Distributionally extended input–output analysis (DE-IOA): A methodological guide to data compilation with an application to study industry shocks on living standards

Jacobo Ferrer
Universidad Politécnica de Madrid
Contact at: jacobo.ferrer@upm.es

Thursday, April 30, 2026

Abstract

Addressing the call from the G-20 Data Gaps Initiative III.8 recommendation to produce integrated micro-macro distributional data, this paper develops a methodology for compiling distributionally extended input-output (DE-IO) tables that simultaneously integrate the income and expenditure dimensions of the household sector into the interindustry system. Building on the Miyazawa interrelational multiplier tradition and recent advances in distributional national accounts, the procedure connects three hitherto separate data infrastructures–Eurostat's FIGARO symmetric input-output tables, the Survey on Income and Living Conditions (SILC), and the Household Budget Survey (HBS)–through statistical matching, grossing up to national accounts totals, and bidirectional linkage via the **C** and **V** matrices. The empirical application to Spain (2010–2023) traces industry-level price shocks through the full circular flow of income, from functional income generation in production to the personal distribution of real purchasing power. The cost-push potential of sectoral shocks is highly concentrated, and principal component analysis of COICOP-level transmission patterns identifies three distinct channels, the most prominent of which groups essentials and is related to systematically regressive outcomes, with sign variability across industries implying that shocks in accommodation and recreational activities can be inequality-reducing. Decomposing real income changes into a price channel and an income channel further reveals a structural misalignment in essential-goods industries, where the price burden falls on the bottom deciles while wage gains accrue to middle ones. The paper contributes to a mesoprudential monitoring framework for inflation inequality and makes publicly available the complete 2010–2023 DE-IOTs series for Spain–replicable across the European Statistical System–as supplementary material.

Keywords: input-output analysis, distributional accounts, inflation inequality.
JEL classification: C67, E01, D31, E31.

1 Introduction

The 2020s have witnessed an unforeseen rise in inflation after decades of price moderation in advanced economies (International Monetary Fund, 2022; Auer et al., 2024). While the postpandemic recovery and Russia's invasion of Ukraine triggered the highest inflation surge since the 1970s (Bernanke and Blanchard, 2024; Weber et al., 2024a; Galbraith, 2023; DeLong, 2023; De Santis, 2024; Boissay et al., 2023; Weber and Wasner, 2023; Eickmeier and Hofmann, 2022), climate change and the latest Gulf war are introducing disruptions to supply chains and the price system as a permanent feature of the global economy (Weber et al., 2024b). Save for a global depression, price instability is here to stay.

From a single industry, cost shocks can ripple through the interrelated structure of the economy (Galbraith, 2023; De Santis, 2024; Bańbura et al., 2023; Chen et al., 2023), aplifying the initial impact and enabling actors to sway the distribution of factor income in their favour (Weber and Wasner, 2023; Uxó et al., 2025). At the hosuehold level, this creates winners and losers whose final position depends not only on the variance in exposer to inflation but also on how households tap into different revenue sources besides labor compensation (Pallotti et al., 2023; Kiss and Strasser, 2024; S. Basso et al., 2023; Jaravel, 2021; Lewbel and Houthakker, 2018; Wimer et al., 2022; Milanović, 2019). Together, the price and income channels identify the accounting connection between material living standards and interindustry dynamics, which is of central importance for the effectiveness and inclusiveness of inflation abatement policies (Van'T Klooster and Weber, 2024; Bernanke and Blanchard, 2024).

The G-20 Data Gaps Initiative III.8 recommendation have hilighted the need for statistics that measure the distributional constraints of policy choice. Substantial institutional effort have beed placed into the construction of such accounts for wealth, income and consumption to bridge the gap between macro benchmarks and micro diversity (Balestra and Oehler, 2023; Coli et al., 2022; OECD, 2024; Morais et al., 2023; UNDESA, 2013). Although this have coexisted with the policy interest on global supply chains leading to the official production of multirregional input-output databases (Remond-Tiedrez and Rueda-Cantuche, 2019), the accounting connection between industries and households represents a fundamental data gap to deliver on the III.8 recommendation in a context of supply disruptions and price instability.

This paper seeks to fill this gap by creating a compilation guide to produce distributionally extended input-output (DE-IO) tables connecting industries and households. Building upon recent work (e.g., Amores et al., 2024; Cazcarro et al., 2022; Alberti et al., 2023), it merges the 2010-2023 FIGARO symmetric input-output tables (IOTs) with the joint distribution of income and consumption of the household sector, which comes from a synthetic dataset matching statistically the Spanish Household Budget Survey (ES-HBS) and the Survey on Income and Living Conditions (ES-SILC) abinding by the national accounts totals. A complete 2010–2023 series of distributionally extended input-output tables for the Spanish economy is provided as supplementary material, along with additional evaluation and intermediate input files.

The analytical possibilities of these data are illustrated by tackling a pressing policy concern. The paper investigates the first-order implications of ., vcincome-driven interdindustry price shocks for the distribution of household purchasing power. This fits into the recent revival of the literature on impact analysis to understand the distributional consequences of inflationary shocks (Amores et al., 2024; Van'T Klooster and Weber, 2024; Weber et al., 2024a; Fahr et al., 2024; Paul et al., 2021; Pichler et al., 2022; Feenstra and Sasahara, 2018), to which it contributes with an improved micro-to-macro linkage approach to the connection between functional and personal real income distribution and a more consitent derivation of a shocked Consumer Price Index (CPI) (Cazcarro et al., 2022; Mongelli et al., 2010; OECD, 2024; Coli et al., 2022). The paper finds substantial heterogeneity in the cost-push potential across industries, three distinct inflationary transmission channels, and sign variability in their distributional effects. While price shocks in some industries (e.g., food or energy production) disproportionately burden households at the bottom of the distribution, others (e.g., accommodation or recreational activities) reduce inequality.

The paper starts in Section 2 with a short review of previous studies on the distributional price shocks in 2.1 and household disaggregation in input-output analysis 2.2. Section 3 documents the methodological steps and choices involved in the creation of a distributional input-output model of the Spanish economy, starting with the statistical matching in Section 3.1, followed by the grossing up to national accounts in Section 3.2, and finally the income and expenditure linkages between the household dataset and the symmetric input-ouput industry tables, in Section 3.3 and Section 3.4, respectively. Section 4 briefly presents the basic Leontief price model and the accounting representation of the distributional model, followed by an empirical overview of the evolution of prices and incomes at the industry level in Section 5.1. Within Section 5, 5.2 quantifies each industry's inflationary potential and transmission channels and 5.3 evaluates the impact at the household level. Lastly, Section 6 summarizes the main conclusions and derives the most relevant implications of the paper.

2 Background studies

2.1 Distributional effects of price shocks

From an input-output modelling perspective, we find plenty of research on industry price shocks (Miller and Blair, 2022; Weber et al., 2024a,b; Valadkhani and Mitchell, 2002; Tunalı Çalışkan and Aydoğuş, 2007; Pichler and Farmer, 2022; Rubbo, 2024), but few contributions addressing the unequal impact of these price shocks on the household sector. Two main reasons explain this. First, in the tradition of the input-output framework the unit of analysis is the industry aggregate, so the interest in price shocks relates to their impact on the macroeconomic balance among sectors. Second, the accounting connection between industries and households, with as few modelling assumptions as possible, is hindered by condiderable data and methodological gaps.

Studying the distributional consequences of industry-specific price shocks requires widening the focus from industry to household heterogeneity. Since we observe differences in consumption baskets, changes to relative prices in addition to the price level will generate inflation inequality among households (Jaravel, 2021). Furthermore, as differences in consumption choices strongly correlate with the position in the income distribution (Lewbel and Houthakker, 2018; Chai and Moneta, 2013; Kaus, 2013), shocks to certain industries would carry a systematic and disproportionate impact on different social groups: if the good whose price rises absorbs a larger expenditure share among lower-income households, as is characteristically the case for food and energy, the shock will be regressive.

The evidence on inflation inequality is robust. Bettarelli et al. (2023) examine 129 economies over 1970–2013 and find that energy price inflation is associated with a significant increase in the Gini coefficient, with a larger effect in developing countries. Kröger et al. (2023) provide granular evidence for Germany, showing that the 2022 natural gas price spike disproportionately burdened poorer households. In Spain, Ferreira et al. (2024) decompose the 2021 inflation surge into wealth, income, and relative consumption channels, while S. Basso et al. (2023) estimate that lower-income households faced inflation rates approximately two percentage points higher than higher-income ones in 2021, driven by food and energy shares. At the EU level, Menyhért (2023; similarly to Claeys et al., 2024) shows that in all member states except the Netherlands the bottom quintile experienced higher inflation than the top in 2022, driven mainly by food and energy expenditure differences. Gürer and Weichenrieder (2020) demonstrate that ignoring group-specific price indices leads to underestimating the European Gini by nearly 0.04 points.

These findings underscore the importance of identifying industries whose price instability carries the greatest potential to increase inequality. The input-output price model offers a natural framework, because it captures both the direct effect on consumption baskets and the indirect propagation through interindustry cost linkages. Weber et al. (2024a) introduce the concept of *systemically significant prices* for inflation, using the Leontief price model to show that a small set of US sectors—petroleum, oil and gas, utilities, chemicals, farms, food and beverages, housing, and wholesale trade—have a disproportionate capacity to trigger economy-wide inflation. Lara Jauregui et al. (2025) extend this framework by incorporating decile-specific consumption baskets, computing the change in the Gini coefficient associated with each sectoral shock. They find a substantial overlap between sectors significant for inflation and those significant for inequality, and show that the combined 2022 price shock to the eight identified sectors produces a Gini increase equivalent to nearly one year of the average annual inequality growth during 1980–2021. Weber et al. (2024b) apply a related methodology to Germany in the context of carbon pricing, confirming that interindustry propagation amplifies the regressive bias beyond what direct consumption shares alone would predict. Recent institutional efforts have sought to embed consumption heterogeneity more formally within multi-sectoral models. Amores et al. (2024) combine the EUROMOD microsimulation model with household survey data to quantify the distributional impact of the 2022 inflation surge across EU member states, estimating that fiscal measures compensated roughly one third of the welfare loss.

This literature establishes three findings upon which the present paper builds: the sectoral origin of a price shock is a first-order determinant of its distributional incidence; the input-output price model provides an appropriate framework for tracing both direct and indirect transmission channels; and the data infrastructure linking household surveys with interindustry accounts, while advancing rapidly, remains incomplete regarding the simultaneous integration of income and expenditure heterogeneity. The distributional extension proposed in this paper addresses this last gap.

2.2 Household disaggregation in input-output models

The analytical foundations for introducing household heterogeneity into the input-output framework were laid by Miyazawa (1976b; see also Miyazawa, 1976a). Miyazawa's central insight was to disaggregate the scalar household income and consumption entries of the standard closed Leontief model into matrices, thereby replacing the single representative household with q socioeconomic groups. He defined a matrix of interrelational income multipliers that captures the feedback loop between household consumption patterns and the generation of factor income across industries, demonstrating that conventional multipliers understate the total economic effects when household-industry feedbacks are ignored, and that distributional consequences depend critically on which household groups are linked to which industries through both channels. In this formulation, both the income received by group h from industry j and the consumption expenditure of group h on industry i jointly determine the interaction between the productive structure and the personal distribution of income.

Miyazawa's framework inspired a broader research programme on social accounting matrices (SAMs), which extend input-output tables by embedding them within a comprehensive representation of the circular flow of income. The SAM tradition acquired its distinctive distributional orientation through Pyatt and Thorbecke (1976) and Pyatt and Round (1985), who formalized the accounting principles for recording transactions between production activities, factors of production, household groups, government, and the rest of the world, making explicit the mapping from the functional distribution of income—how value added is divided between labour compensation, operating surplus, and mixed income—to the personal distribution across household groups. Pyatt and Round (1979) provided the foundational multiplier decomposition, distinguishing between transfer, open-loop, and closed-loop effects (Miller and Blair, 2022, p. 517). Stone (1985) demonstrated how the household sector could be disaggregated within the national accounts using SAM principles, showing that a consistent disaggregation requires the simultaneous partitioning of both the factor income receipts row and the consumption expenditure column, and observing that higher-order circular effects of demand injections on the distribution of income tend to be relatively invariant to the sectoral composition of the initial shock. Defourny and Thorbecke (1984) refined this tradition by applying structural path analysis to the SAM framework, decomposing the global multiplier into elementary transmission channels and tracing the specific paths through which factor payments in one industry reach particular household groups via the income-expenditure circuit.

