANNEX




**Figure 15.** Eurostat vs. DIAMOND: Distribution of Household Income Items across Deciles (EU Member States)

Funded by
the European Union




**Figure 16.** Eurostat vs. DIAMOND: Distribution of Household Consumption Categories across Deciles (EU Member States)

# Bibliography

Bond-Lamberty, B., Patel, P., Lurz, J., et al. (2024). *JGCRI/Gcam-Core: GCAM 7.1*. Version gcam-v7.1. Zenodo, released June. https://doi.org/10.5281/zenodo.11481167

Bracke, I., & Meyermans, E. (1997). Specification and estimation of an allocation system for private consumption in Europe. *Federal Planning Bureau Working Paper*, (03-97). https://www.plan.be/sites/default/files/documents/WP_199703_EN.pdf

Budolfson, M., Dennig, F., Errickson, F. Feindt, S., Ferranna, M., Fleurbaey, M., Klenert, D., Kornek, U., Kuruc, K., Méjean, A., Peng, W., Scovronick, N., Spears, D., Wagner, F., & Zuber, S. (2021). Climate action with revenue recycling has benefits for poverty, inequality and well-being. *Nature Climate Change*, *11*(12), 1111–1116. https://doi.org/10.1038/s41558-021-01217-0

Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794. https://doi.org/10.1145/2939672.2939785

Coli, A., & Tartamella, F. (2017). Developing Households' sub-sectors accounts: Pros and cons of the top-down and the bottom-up methods. *Statistical Journal of the IAOS*, 33(2), 525–535. https://doi.org/10.3233/SJI-161027

Emmerling, J., & Tavoni, M. (2021). Representing inequalities in integrated assessment modeling of climate change. *One Earth*, *4*(2), 177–180. https://doi.org/10.1016/j.oneear.2021.01.013

Engle, R. F., & Granger, C. W. J. (1987). Co-Integration and Error Correction: Representation, Estimation, and Testing. *Econometrica 55*(2), 251. https://doi.org/10.2307/1913236

Eurostat (2024a). *Distributional national accounts – national data and metadata*. Experimental statistics. Data. December. https://ec.europa.eu/eurostat/web/experimental-statistics/income-consumption-social-surveys-national-accounts

Eurostat (2024b). *Income and consumption: social surveys and national accounts*. Experimental statistics. Data. Decembre 1. https://ec.europa.eu/eurostat/web/experimental-statistics/income-consumption-social-surveys-national-accounts.

Eurostat (2025a). *National accounts and GDP — Table 6: Consumption expenditure of households, 2005–24* [Data set]. Source: Eurostat (nama_10_gdp) and (nama_10_pc). Retrieved September 1, 2025, from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=National_accounts_and_GDP

Eurostat (2025b). *How to apply for microdata access: Guidelines and application forms*. European Commission. https://ec.europa.eu/eurostat/documents/203647/771732/How_to_apply_for_microdata_access.pdf

Eurostat (2025c). *Complete Energy Balances*. Nrg_bal_c. Data. https://doi.org/10.2908/NRG_BAL_C

Eurostat (2025d) *Disaggregated Final Energy Consumption in Households - Quantities*. Nrg_d_hhq. Data. https://doi.org/10.2908/NRG_D_HHQ

Eurostat (2025e). *Mean consumption expenditure per household by COICOP consumption purpose*. Data. February. https://ec.europa.eu/eurostat/databrowser/view/hbs_exp_t121__custom_15084559/default/table?lang=en

Eurostat (n.d.-a). *European Union Statistics on Income and Living Conditions (EU-SILC)*. Webpage. https://ec.europa.eu/eurostat/web/microdata/european-union-statistics-on-income-and-living-

conditions

Eurostat (n.d.-b). *Household Budget Survey (HBS)*. Webpage. https://ec.europa.eu/eurostat/web/microdata/household-budget-survey

IEA - International Energy Agency (2022). *World Energy Balances 2022* (Database). www.iea.org.

Kratena, K., Tomás, M., Arto, I., Knaus, K., Goritschnig, W., & Cazcarro I. (2021). Household/Consumption sub-module of WILIAM, WP4, Task 4.1, D4.1, Low Carbon Society: *An Enhanced Modelling Tool for the Transition to Sustainability* (LOCOMOTION), H2020-LC-CLA-2018-2, 28th May 2021. https://www.locomotion-h2020.eu/resources/main-project-reports/)

Landis, F., Fredriksson, G., & Rausch, S. (2021). Between- and within-country distributional impacts from harmonizing carbon prices in the EU. *Energy Economics*, *103*, 105585. https://doi.org/10.1016/j.eneco.2021.105585

sOFS (2024). *Dépenses détaillées de l'ensemble des ménages selon l'année*. Data. June 1. https://www.bfs.admin.ch/bfs/en/home/statistics/economic-social-situation-population/income-consumption-wealth/household-budget/household-expenditure.assetdetail.32667039.html

ONS (2023). *The effects of taxes and benefits on household income, disposable income estimate*. Data. June. https://www.ons.gov.uk/peoplepopulationandcommunity/personalandhouseholdfinances/incomeandwealth/datasets/householddisposableincomeandinequality

ONS (2024). *Family spending workbook 2: expenditure by income*. Data. September. https://www.ons.gov.uk/peoplepopulationandcommunity/personalandhouseholdfinances/expenditure/datasets/familyspendingworkbook2expenditurebyincome

Rao, N. D., van Ruijven, B. J., Riahi, K., & Bosetti, V. (2017). Improving poverty and inequality modelling in climate research. *Nature Climate Change*, *7*(12), 857–862. https://doi.org/10.1038/s41558-017-0004-x

Sampedro, J., Waldhoff, St. T., Edmonds, J. A., Iyer, G., Msangi, S., Narayan, K. B., Patel, P., & Wise, M. (2024). Residential Energy Demand, Emissions, and Expenditures at Regional and Income-Decile Level for Alternative Futures. *Environmental Research Letters*, 19(8), 084031.

Sampedro, J. (2025). *GCAM-GRAPHICS-7.0.0*. Version GCAM-GRAPHICS-v7.0.0. Zenodo, released August. https://doi.org/10.5281/zenodo.16960260

Sampedro, J., van de Ven, D.-J., Horowitz, R., Rodés-Bachs, C., Roman, M. V., Tomás, M., Narayan, K., O'Neill, B., Patel, P., Waldhoff, St., & Arto, I. (2025). Improved Living Standards through Reduction of Income Inequality Does Not Jeopardize the Achievement of Existing Climate Pledges. Preprint, *Research Square*, March 19. https://doi.org/10.21203/rs.3.rs-5433539/v1.

Statistics Iceland (2023). *Household final consumption expenditure at constant prices, 1990-2024*. Data. June. https://px.hagstofa.is/pxen/pxweb/en/Efnahagur/Efnahagur__thjodhagsreikningar__einkaneysla/THJ02105.px/

Funded by the European Union