Subsequent contributions have further developed these analytical tools. Sonis and Hewings (1999) formalized the mathematical properties of the interrelational income multiplier, while Kimura and Kondo (1999) offered alternative derivations highlighting the conditions under which the income distribution effects are well-defined. Okuyama et al. (1999) adapted the Miyazawa multiplier to assess distributional consequences of supply-side disruptions. Pyatt (2001) traced the historical convergence of Miyazawa's interrelational multiplier and the SAM-based decomposition of Pyatt and Round. More recently, Steenge and Serrano (2012) revisited the conditions under which the Miyazawa framework yields economically meaningful distributional results, and Steenge et al. (2020) examined the stability and convergence properties of the multiplier process, clarifying the theoretical conditions for consistent and interpretable distributional multipliers. Pyatt and Round (2006) have further demonstrated how the SAM multiplier framework can be applied directly to the analysis of poverty reduction, linking changes in the functional distribution of income to the household distribution.

In parallel, a distinct body of work on demographic-economic extensions emerged through Batey and collaborators. Batey and Madden (1983) developed extended input-output models incorporating household demographic characteristics—notably employment status—into the interindustry system, converting the household activity variables from monetary units to population counts. Batey (1985) provided a comparative survey of the many ways the static Leontief model can accommodate demographic-economic relationships, identifying a family of models ranging from simple household closures to comprehensive formulations with multiple household types. Batey et al. (1987) demonstrated the substantial sensitivity of multiplier estimates to the level of household disaggregation—a finding of direct relevance to the distributional extensions pursued here. Batey and Rose (1990) synthesized this programme in an influential survey, emphasizing both the progress achieved and the remaining potential for richer socioeconomic modelling through extended input-output structures.

On the empirical side, substantial effort has been devoted to operationalizing the disaggregation of household accounts within the SAM framework. Roland-Holst and Sancho (1992) constructed a disaggregated SAM for the United States in which the household sector was partitioned by income bracket, and used the linear multiplier structure to study relative income determination, demonstrating that the distributional effects of sectoral shocks are highly sensitive to the configuration of the factor-income-to-household mapping. Roland-Holst and Sancho (1995) subsequently extended this analysis to the price side, developing an intersectoral price model that captures the interdependence among activities, households, and factors, providing a complete set of accounting-consistent prices. In a related effort, Rose et al. (1999) presented the full methodology for constructing an input-output income distribution matrix for the US economy, disaggregating personal income accounts by eleven income brackets and eighty sectors, and introducing for the first time a procedure that accounts for multiple jobholders and multiple-earner households. Rose and Beaumont (1988) applied the Miyazawa interrelational income distribution multiplier to a regional SAM, quantifying the feedback effects between industrial structure and household income groups. Mongelli et al. (2010) integrated a heterogeneous household demand system into the input-output framework, demonstrating that the assumption of a single representative consumer introduces significant bias into impact assessments. A series of papers by Kim and Hewings (2012), Kim et al. (2015), and Kim and Hewings (2015) developed extended econometric input-output models accommodating household heterogeneity along dimensions such as income, migration status, and age structure, while Katris et al. (2017) disaggregated the UK Social Accounting Matrix to report household income quintiles, providing a practical template for distributional SAMs in the European context. Round (2003) synthesized the methodological lessons from decades of applied SAM work, noting that while the disaggregation of the consumption vector had become relatively standard, the simultaneous partitioning of the labour compensation row remained comparatively underdeveloped.

Most recently, Cazcarro et al. (2022) have presented the full procedure for linking European household budget survey microdata with multi-regional input-output databases, providing a systematic estimation of the bridge matrices needed at the EU level. Their work addresses the well-documented difficulties arising from differences in classification systems and coverage between micro and macro sources. Alberti et al. (2023) have extended this approach to examine poverty and the functional distribution of income within the input-output framework. The FIDELIO model (Hu et al., 2020) at the European Commission's Joint Research Centre has operationalized the SAM module within a broader framework linking input-output and general equilibrium approaches, endogenizing households and government to capture distributional effects. At the intersection with microsimulation, Peichl (2015) has surveyed approaches for linking microsimulation and computable general equilibrium models. Institutional efforts to develop distributional national accounts have further expanded the scope for this integration (Coli et al., 2022; OECD, 2024; Balestra and Oehler, 2023).

Despite these advances, no unified methodology has been presented for extending both the income and expenditure sides of the input-output system simultaneously using household survey data integrated with national accounts totals. Much of the existing work has focused on one channel at a time—typically the consumption side—

or has relied on aggregate SAM structures without micro-founded distributional detail. The present paper seeks to address this gap by building on the accumulated insights of the Miyazawa-SAM tradition, the demographic-economic extensions of Batey and collaborators, and the recent empirical contributions on micro-macro linkage.

3 Distributionally extended input-output tables: Linking industry and household data

The possibility of adding a distributional dimension into input-output data is premised on the compatibility of different information sources. Whereas household surveys are one among the many inputs that come into the compilation of macroeconomic statistics, they are not fully consistent with the latter (Piketty et al., 2018). A main advantage of national accounts (NA) despite averaging over household heterogeneity is that they follow a "methodology harmonised at global level (the System of National Accounts) and are thus generally comparable worldwide" (Coli et al., 2022, p. 7). It is a consistent framework developed over decades that uses multiple data sources to provide consistent and comprehensive data on income, consumption and wealth (Lequiller and Blades, 2014).

In general, there are at least three main fault lines that separate micro and macro sources. First, the concepts used often fail to show a complete overlap in the definition of common variables or reference dates, such that reclassifications or auxiliary assumptions are needed for conceptual correspondance (OECD, 2020). Second, the population scope of household surveys typically exclude some units (e.g., people living in prisons or retirement homes), which requires upscaling survey weights to match the population totals from NA (OECD, 2020). Third, household survey totals do not match reference macro targets, sometimes by a substantial margin (Honkkila and Kavonius, 2013; Morais et al., 2023). Put together, these issues can create considerable bias in our understanding of the connection between micro to macro states (Golan, 2017; dos Santos, 2017).

The input-output framework, on the other hand, presents a consistent representation of the productive structure of the economy in complience with the main macromagnitudes (Miller and Blair, 2022). In the same way that household heterogeneity challenges the assumption of a representative agent (Kirman, 1992; Shaikh, 2016), the introduction of structural diversity into a single-sector macroeconomic model has shed light on the importance of sectoral composition for development (Rodrik, 2006) or, as more recently, to improve our understanding of inflation transmition channels (Van'T Klooster and Weber, 2024).

By linking household survey data and conventional input-output tables we can introduce heterogeneity in the relationship between the household sector and the interindustry system. Although this connection has been explored in the past, no fleshed-out, coherent method has been presented on how to extend distributionally

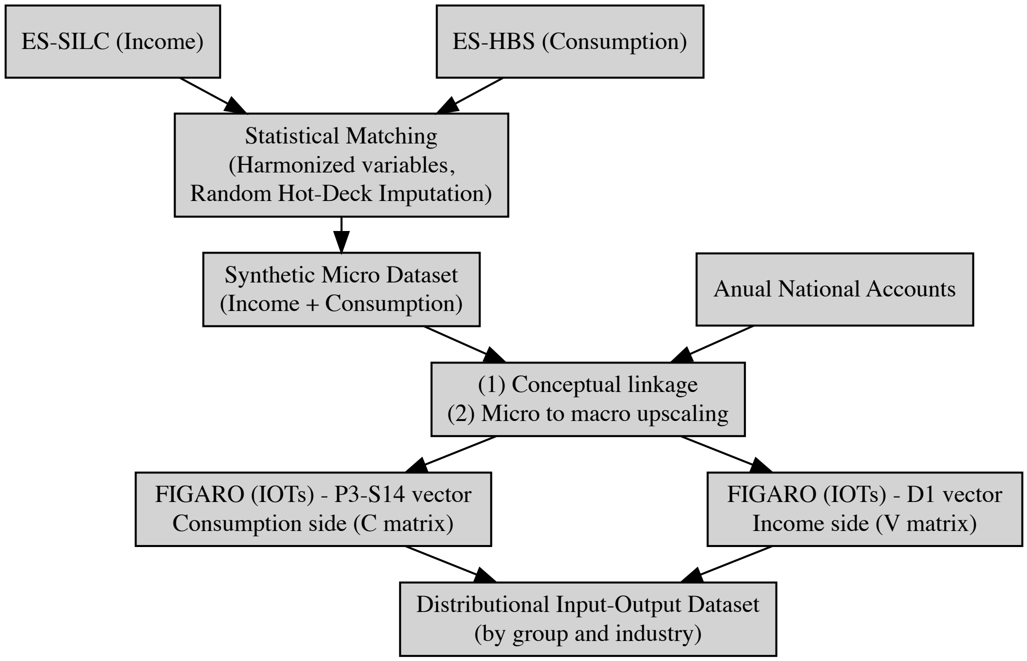


Figure 1: Distributionally extended input-output dataset compilation procedure

the household sector in the input-output system from the income and expenditure sides simultaneously. While Figure 1 portrays the four-stage process for integrating micro-level household survey data with the conventional input-output framework: statistical matching to derive joint income-expenditure distributions, conceptual mapping, gap allocation, and bidirectional linkage between households and industries through both consumption and income channels, which we will refer to as the **C** and **V** matrices, Figure 3 shows the structure of a DE-IOT.

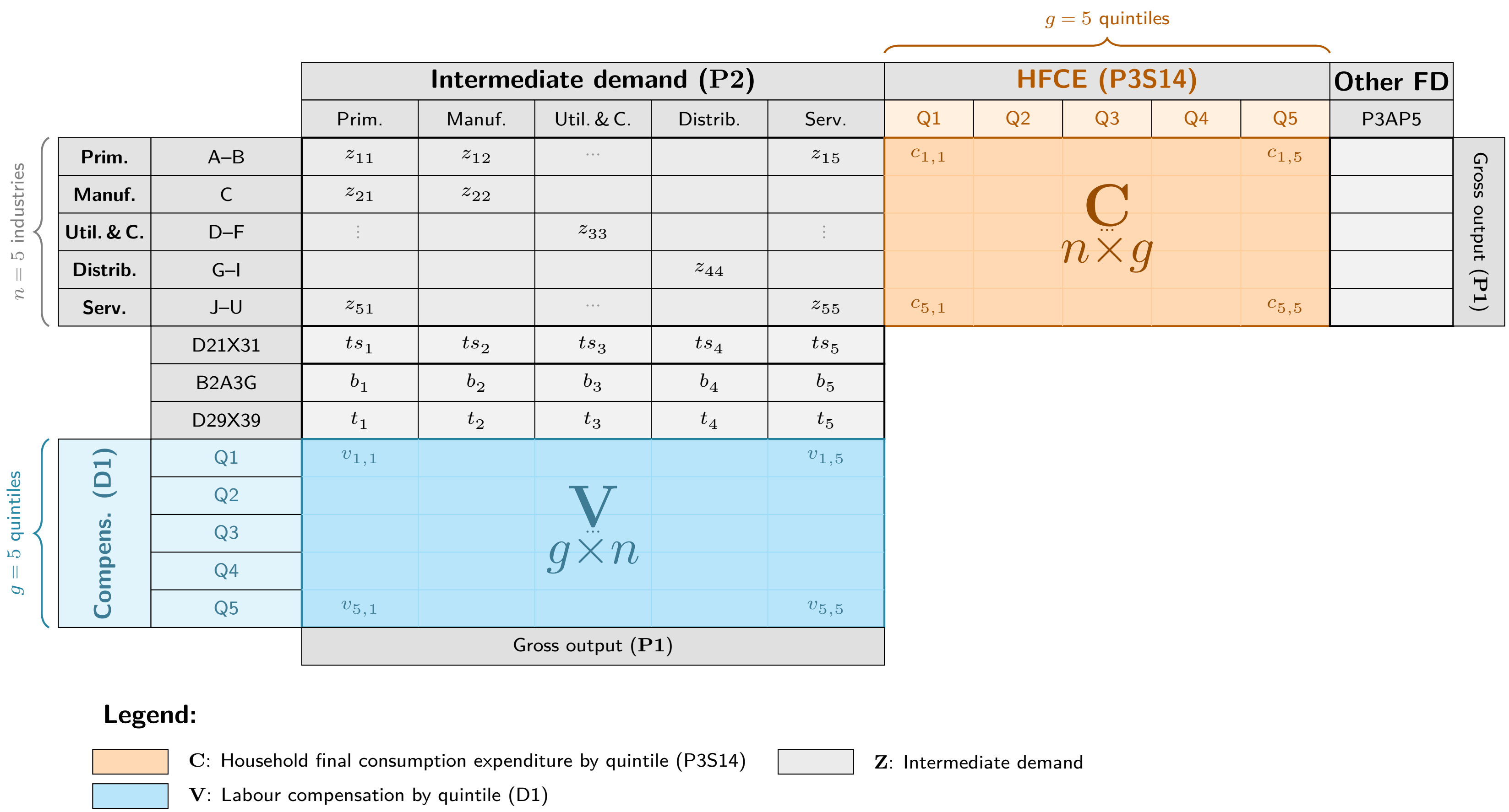


Figure 2: Stylised structure of the distributionally extended IOT. Industries are aggregated into five broad groups (Primary: A–B; Manufacturing: C; Utilities & Construction: D–F; Distribution: G–I; Services: J–U). The highlighted blocks correspond to the distributional extensions: the **C** matrix (household final consumption expenditure, P3_S14, disaggregated by income quintile) and the **V** matrix (labour compensation, D1, distributed by income quintile). *Source*: Author's elaboration.

Definition	ES-SILC	Hellinger	Frequency
Type of contract	PL140	0.010	13
Tenure status	HH020	0.016	16
Ventile (5%)	-	0.021	16
Degree of urbanisation	DB100	0.027	16
Gender	PB150	0.042	16
Profesional status	PL040A	0.043	14
Age bracket	PB140	0.047	16
Economic status	PL032	0.048	16
Marital status	PB190	0.049	16
Ocupation	PL051A	0.077	16
Education level	PE041	0.114	9

Table 1: Matching variables ranked according to the average Hellinger distance metric across the period

3.1 Deriving the joint distribution of income and expenditure

First, a single synthetic dataset is derived out of the statistical matching of the 2010-2023 series of the Survey on Living Conditions (ES-SILC) and the Household Budget Survey (ES-HBS). Since ES-SILC provides very complete information on household disposable income, unlike ES-HBS, but no information on expenditure (López-Laborda et al., 2021), it becomes necessary to link observations from both surveys to derive the joint distribution of income and expenditure at the household level. For this case, we choose to apply a statistical matching approach to create synthetic observations that approximate as much as possible actual households' income and consumption profiles (Balestra and Oehler, 2023; Serafino and Tonkin, 2017). This method assumes that similarity across household characteristics implies exchangeable income or expenditure information on average (Balestra and Oehler, 2023; D'Orazio et al., 2006). For this case, we prioritize completeness in the income information by making ES-SILC the recipient dataset and ES-HBS the donor dataset (Balestra and Oehler, 2023; López-Laborda et al., 2021).

Prior to the statistical matching, a set of common variables needs to be identified and harmonized across the two surveys. This includes the selection of the reference person of the households, whose sociodemographic information is used for group aggregation at later stages. According to Balestra and Oehler (2023), the matching variables have to display similar distributions and bear a significant correlation to the variables of interest to guarantee the conditional independence assumption (CIA). The selection of these variables is guided by the < 0.15 threshold on the calculated annual Hellinger distance measure for all sociodemographic characteristics of the household. The list of matching variables with their corresponding ES-SILC codes and the average Hellinger distance in the 2010-2023 period is provided in Table 1.

The statistical matching is performed via random hot-deck imputation with constrained replacement (Balestra and Oehler, 2023; D'Orazio et al., 2006). This is illustrated in Figure 3. From the stratum of donor households matching the chosen characteristics of the recepient observation, we select one in random. This donor household is constrained to appear a maximum of three times in the donor pool. In case no donor household meet the criteria, the number of matching variables can be relaxed by dropping each one at a time (D'Orazio et al., 2006). A fallback order is defined by the Hellinger distance ranking, where most disimilar variables are dropped first. After running sensistivity analyses, we decide that in no case more than 5 out of 11 (55%) matching variables could be discarded. In the few instances that an unmatched observation remains (e.g., 0.6% in 2023) we relax the matching set down to the distributional variables with no replacement constraint. We flag that observation and check that on average it does not bias the results. In general, the variable attrition rate is very sensitive to the relative size of the donor dataset, which tended to be 60% relatively larger up to 2021 (e.g., 13597 vs 22203 in 2010 and 21007 vs 19394 in 2021, for ES-SILC and ES-HBS, respectively). Figure 15 in the appendix shows the actual distribution of donor uses and variables dropped.

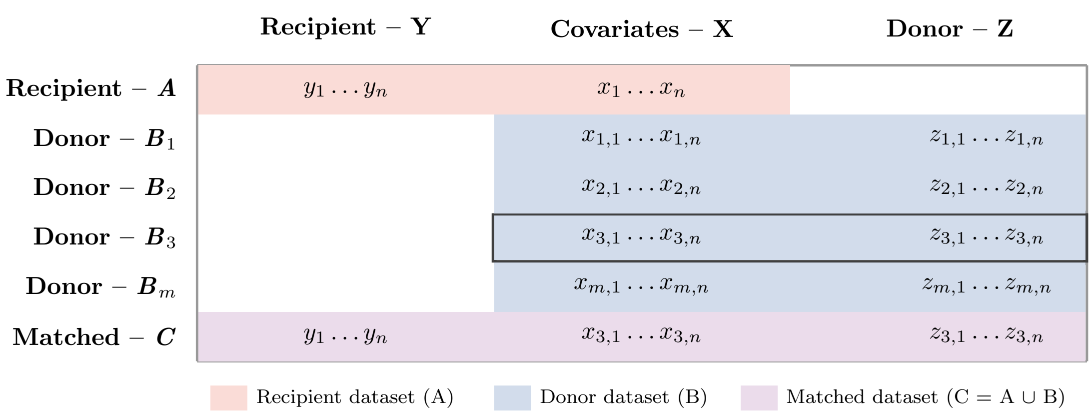


Figure 3: Schematic representation of two survey matching

To verify the robustness of the match, we apply several sanity checks to the procedure (e.g., matching number of observantions) and rely on the visual inspection of the marginal distributions of expenditure at different aggregation levels, as illustrated in Figure 4 for household expenditure on food products. We find an almost complete overlap in the aggregate, as well as at 2 and 3 digits. We provide additional robustness tests in Appendix B.

3.2 Meeting the distributional national accounts requirements

The synthetic dataset resulting from the statistical matching procedure imputes the joint distribution of income and consumption on ES-SILC. However, survey income

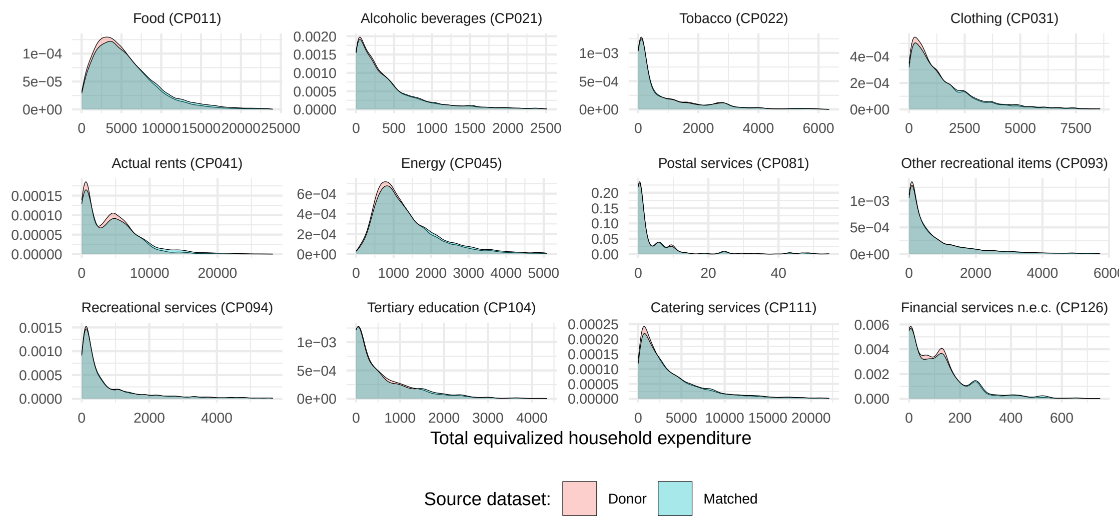


Figure 4: Household expenditure on selected 3-digit consumption purposes before and after the statistical matching, 2023

and expenditure totals are not sufficient to produce a linkage with industry data. In fact, due to their different objectives and compilation procedures, survey totals are well-known to fall short of the benchmark set by the Anual National Accounts (ANA) (Morais et al., 2023). Following the call by Stiglitz et al. (2009) for better statistics to improve our measurement of welfare, the G-20 Data Gaps Initiative (DGI) has framed multiple efforts to bring the heterogeneity of household microdata in line with the completeness of aggregate statistics to render visible the distributional impacts of policy choices. Most importantly for this work, this linkage would allow to create a connection between micro observations and IOTs.

Drawing from the methodological work of the Eurostat-OECD Expert Group on joint distributions of Income, Consumption and Wealth at micro level (EG ICW), documented by Coli et al. (2022) and the recent OECD Handbook on the compilation of distributional national accounts (OECD, 2024), this paper follows the approach to bring household distributional results in line with national accounts totals. This consists of a three-fold procedure. Firstly, we implement a proportional adjustment of survey weights to match the target population of ANA –this typically involve small discrepancies related to different statistical frames.

Secondly, a conceptual correspondance between survey variables and the national accounts framework is required to produce the vertical linkage, which, in turn, depends on the level of aggregation of the macrodata. For instance, FIGARO IOTs presents a single vector of industry observations for the gross operating surplus and mixed income (B2AB3G), but in surveys we find information about, on the one hand, imputed rents and, on the other, cash benefits or losses from self-employment and value of goods produced for own consumption, which are themselves proxies for *part of* the total gross operating surplus (B2G) and mixed income (B3G). This paper relies com-

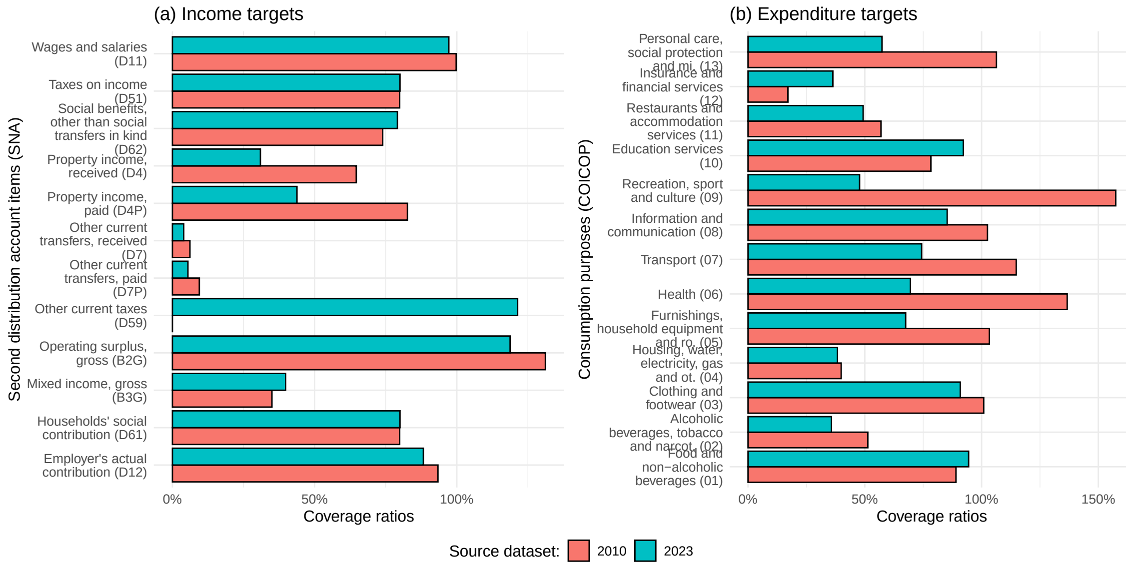


Figure 5: Initial SILC-HBS to national accounts coverage ratios, 2023.

pletely on the EG ICW expertise by adopting the correspondence between ES-SILC and ANA income items (see Coli et al., 2022, p. 23), whose conceptual mapping is reproduced in Table 5. As for consumption, the correspondance is strightforward since ANA and the household budget survey present their expenditure information under the Classification of individual consumption by purpose (COICOP).

Thirdly, to arrive at comparabe totals, it is necessary to gross up "the relevant components from the micro data sources with the income and consumption variables from national accounts" (Coli et al., 2022, p.13). This is carried out item by item using a specific gap allocation method. Coverage across income items is more uneven than for expenditure, with cases such as other current transfers, received (4%) and other current taxes (119.6%) in 2023. Assuming macro measurement dominance, any deviation (up or down) from a 100% coverage ratio needs to be addressed. When the nature of the gap is unknown, an allocation method is needed. Eurostat proposes several alternatives: proportional allocation, Pareto tail modelling, progressive quantile scaling, or a combination of more than one method. Each one relies on specific distributional assumptions. Proportional allocation keeps the distribution invariant, while Pareto tail modelling assumes a power law distribution above a given threshold (Morais et al., 2023). Most importantly, the choice of a gap allocation method might differ for income and consumption.

Following Eurostat's centralized exercise for income, this paper adopts a combined gap allocation strategy that differentiates by income item. First, it assigns proportional scaling for items where the survey distribution is expected to be close to the true distribution, such as wages and salaries (D11), employers' social contributions (D12), social benefits other than STiK (D62), net social contributions (D61), taxes on income (D51), gross operating surplus (B2G), and property income paid (D4P). Second, it applies Pareto tail modelling for property income received (D4R), which

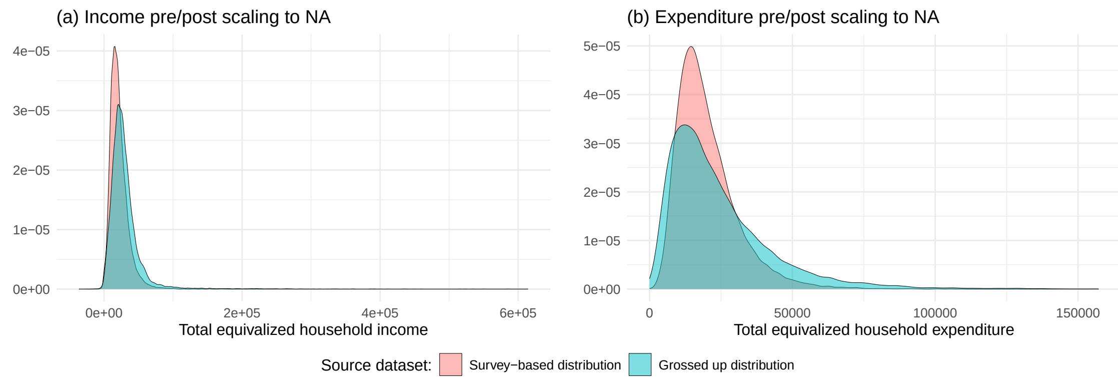


Figure 6: Marginal distribution of total equivalized household income pre/post grossup to national accounts, 2023

we simplify as allocating proportionally all the gap to the top decile. Third, for mixed income (B3G) and other current taxes (D59P) the method follows standard bottom-up ascending decile gap shares. Fourth, a flatter ascending decile gap shares allocation for other current transfers, received and paid (D7). The different gap shares are reported in Table 2.

A few additional corrections are needed to prevent the gap allocation of items with low coverage ratios from producing extreme values that distort household ordering. This concerns other current transfers (D7), received and paid, as well as other current taxes (D59). For the former, we redistribute the gap uniformly across all households, following the rationale of Eurostat's allocation albeit with a simpler distributional assumption. As for the latter, we allocate the gap uniformly across households either receiving interest, dividends, or profit income from capital investments in unincorporated business or income from rental of a property or land .

Allocation gradient	**D1**	**D2**	**D3**	**D4**	**D5**	**D6**	**D7**	**D8**	**D9**	**D10**
Income (to-the-top)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
Income (standard bottom-up)	0.00	0.00	0.02	0.04	0.06	0.08	0.14	0.16	0.18	0.22
Income (flatter bottom-up)	0.04	0.06	0.07	0.08	0.09	0.11	0.12	0.13	0.14	0.16
Expenditure (non-necessities)	0.00	0.00	0.02	0.04	0.06	0.08	0.14	0.16	0.18	0.22

Table 2: Gap allocation gradients by decile. *Notes:* Values represent the share of the aggregate micro-to-macro gap allocated to each decile, summing to unity. The to-the-top income gradient assigns the entire gap to the top decile. The standard bottom-up income gradient follows Eurostat's progressive top-concentration approach (M3.1) adapted to deciles. The flatter bottom-up gradient distributes the gap along a smoother progression akin to Eurostat's M3.3 for other current transfers. The expenditure gradient for non-necessities applies a progressive quantile scaling identical to the standard income distribution. Necessities are allocated proportionally across all deciles.

For expenditure, the grossing up is carried out at 3 digits, which involve 46 consumption purposes—redistributing proportionally the total for narcotics (CP023). The allocation leverage Engel's Law by proportionally distributing the gap for necessities, but according to a progressive quantile scaling for non-necessities as reported in Table 2. Our definition of necessities requires that the spending share of any 3-digit COICOP item is larger for the first than for the last quintile.

		Saving rate			
Year	Quintile	Original	Matched	Final	Eurostat
2010	Q1	-50.7	-77.6	-57.9	-20.3
2010	Q2	-21.2	-3.1	14.2	3.3
2010	Q3	-14.5	18.8	28.3	15.2
2010	Q4	-7.0	24.7	33.0	24.7
2010	Q5	4.5	42.8	46.5	37.0
2015	Q1	-51.6	-86.3	-51.5	-19.1
2015	Q2	-19.2	-2.8	11.7	6.4
2015	Q3	-9.8	18.8	28.5	18.1
2015	Q4	-4.6	34.4	36.8	27.3
2015	Q5	7.0	49.8	44.5	39.3
2020	Q1	-18.2	-22.4	-21.3	-6.7
2020	Q2	7.2	26.5	20.4	22.3
2020	Q3	15.8	43.5	35.9	32.6
2020	Q4	23.5	55.2	46.3	42.3
2020	Q5	31.8	70.3	58.3	53.9

Table 3: Comparison of the quintile saving rate at different stages of the compilation process with the official data by Eurostat. *Source*: Author's elaboration using data from INE and Eurostat.

We can compare the initial and final distribution of equivalized household income by looking at Figure 6 (a), where they display very similar shapes, albeit flatter and right-skewed in the case of the post-adjustment distribution. Figure 6 (b) shows the comparison between the pre- and post-adjustment distributions for income. The difference is larger than in the case of income, but the characteristic skewness to the right and the smoothness of the distribution remains. At a more disaggregated level we find a very similar pattern of close-resembling distributions, with no multimodality or degenerate behavior when not present in the original distribution. Despite following a more strightforward procedure to bring consumption records in line with national accounts, the pre- and post-adjustment income distributions are much closer to one another. This suggest the possibility of a higher degree of uncertainty in consumption data.

Lastly, we can investigate how well our implementation of the methodology works in comparison with Eurostat's own experimental statistics on distributional national accounts (Coli et al., 2022). Table 2 shows a comparison of savings rates by quintile across the original ES-HBS, the SILC-HBS statistical matching, the final grossed-up microdata, and Eurostat's. Since savings results from the interaction between income and consumption, it is an efficient way to test our results against official statistics. As we can see, the comparison suggests a reasonable approximation. In addition to comparing marginal distributions and sanity checks at each step, we can further inspect the robustness by looking at a few additional tests, such as the percentage of households changing decile as a consequence of the grossup process in **?@fig-flow** from Appendix B. Lastly, Table 3 shows the saving rate by income quintile at different moments during the compilation of the data. It also includes Eurostat's own distributional national accounts estimates, which are only available for every five years since 2010. Note that savings in this case are defined as total disposable income (B6G) minus total consumption expenditure (P31NC) (OECD, 2024, p.27). The first thing to notice is that the most dissimilar rates are those produced by the statistical matching. The final adjustment to national accounts totals corrects them and brings them fairly close to Eurostat's. This is a remarkable result given the few different methodological choices and the fact that the results in this paper rely entirely on open data. Nonetheless, further investigation on the allocation to the first quintile is required.

3.3 The C matrix: bridging household expenditures on products to industry production aggregates

In the input-output framework, household final consumption expenditure (P3_S14) is aggregated into a single column vector, averaging over sociodemographic differences among households (Remond-Tiedrez and Rueda-Cantuche, 2019; UNDESA, 2018; Lequiller and Blades, 2014; Mongelli et al., 2010; Stiglitz et al., 2009). Since Miyazawa (1976a), ubstantial efforts have been put in developing a consistent methodology to disaggregate this vector distributionally (Mongelli et al., 2010; Kim et al., 2015; Kim and Hewings, 2015, 2012; Katris et al., 2017). A recent paper by Cazcarro et al. (2022) has filled the methodological gap in the literature and provided the most detailed and robust connection between household expenditure surveys and the corresponding vector from the IOTs.

Constructing such a disaggregation requires overcoming three systematic obstacles. First, survey expenditure data are recorded according to the Classification of Individual Consumption by Purpose (COICOP), whereas IOTs report consumption according to the Statistical Classification of Products by Activity (CPA). Second, survey expenditures are valued at purchaser's prices, while IOTs are compiled in basic prices (Remond-Tiedrez and Rueda-Cantuche, 2019; Lequiller and Blades, 2014). Third, the IOTs are structured around industry output rather than product output, which requires a further transformation from products to industries. Each of these steps involves non-trivial data manipulation whose quality directly conditions the reliability of the distributional extension. The procedure in Section 3.2 has already addressed the well-known gap between survey and National Accounts totals (Coli et al., 2022); from the grossed-up COICOP expenditure totals therein, we follow three additional transformations to arrive at a matrix of household consumption that is fully compatible with the reference IOTs.

This procedure permits grouping by any sociodemographic characteristic, such that we can split the single aggregate HFCE vector into its group components (e.g., income deciles or age groups). Borrowing from Miyazawa (1976a), we call this subset of consumption vectors the $\underset{n\times g}{\mathbf{C}} = \{c_{i,d}\}$ matrix, where each element $c_{i,d}$ represents the flow from industry $i = 1, 2, ..., n$ to the final consumption expenditure demand (P3_S14) from households in group $d = 1, 2, ..., g$.

3.3.1 Classification bridge: COICOP to CPA

While surveys report consumption data according to the COICOP classification, expenditure aggregates in the IOTs follow the European Classification of Products by Activity (CPA). To perform this reclassification, we employ a contingency or bridge matrix $\mathbf{B} = \{b_{i,j}\}$, in which each element $b_{i,j}$ represents the expenditure amounts that have a direct and unique correspondence between the two classifications. In Europe, however, only a few countries report this information. To overcome this difficulty, we draw from the estimated bridge tables made available by Cazcarro et al. (2022), which link 64 CPA products to 46 three-digit COICOP items (excluding narcotics, CP023). To match our 2010–2023 series, a set of annual bridge matrices are projected using the RAS algorithm (Miller and Blair, 2022, ch. 4). The transformation is as follows,

$$\underset{cpa}{\mathbf{C}^{pp}} = \underset{coicop}{\mathbf{C}^{pp}} \cdot \mathbf{RAS}(\mathbf{B}^0, \mathbf{u}, \mathbf{v}) \tag{1}$$

where $\mathbf{B}^0$ is the 2010 bridge matrix (46×64), with target row sums $\mathbf{u}$ derived from COICOP aggregates and target column sums $\mathbf{v}$ from the CPA use table vector of household final consumption expenditure.

3.3.2 Valuation adjustment: purchaser's prices to basic prices

At this point, we obtain a $64 \times g$ matrix of product values at purchaser's prices. The IOTs, however, are compiled in basic prices (Lequiller and Blades, 2014). The accounting identity linking the two valuations is:

$$p_i^{pp} = p_i^{bp} + T_i + M_i \tag{2}$$

where p_i^{pp} and p_i^{bp} denote household consumption of product i in purchaser's and basic prices, respectively, T_i represents net taxes on products (taxes less subsidies), and M_i captures trade and transport margins (Eurostat, 2013, p. 17). Moving from purchaser's prices to basic prices requires knowledge of net tax rates and trade and transport margin rates, which are confidential for several countries, including Spain and Germany. Following Mongelli et al. (2010) and Cazcarro et al. (2022), we derive implicit rates from available supply-use table components and apply a stepwise transformation algorithm $F(\cdot)$ whose details are presented in Appendix A. The resulting transformation can be written compactly as:

$$\underset{cpa}{\mathbf{C}^{bp}} = \mathbf{F}(\underset{cpa}{\mathbf{C}^{pp}}) \tag{3}$$

We ensure that the reference vectors used in the calculation match the corresponding entries in the FIGARO IOTs, even where this implies some deviation from other intermediate annual national accounts figures.

3.3.3 Industry transformation: CPA to NACE

Lastly, we need to match the CPA product data at basic prices to the NACE industry classification used in the reference IOTs. Following Eurostat (2008, p. 349), we convert CPA data using the market shares matrix $\mathbf{D} = \mathbf{V}'\hat{q}^{-1}$, where $\mathbf{V}'$ is the transpose of the supply matrix (product by industry) and $\hat{q}^{-1}$ the inverse of the diagonalized column vector of product outputs. This corresponds to Model D, the fixed product sales structure assumption (Miller and Blair, 2022, p. 203; Eurostat, 2008, p. 349):

$$\underset{nace}{\mathbf{C}^{bp}} = \mathbf{D} \cdot \underset{cpa}{\mathbf{C}^{bp}} \tag{4}$$

The resulting $n \times g$ matrix is directly compatible with the industry-by-industry FIGARO tables and constitutes the $\mathbf{C}$ matrix entering the extended model in Equation 9. When deriving synthetic CPI indices in Section 4, we reverse this process using commodity-technology Model B and the same bridge matrix, redistributing price changes from industries to consumption purposes at purchaser's prices.

3.4 The V matrix: linking household disposable income to the distribution of value added by industry

Although sufficient for many analyses (Mongelli et al., 2010), disaggregating only the household consumption vector provides a one-sided distributional extension of the input-output framework. Because households are conventionally ranked by their position in the income distribution, a complete specification requires integrating the income side of the industry-household connection as well. Following Miyazawa (1976a), we define $\mathbf{V} = \{v_{h,j}\}$ as the matrix of total labour income coefficients paid to wage earners in each income or socio-demographic bracket $h = \{1, 2, \dots, q\}$ per unit of output j. For the standard 64-industry breakdown and income deciles, $\mathbf{V}$ is a 10×64 matrix that replaces the 1×64 aggregate vector of labour compensation (D1).

Constructing $\mathbf{V}$ requires tracing household disposable income (B6G) back to gross value added (B1G) through the ESA 2010 sequence of accounts (Lequiller and

Blades, 2014). In practice, this means reversing the consolidation of transactions from the secondary distribution account to the production account, separating and reallocating each source of household revenue to its corresponding factor of production. For studies concerned solely with linking micro and macro data at the level of household disposable income, this reverse mapping is unnecessary (Coli et al., 2022; Morais et al., 2023). The difficulty here arises from the need to allocate each income transaction to one of the three components of gross value added from the income side—compensation of employees (D1), gross operating surplus and mixed income (B2A3G), and other net taxes on production (D29X39)—while respecting the consolidation of accounts across institutional sectors.

Three ESA 2010 accounts mediate between gross value added and household disposable income (Eurostat, 2013; Lequiller and Blades, 2014, pp. 176–179). The *generation of income account* distributes value added into compensation of employees (D1) and gross operating surplus and mixed income (B2A3G), net of other taxes and subsidies on production (D29XD39). The *primary income distribution account* adds property income (D4) to the operating surplus (B2G) and mixed income (B3G) of the household sector, together with the economy-wide total of employee compensation (D1). These two accounts record factor incomes proper; by folding property income into operating surplus, a vertical linkage with the gross value added vector in the IOTs can be established. The *secondary distribution account* traces the redistribution of primary income through taxes on income (D5), social contributions (D61), social benefits (D62), and other current transfers (D7), yielding disposable income (B6G) as its balancing item. Crucially, the ES-SILC definition of total household income corresponds to this last level of aggregation, not to the primary income distribution. For the purposes of this paper, we establish the vertical linkage at the primary income level—for B2A3G, D1, and D4—and treat the net government transfers from the secondary distribution account as exogenous to the input-output system.

Table 4 shows the resulting partial vertical linkage between the survey variables, consolidated at the household level, and the production account. It excludes other taxes and subsidies on production (D29X39), which have no counterpart in the survey, and the share of gross operating surplus (B2A3G) that accrues to the corporate sectors—whether reinvested or retained as undistributed earnings. Once the survey variables are harmonised with their national accounts counterparts, the data can be consolidated at any desired level of household disaggregation, for instance by income decile, occupation, gender, or education level of the reference person.

In addition to this sociodemographic grouping, households must be classified by industry to produce an effective linkage with the IOTs—yielding a matrix that is distributional along both the household and the industry dimensions. ES-SILC records the economic activity of the local unit for the main job of each household member (PL111A), classified according to NACE Rev. 2. As an approximation, we assign each household's total labour income to the NACE section of the reference person's primary occupation. This simplification abstracts from multi-earner

Production Account	SNA	Distribution Account	SILC	Definition	Variable
Operating surplus	B2A3G	Gross operating surplus	B2G	Imputed rent	HY030G
		Mixed income	B3G	Self-employment cash flows	PY050G
				Value of own consumption	HY170N
		Property income	D4	Interest, dividends, profits	HY090G
				Income from property or land	HY040G
				Interest repayments on mortgage	HY100G
Labor compensation	D1	Wages and salaries	D11	Employee cash or near cash income	PY010G
				Non-cash employee income	PY020G
				Company car	PY021G
		Employers' contributions	D12	Employers' insurance contributions	PY030G
				Contributions to private pension	PY035G

Table 4: Vertical linkage between survey variables and the production account.

households.[1]. Where industry information is missing for households with positive wages and salaries, the NACE code is imputed via a random hot-deck procedure with unlimited donor reuse and a six-variable fallback sequence.

A further challenge is that ES-SILC provides industry information only at the NACE 1-digit level (21 sections, A to U), whereas FIGARO disaggregates up to 64 industries. To bridge this gap, we exploit the Structure of Earnings Survey (ES-SES), which contains microdata on gross annual earnings by 3-digit NACE activities for all non-agricultural sectors. From these data, we derive the distributional shares of labour compensation across sub-industries within each NACE section, stratified by earnings decile. Although the correspondence between individual earnings deciles and equivalised household income deciles is imperfect, it provides a workable approximation. For the primary sector, where the SES has no coverage, we distribute aggregate labour income across sub-industries (A01 to A03) in proportion to each sub-industry's share of total agricultural labour compensation (Section A) as reported in the IOTs. The resulting $\mathbf{V}$ matrix matches FIGARO's 64-industry breakdown. Lastly, an additional $\mathbf{N}$ matrix, produced following partly the same agrregation procedure, reflects a 64-industry breakdown of the number of people employed by industry and group.

4 A distributional extension of price impact analysis of the household sector

4.1 The Leontief price model

Leontief's linear production model (1986) characterises an economy in static equilibrium with n industries, each producing a single composite commodity. The $n \times n$

[1] A refinement that allocates individual earnings to each member's own industry would improve precision but requires assumptions on intra-household income pooling that lie beyond the scope of this paper

matrix $\mathbf{Z} \equiv \{z_{ij}\}$ records interindustry transactions, the $n \times 1$ vector $\mathbf{y} \equiv \{y_i\}$ collects exogenous final demand, and the $n \times 1$ gross output vector $\mathbf{x} \equiv \{x_i\}$ satisfies the accounting identity $\mathbf{x} = \mathbf{Zi} + \mathbf{y}$ from the row side and $\mathbf{x}' = \mathbf{i}'\mathbf{Z} + \mathbf{v}'$ from the column side, where $\mathbf{v}'$ is the $1 \times n$ vector of value added (Miller and Blair, 2022, p. 9).[2] Dividing each column of $\mathbf{Z}$ by the corresponding industry output yields the $n \times n$ matrix of technical coefficients,

$$\mathbf{A} = \mathbf{Z}\hat{\mathbf{x}}^{-1} \tag{5}$$

where $\hat{\mathbf{x}} \equiv \text{diag}(x_1, \dots, x_n)$. The open Leontief system $(\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{y}$ can be solved for gross output as $\mathbf{x} = \mathbf{Ly}$, where $\mathbf{L} = (\mathbf{I} - \mathbf{A})^{-1}$ is the Leontief inverse. Post-multiplying $\mathbf{L}$ by a final demand vector traces how an exogenous demand change propagates through the interindustry network–the basis of *demand-pull* impact analysis (Miller and Blair, 2022).

Alternatively, the same framework can be read from the supply side. Normalising the column-sum identity $\mathbf{x}' = \mathbf{i}'\mathbf{Z} + \mathbf{v}'$ by the output vector, we obtain $\mathbf{i}' = \mathbf{i}'\mathbf{A} + \mathbf{v}'_c$, where $\mathbf{v}'_c$ is the vector of primary-input shares in total output. Since all values are expressed per unit of output, they can be interpreted as base-year index prices equal to unity. Denoting these prices by $\tilde{\mathbf{p}}' = [\tilde{p}_1, \dots, \tilde{p}_n]$, the model can be solved with respect to prices (Miller and Blair, 2022, p. 45):

$$\tilde{\mathbf{p}} = (\mathbf{I} - \mathbf{A}')^{-1}\mathbf{v}'_c \tag{6}$$

Technology, represented by $\mathbf{A}$, is taken as given, whereas primary-input costs are treated as exogenous. Changes in unit factor costs propagate through the interindustry structure, capturing the mechanics of *cost-push* inflation.

4.2 The accounting structure of distributionally extended IOTs

The distributional extension developed in Section 3 decomposes the income and consumption dimensions of the household sector into subgroups. In the standard open Leontief model, household income and consumption are part of exogenous final demand. To internalise them, we close the technical coefficient matrix with respect to households by moving the household consumption column and the labour-payments row inside the system (Miller and Blair, 2022, p. 36). The resulting $(n + g)$-equation system can be partitioned as:

$$\begin{bmatrix} \mathbf{x} \\ x_{n+1} \end{bmatrix} = \begin{bmatrix} \mathbf{I} - \mathbf{A} & -\mathbf{h}_c \\ -\mathbf{h}_r' & (1 - h) \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{f}^* \\ f^*_{n+1} \end{bmatrix} \tag{7}$$

[2] The standard IO assumptions include product homogeneity, fixed-proportions production functions, no input substitution, no economies of scale, exogeneity of primary inputs and final demand (unless the model is closed with respect to any of them), and semi-positivity of the interindustry matrix (Miller and Blair, 2022, ch. 2).

where $\mathbf{h}_r^{'}$ is the vector of income received by households per unit of industry output and $\mathbf{h}_c$ the vector of each industry's household sales divided by total household income x_{n+1}. Both $\mathbf{h}_r^{'}$ and $\mathbf{h}_c$ can be separated into g distinct vectors corresponding to grouped households (e.g., income deciles), which is the purpose of the distributional extension.

Following Miller and Blair (2022, p. 517), the full accounting structure can be represented by a partitioned matrix $\bar{\mathbf{G}}$ containing the endogenous block $\bar{\mathbf{Z}}$, the matrix of exogenous final expenditures $\mathbf{F}$, the matrix of exogenous income categories $\mathbf{W}$, and the matrix of exogenous income allocations to final expenditures $\mathbf{B}$:

$$\bar{\mathbf{G}} = \begin{bmatrix} \bar{\mathbf{Z}} & \mathbf{F} \\ \mathbf{W} & \mathbf{B} \end{bmatrix} \tag{8}$$

In this paper, $\mathbf{W}$ contains other taxes on products (D21X31) and gross operating surplus (B2A3G). The endogenous block $\bar{\mathbf{Z}}$ is itself partitioned into the interindustry transaction matrix $\mathbf{Z}$, the matrix of endogenous value-added inputs $\bar{\mathbf{V}}$, and the matrix of endogenous final-demand expenditures $\bar{\mathbf{C}}$. Both $\bar{\mathbf{V}}$ and $\bar{\mathbf{C}}$ can contain as many rows and columns as subgroups in the distribution. Normalising as in Equation 7 yields the coefficient matrix:

$$\mathbf{S} = \begin{bmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{V} & \mathbf{0} \end{bmatrix} \tag{9}$$

where, for a decile breakdown ($g = 10$), $\mathbf{A}$ is the 64×64 matrix of technical coefficients, $\mathbf{C}$ the 64×10 matrix of endogenous final demand coefficients, and $\mathbf{V}$ the 10×64 matrix of endogenous value-added input shares (Miller and Blair, 2022, p. 517). The construction of $\mathbf{C}$ and $\mathbf{V}$ from household microdata is documented in Section 3.3 and Section 3.4, respectively. By building a correspondence between these model coefficients and the underlying microdata, we can connect industry-level shocks with household-level outcomes through both the price and income channels.

4.3 Deriving a synthetic CPI

To link industry prices with household purchasing power, we exploit the price model in Equation 6 together with the distributional structure of the $\mathbf{S}$ matrix. Let $\mathbf{p}$ denote the base-year price vector satisfying $\mathbf{p} = \mathbf{v}_c'(\mathbf{I} - \mathbf{A})^{-1}$. A shock to any component of the value-added vector–for instance, a uniform increase in labour compensation for the bottom three deciles–can be represented by a diagonal matrix of scaling factors $\Gamma = \mathrm{diag}(\gamma_1, \dots, \gamma_g)$ applied to the corresponding rows of $\mathbf{V}$. The resulting post-shock price vector is:

$$\tilde{\mathbf{p}} = \mathbf{i}'\Gamma\mathbf{V}(\mathbf{I} - \mathbf{A})^{-1} \tag{10}$$

where $\gamma_d = 1$ for unaffected groups and $\gamma_d > 1$ for those receiving the increase. A two-industry, two-group numerical example with a 50% shock to the first group illustrates the mechanism:

$$\tilde{\mathbf{p}} = \begin{bmatrix}1.5 & 1\end{bmatrix} \begin{bmatrix}0.20 & 0.35 \\ 0.45 & 0.15\end{bmatrix} \begin{bmatrix}1.20 & 0.35 \\ 0.25 & 1.15\end{bmatrix}$$

Note that the shock applies to only one of the two groups, but the formulation accommodates any combination of heterogeneous shocks across groups and income categories (e.g., profits).

With the post-shock $\tilde{\mathbf{p}}$ and baseline $\mathbf{p}$ price vectors, we derive the implied nominal change in the household final consumption expenditure vector $\mathbf{c}$ at the industry level:

$$\tilde{\mathbf{c}} = (\tilde{\mathbf{p}} - \mathbf{p}) \cdot \mathbf{c} \tag{11}$$

To express the price changes in terms that match actual consumer expenditure categories, we reverse the three-step transformation described in Section 3.3. First, we reclassify $\tilde{\mathbf{c}}$ from industries to products using the commodity-technology Model B transformation matrix $\mathbf{D}' = \widehat{g}^{-1}\mathbf{V}'$, where $\widehat{g}^{-1}$ is the inverse of the diagonalised vector of industry outputs and $\mathbf{V}'$ the transpose of the supply matrix (Eurostat, 2008, p. 349). Second, we apply the valuation algorithm $\mathbf{F}(\cdot)$ to move from basic to purchaser's prices (see Section 3.3.2 and Appendix A). Third, we apply the bridge matrix $\mathbf{B}$ to convert from CPA products to COICOP consumption purposes. In compact notation:

$$\tilde{\mathbf{c}}_{coicop} = \mathbf{F}\big(\mathbf{D}' \cdot \tilde{\mathbf{c}}\big) \cdot \mathbf{B} \tag{12}$$

At this stage, weighting the COICOP expenditure changes by the appropriate consumption shares ω_i yields a household-specific synthetic CPI:

$$CPI_h = \sum_{i=1}^{N} \omega_{h,i} \cdot \tilde{c}_{h,i} \tag{13}$$

where the subscript h indexes individual households (or household groups). Using this scalar, we can deflate nominal equivalised disposable income to obtain real income measures and compute, for instance, the corresponding Gini coefficients of the real income and expenditure distributions. To capture the full personal distribution effect, one would additionally allocate the increased labour incomes $\mathbf{\Gamma V}$ to the corresponding wage-earning households. Although the possibilities of this methodology are several, in Section 5.3 we illustrate it by studying the distributional impact of industry-specific price shocks in Spain.

5 Empirical results

This section presents a few stylized facts on incomes and prices and two sets of applications of the Spanish estimation of the distributional input-output tables. Section 5.1 provides a descriptive introduction to the accounting relationship between prices and incomes in the context of shock propagation. Section 5.2 presents first-order analysis of industries' inflationary potential, which includes the identification of possible inflation channels via principal components analysis (PCA). Section 5.3 delves into the distributional impact of the shocks analyzed in the previous section, with special attention to how inflation of necessities affect the bottom of the distribution of household resources.

5.1 From functional to personal real distribution

This section illustrates the relationship between factor revenues, prices and real purchaising power by presenting the macroeconomic evolution of the gross value added (GVA) deflator, functional distribution, and the Harmonised Indices of Consumer Prices (HICP) together with budget shares for the main expenditure groups. Through these data, the section shows that post-2020 inflation surge has been concentrated in a handful of systemically significant industries, has coincided with a divergence between profits and labour compensation, and has propagated to consumer prices through expenditure groups whose share of household budgets is highly skewed along the income distribution.

Figure 7 decomposes the growth of the GVA deflator and of nominal GVA itself, both by 11-industry NACE grouping and by factor of income, over 2010–2024. Three regularities stand out. First, aggregate deflator growth was muted—around 0.4% on average from 2010 to 2019—before the post-pandemic surge to an annual average of 5.5% in 2021–2023. Second, the surge was highly concentrated: in 2021 alone, the top three contributing industries—Mining & utilities (BDE), Distribution (GHI), and Public services (OPQ)—accounted for the bulk of aggregate price growth, a pattern consistent with the cost-push transmission via systemically significant sectors documented by Weber and Wasner (2023). Third, the factor-wise decomposition (panel c) reveals a clear post-2020 divergence: in 2022 gross operating surplus expanded by 9.3%, well ahead of wages and salaries. Combined with the industry concentration of the deflator surge, this functional pattern indicates that the inflationary episode was disproportionately driven by profits in a small set of industries that simultaneously dominate the deflator and the nominal GVA decomposition.

The functional surge translated into consumer prices through specific COICOP groups. Figure 8 (a) shows that, since 2021, the three most cumulative contributors to HICP growth have been food & beverages, alcohol & tobacco, and restaurants & hotels—precisely the categories most directly tied to the industries identified above. Importantly, the budget incidence of these categories is far from neutral: as Figure 8 (b) shows, food and non-alcoholic beverages absorb 25.3% of total expenditure for the bottom quintile in 2023, against 9.6% for the top quintile—a near three-fold gap consistent with Engel's Law (Houthakker, 1957; Chai and Moneta, 2010; Kaus, 2013). The opposite gradient holds for restaurants and hotels (8.6% in Q1 versus 23.9% in Q5), implying that the distributional sign of an industry-specific shock depends jointly on which COICOP categories absorb it and how those cate-

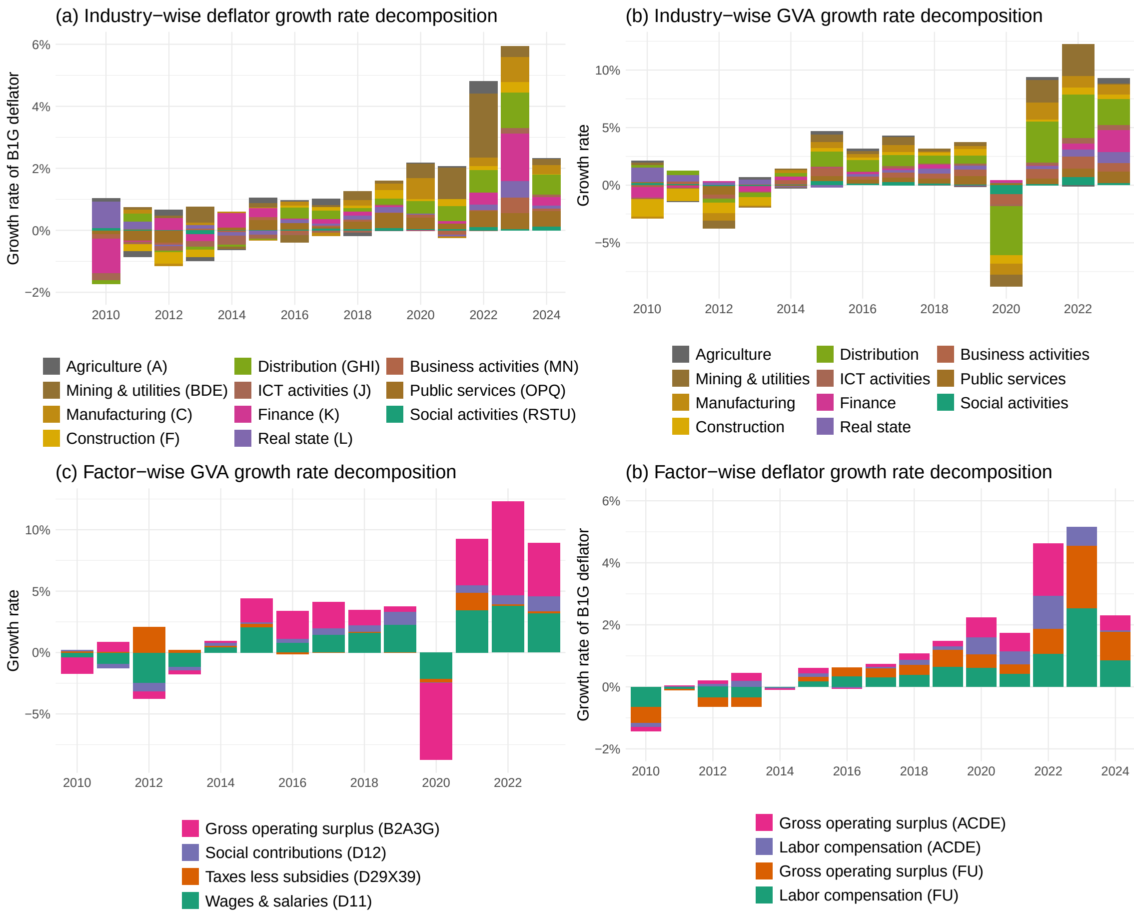


Figure 7: Decomposition of the growth rate of the gross value added and its deflator. *Source*: Author's elaboration using Eurostat data.

gories rank in the household budget. This is precisely the joint pattern of industry concentration, factor divergence, and skewed consumer exposure that the cost-push simulations in Section 5.2 and the household-level decomposition in Section 5.3 are designed to quantify.

5.2 Inflationary potential by industry shock

Figure 9 show the inflation rate resulting from a 50% shock increase to the gross operating surplus of one industry at a time. The top five sectors by the size of their impact in 2023 were accommodation (3.6%), real estate activities (3.5%), crop & animal production (2.2%), financial services (2.1%), and electricity (1.8%). While this implies shock sizes in proportion to the profit mass of each industry, if the magnitud of the shock corresponds to a 50% increase in the average profit mass, we obtain very different results. We find that mining & quarrying (12%), other personal services (3%), fishing (3%), electronics repair (2.8%), and water transport (2.6%) are the ones with the largest impact potential.

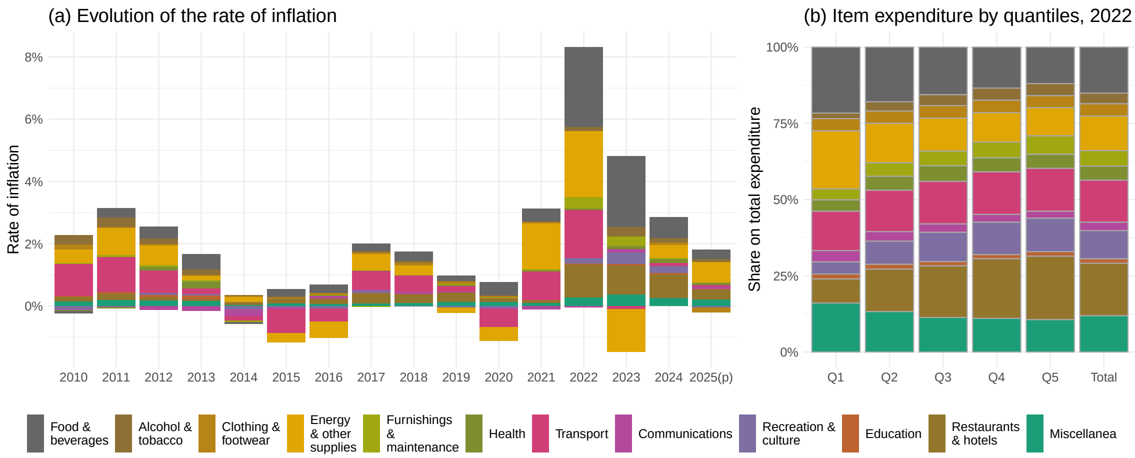


Figure 8: Inflation decomposition by consumption purpose. *Note*: Data for 2025 is a monthly average up to May. *Source*: Author's elaboration with Eurostat data.

A majority of industries show no changes across the years, with a few exceptions. The largest effect reductions arise in telecommunications (-0.6pp), textile products (-0.5pp), and crop & animal production (-0.4pp). Conversely, some modest increases in inflation potential can be found in financial services (1.4pp), mining & quarrying (0.9pp), and accommodation (0.7pp). This corresponds to a proportional shock to each industry's profits. Alternatively, if we try to normalize the shock by allocating a 50% increase to the average profit mass, we find much larger decreases in textile products (-2.1pp), electronics repair (-1.5pp), and coke & refined petroleum (-1pp) on the lower bound, and larger increases textile products (-2.1pp), electronics repair (-1.5pp), and coke & refined petroleum (-1pp) on the upper bound. This difference between shock implementations puts the emphasis on the weight of each industry on total value added. Larger industries with lower profit shares could have the same impact potential than smaller ones with higher shares. Given that we care most about the absolute inflation potential, a proportional increase in the income component is preferable.

While we have obtain a ranking of industries by the inflationary potential, the analysis of the channels through which the CPI is affected admits futher disaggregation. Instead of deriving a synthetic CPI using average consumption shares as weights, we can try to find patterns of activation of 3-digit consumption purposes across industries. To synthesize such information, we can apply principal components analysis, as in Figure 10.[3] For our case, PCA identifies the largest cluster of related variables (i.e., 3-digit COICOP items) across industries. The first PC explains 18.2% and the

[3]PCA is an unsupervised learning method of dimensionality reduction similar to clustering. In essence, it is used to break complex data into its fundamental patterns by thining out less relevant features (Greenacre et al., 2022; Lever et al., 2017). Each principal component (PC) consists of the geometrical projection of the data onto a lower-dimensional vector. There are as many PCs as variables we find in the data. The first PC provides the best summary of the relationship among the data while retaining the most variability. Each subsequent PC will provide a second- and third-best fit to the remaining data variability, and so on until all variables and variance are exhausted.

Figure 9: Top 30 industries by the estimated inflation rate resulting from a 50% shock increase of the gross operating surplus of each industry in Spain. Source: Author's elaboration

first six PCs 59.7% of the total variability. Thus, focusing on the first six PCs seems reasonable to identify the strongest patterns.

While Figure 10 (a) indicates how strong is the correlation of each original 2-digit item with the first two PCs by means of the direction and length of the arrows, Figure 10 (b) replicates the analysis for a 3-digit COICOP breakdown. The gradient

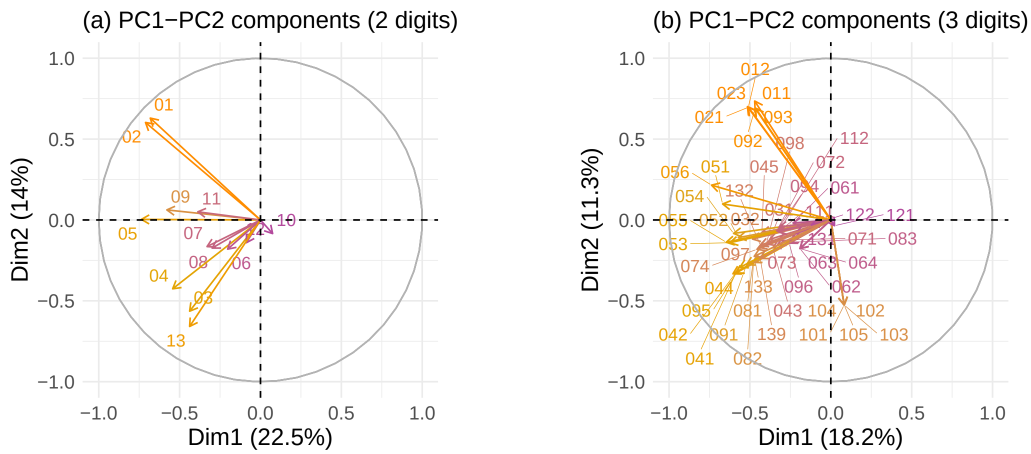


Figure 10: Representation of the first two principal components for 2- and 3-digit COICOP item breakdown. *Note*: The complete results for 3-digit COICOP items can be consulted in the appendix. Source: Author's elaboration with Eurostat data.

reflects the quality of the representation, with warmer colors indicating a weaker one. At two digits, the chart indicates that there are three consistent clusters. One highly populated group encompassing almost all items, and another two smaller ones. While the first smaller cluster gathers furnishings and home appliances, recreation, and restaurants, the second one collects a bundle of which several headings can be considered essential: food and beverages (01), alcohol & tobacco (02), furnishings & appliances (05), recreation & culture (09), and restaurants and hotels (11). At 3 digits the stark separation thins down, with the relevant cluster including a variaty of headings: food (011), beverages (012), alcoholic beverages (021), tobacco (023), furniture (051), home maintenance purchases (056), other recreational goods (092), and garden products and pets (093).

Figure 11 protrays with greater clarity the cluster of expenditure groups. In general, a persistent association between food and beverages and alcohol and tobacco across vectors seems to exist. The first two PCs showed that this group extends to some recreation good and services and homecare, which is likewise observable from this perspective. As for the rest, the main associations are in recreational and hospitality services, transport related purchases, homecare goods, and other personal expenditures.

5.3 Distributional consequences of local industry price shocks

The aggregate CPI computed in Section 6.1 is a population-weighted average that conceals systematic differences in consumption structure across the income distribution (Gürer and Weichenrieder, 2020; Jaravel, 2021). Applying the 3-digit COICOP price indices derived from the Leontief price model to the matched household mi-

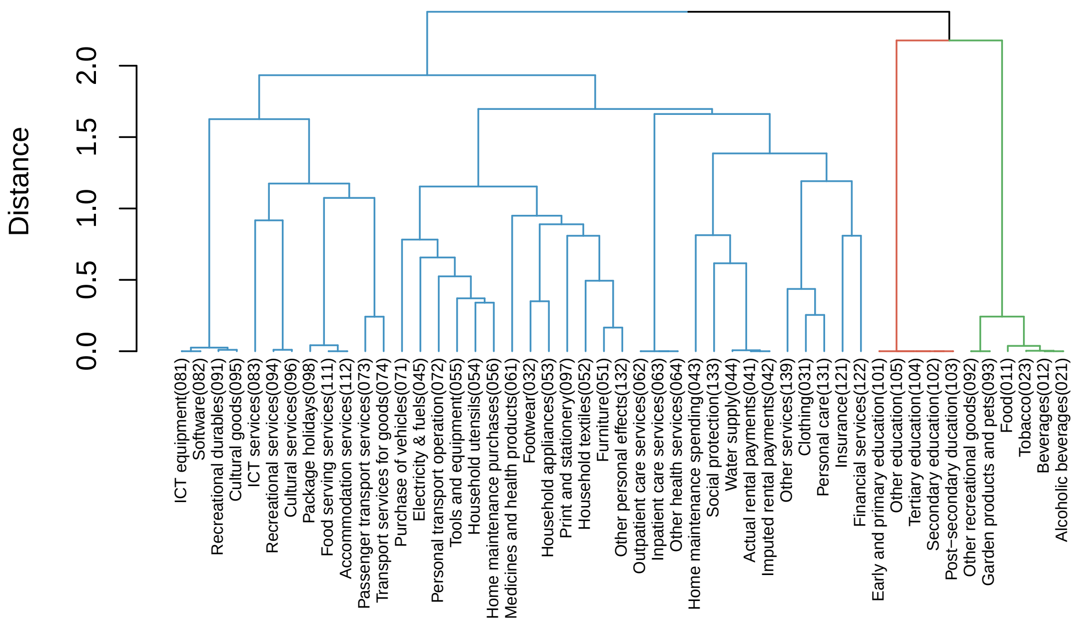


Figure 11: Hierarchical representation of item association. Source: Author's elaboration with Eurostat data.

crodata, we construct a household-specific synthetic price index CPI_h and deflate nominal equivalised disposable income accordingly, yielding the full distribution of real income losses for each industry shock. As suggested by Figure 10 (b), the PCA of COICOP-level transmission patterns already identifies a necessity cluster; the household-level analysis confirms that this structural feature translates into a systematic distributional bias. In 2023, the industries with the largest estimated increase in the real income Gini coefficient under the price-only scenario are crop & animal production (0.1pp), financial services (0.1pp), real estate activities (0.1pp), food products (0.1pp), and electricity (0.1pp). Conversely, shocks most conducive to inequality reduction originate in accommodation (-0.2pp), land transport (0pp), and motor vehicle equipment (0pp).

A more granular characterisation of the asymmetry is provided by the inflation inequality gap $\Delta_{D1,D10} = CPI_{D1} - CPI_{D10}$, measuring whether bottom-decile households face a systematically larger price increase than top-decile ones for each industry shock. Figure 12 ranks all 64 NACE industries by this scalar: positive values indicate a regressive transmission pattern, negative values a progressive one. The most regressive shocks originate predominantly in industries that constitute the necessity cluster of Section 6.1—food processing, agriculture, electricity generation, and real estate—consistent with well-documented European evidence (Bettarelli et al., 2023; Menyhért, 2023; Ferreira et al., 2024). A smaller set of industries, chiefly accommodation and food services and recreational activities, exhibits a negative gap, reflecting the higher expenditure shares that affluent households allocate to those

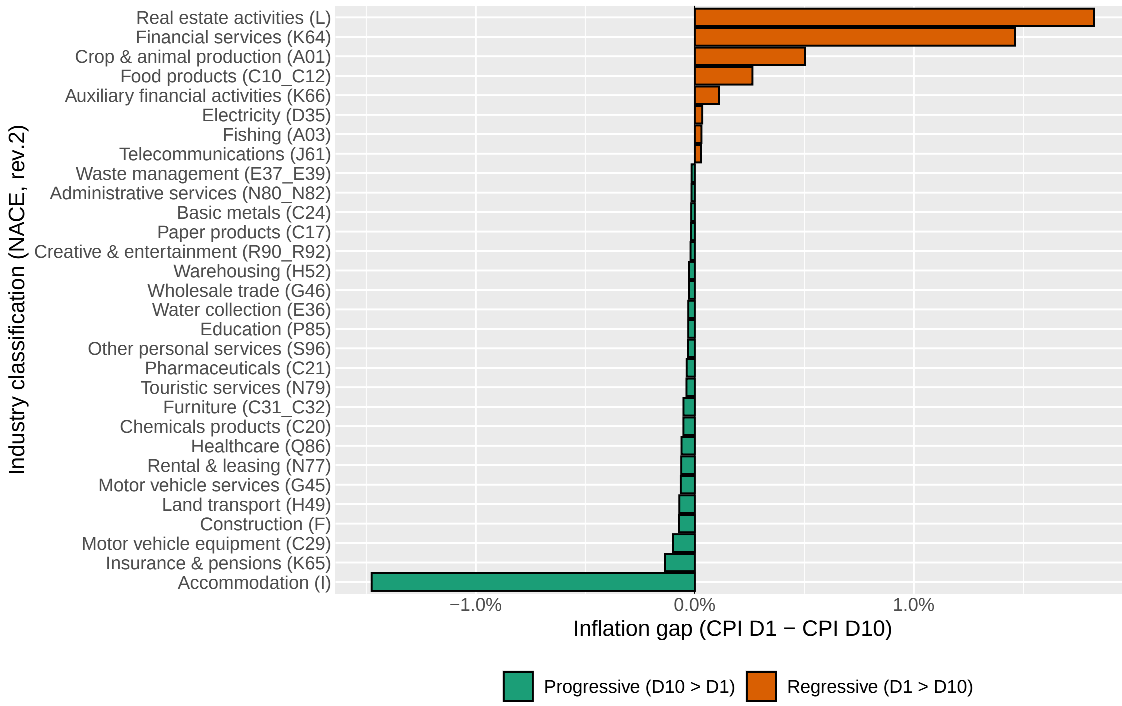


Figure 12: Inflation inequality gap ($CPI_{D1} - CPI_{D10}$) for all 64 NACE industries following a 50% shock to gross operating surplus, Spain, 2023. Orange bars denote regressive outcomes ($CPI_{D1} > CPI_{D10}$); green bars denote progressive ones. *Source*: Authors' own elaboration.

categories.

The distributional extended IOT also allows activation of the income channel through the V matrix. For a given shock to the gross operating surplus of industry j, workers in that sector receive a proportional increment to labour compensation $\gamma \mathbf{V}_j$ alongside the expenditure-side price effect. Figure 13 decomposes the total change in equivalised real income into three components across income deciles for a selection of key industries: (i) the *price channel*—the reduction in purchasing power due to higher consumer prices; (ii) the *income channel*—the wage gain accruing exclusively to households whose reference person is employed in the shocked industry; and (iii) the *net effect*. The decomposition reveals a structural misalignment: for food and energy industries, the price burden is monotonically regressive, depressing real incomes most in D1 and D2, whereas the income gain concentrates in the middle deciles where worker coverage is highest. The income channel therefore provides limited relief to the poorest households, and the net effect remains regressive for necessity-goods industries (Lara Jauregui et al., 2025; Weber et al., 2024a).

Figure 14 synthesises the two-channel interaction by presenting the change in the Gini coefficient under the price-only and the net (price plus income) scenarios for the 20 industries with the largest absolute distributional impact. In every case the net Gini change is smaller in absolute value than the price-only counterpart, confirming

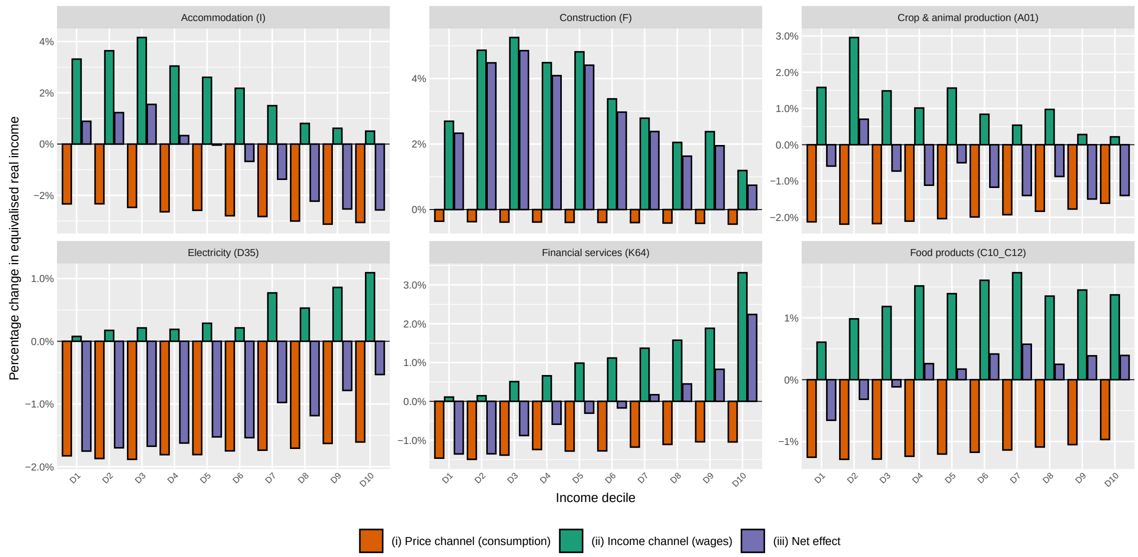


Figure 13: Decomposition of the real income change into the price channel (i), the income channel (ii), and the net effect (iii), by income decile, for selected NACE industries. 50% shock to gross operating surplus, Spain, 2023. *Source*: Authors' own elaboration.

that exclusive focus on the expenditure channel overstates the disequalising effect of a profit shock. For hospitality and recreation industries the attenuation is more pronounced, as workers tend to be distributed across the lower and middle deciles and their wage gain partially offsets the inequality contribution of higher prices. For food and energy industries, by contrast, the income and price channels remain structurally misaligned: the price burden falls most on D1 while the wage gain accrues to middle-decile households, reinforcing the case for targeted fiscal measures when shocks originate in necessity-goods sectors.

6 Implications and conclusions

This paper interpellates the G-20 Data Gaps Initiative III.8 recommendation by developing a replicable methodology for compiling distributionally extended input-output tables (DE-IOTs) that simultaneously integrate the income and expenditure dimensions of the household sector into the interindustry system. Applied to the Spanish economy over 2010–2023, the resulting dataset enables the tracing of industry-level price shocks through the full circular flow of income—from functional income generation in production to the personal distribution of real purchasing power across households. This final section distils the methodological and policy implications of the analysis.

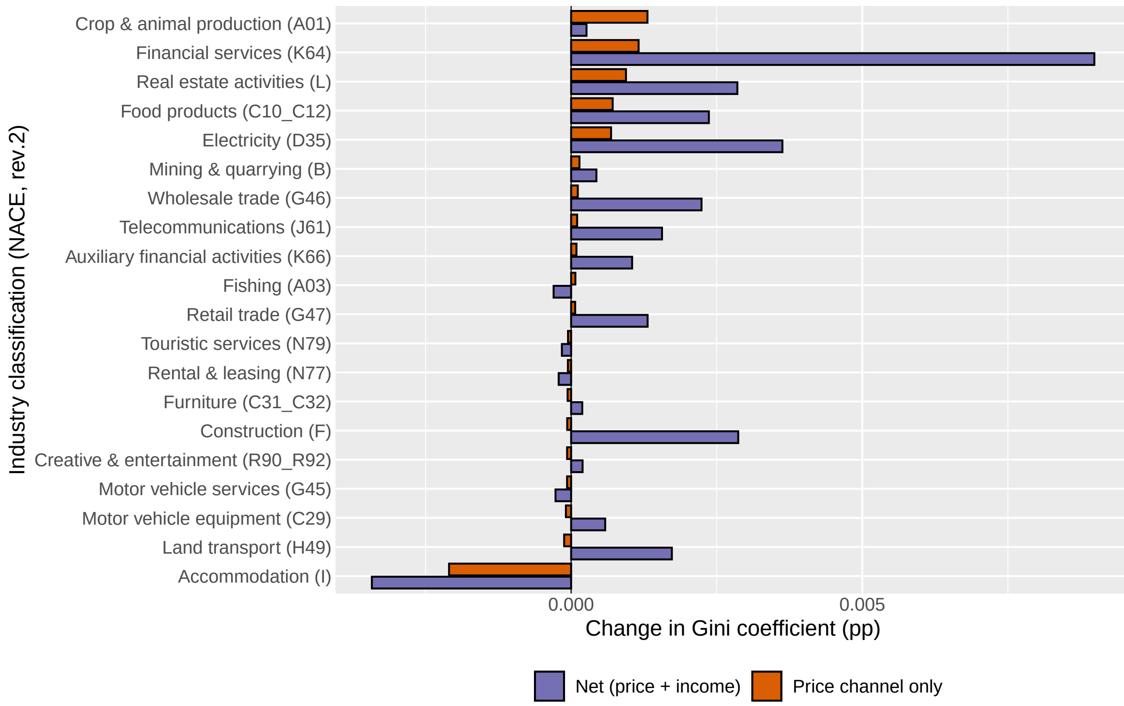


Figure 14: Change in the Gini coefficient of equivalised real household income under the price-only and net (price plus income) scenarios, for the 20 industries with the largest absolute distributional impact. 50% shock to gross operating surplus, Spain, 2023. Orange bars: price channel only; purple bars: net effect. *Source*: Authors' own elaboration.

6.1 Methodological implications

The first contribution of the paper is procedural. Building on the Miyazawa interrelational multiplier tradition (Miyazawa, 1976a,b) and recent advances in distributional national accounts (Coli et al., 2022; OECD, 2024; Piketty et al., 2018), the paper offers a step-by-step compilation guide that connects three hitherto separate data infrastructures—FIGARO symmetric IOTs, ES-SILC, and ES-HBS—through statistical matching, grossing up to national accounts totals, and bidirectional linkage via the **C** and **V** matrices. Although the SAM tradition has long advocated for household disaggregation (Pyatt and Thorbecke, 1976; Pyatt and Round, 1985; Stone, 1985), most empirical implementations have been confined to the expenditure channel alone (Mongelli et al., 2010; Cazcarro et al., 2022). By operationalising the income channel through the **V** matrix, the DE-IO framework renders the interaction between functional and personal income distribution directly observable at the household level, a feature whose analytical value is demonstrated by the two-channel decomposition presented in Section 5.3.

A second methodological implication concerns the derivation of household-specific consumer price indices from the Leontief price model. Rather than relying on aggregate CPI weights, the combination of 3-digit COICOP bridge matrices with household-level expenditure shares yields a price index CPI_h that captures the heterogeneous exposure of different income groups to interindustry price transmission. The PCA of COICOP-level transmission patterns further reveals that the price effects of sectoral shocks are not randomly distributed across consumption categories but cluster along identifiable channels, the most prominent of which loads heavily on necessity goods—food, energy, and housing. This finding reinforces, from the supply side, the demand-side evidence on inflation inequality documented in the literature (Jaravel, 2021; Bettarelli et al., 2023; Menyhért, 2023; Ferreira et al., 2024).

Third, the analysis confirms that summary inequality measures must be chosen with care. The Gini coefficient, while responding to industry-level shocks, understates the severity of tail effects. The inflation inequality gap $\Delta_{D1,D10}$ and decile-specific real income losses reveal a distributional incidence that the Gini alone would mask, consistent with the wider critique of centre-sensitive measures in contexts of price-driven welfare erosion (Pallotti et al., 2023; Amores et al., 2024).

More broadly, the DE-IO methodology is portable. Because it relies on harmonised data sources—FIGARO tables are available for all EU member states and major trading partners, while SILC and HBS microdata follow Eurostat regulations—the compilation procedure can be replicated across the European Statistical System and, with suitable adaptations, for OECD and G-20 economies. The provision of a complete 2010–2023 panel as supplementary material is intended to facilitate replication and extension by the research community.

6.2 Policy implications

The empirical application to the Spanish economy yields three policy-relevant findings. First, the inflationary potential of industry shocks is highly concentrated: a small number of sectors—real estate, food processing, agriculture, electricity, and financial services—account for the bulk of the aggregate CPI response and, critically, for the most regressive distributional outcomes. Early monitoring of cost pressures in these sectors could provide actionable lead indicators for stabilisation policy, contributing to the *mesoprudential* framework advocated in the paper's analytical design.

Second, the two-channel decomposition of real income changes demonstrates that the price and income effects of a given shock are structurally misaligned for precisely those industries that matter most for the poor. When food or energy prices rise, the resulting purchasing power loss falls disproportionately on the bottom deciles, whereas the associated wage gains accrue to middle-income households employed in those sectors. This misalignment implies that market-mediated compensation through the income channel is insufficient to offset regressive price effects in necessity-goods industries (Lara Jauregui et al., 2025; Weber et al., 2024a). The case for targeted fiscal transfers or price stabilisation mechanisms is therefore strengthened on structural rather than merely conjunctural grounds.

Third, the sign variability of distributional effects across industries challenges the assumption that all inflationary episodes are uniformly regressive. Shocks originating in accommodation, recreational services, or motor vehicle activities reduce measured inequality, reflecting the concentration of expenditure on these categories among affluent households. This heterogeneity suggests that the net distributional impact of a multi-sectoral inflationary episode—such as the post-2021 European price surge—depends on the composition of the shock vector, an insight that aggregate inflation targeting alone cannot capture (Van'T Klooster and Weber, 2024; Fahr et al., 2024).

6.3 Limitations and avenues for future research

The paper is subject to several limitations that I aim to tackle in future developments of this research line. First, the absence of uncertainty bands around the estimates is a shared weakness of the distributional national accounts literature, for which no consensus methodology yet exists (Coli et al., 2022; OECD, 2024). Sensitivity analysis with respect to the gap allocation procedure, particularly for low-coverage expenditure items, would strengthen the robustness of the results. Second, the proportional shock calibration employed in Section 5 remains arbitrary; historically grounded shock magnitudes, calibrated to observed episodes such as the 2021–2023 inflationary surge, would improve the external validity of the simulations. Third, the time-series dimension of the 2010–2023 panel remains underexploited: structural decomposition analysis tracing the evolution of distributional sensitivities over the business cycle constitutes a promising extension.

In sum, the distributional extension of the input-output framework offers a versatile infrastructure for analysing the heterogeneous interaction between industries and households. A large body of existing structural decomposition, multiplier, and impact analyses in the input-output tradition is suitable for this extension, gaining an additional layer of distributional detail (Steenge and Serrano, 2012; Alberti et al., 2023). By rendering the functional-to-personal distribution connection empirically tractable, DE-IOA contributes to the broader effort of integrating household heterogeneity into macroeconomic accounting and asserts the relevance of input-output analysis for the study of inequality and living standards.

Data Availability

The FIGARO supply, use, and input–output tables are publicly available from Eurostat. The Spanish waves of the Household Budget Survey (HBS) and the Survey on Income and Living Conditions (SILC) microdata used to build the distributionally extended input–output tables are publicly available from the Instituto Nacional de Estadística. The full 2010–2023 series of distributionally extended input–output tables for the Spanish economy, together with the replication code, is provided as SUPPLEMENTARY MATERIAL to this paper.

Acknowledgments

I would like to express gratitute to Henning Anhert, Pierre Sola, Nina Blatnik, and Marco Felici for the generous training they provided on the principles of distributional (wealth) accounts during my time at the DWA team of the ECB. Likewise, an early instance of this research piece was presented at the 10th Permanent Workshop of the Hispanic-American Input-Output Society and the 2024 YSI Madrid meeting, benefitting from comments made from the audience. All remaining errors are my own.

Disclaimer

To the best of my knowledge, I have no conflicts of interest to disclose.

References

Akoğuz, E. C., Capéau, B., Decoster, A., De Sadeleer, L., Güner, D., Manios, K., Paulus, A., and Vanheukelom, T. (2020). A new indirect tax tool for EUROMOD: Final Report. Technical Report JRC/SVQ/2018/B.2/0021/OC, KU Leuven, Department of Economics.

Alberti, T. M., De Souza, K., and Porsse, A. (2023). Poverty and the functional distribution of income in the input–output framework: in pursuit of strategies for inclusive growth. *Economic Systems Research*, 35(4):614–633.

Amores, A. F., Basso, H., Bischl, J. S., De Agostini, P., De Poli, S., Dicarlo, E., Flevotomou, M., Freier, M., Maier, S., García-Miralles, E., Pidkuyko, M., Ricci, M., and Riscado, S. (2024). Inflation, fiscal policy and inequality. The distributional impact of fiscal measures to compensate for consumer inflation. Documentos Ocasionales 2418, Banco de España, Madrid. Series: Documentos Ocasionales.

Auer, R., Pedemonte, M., and Schoenle, R. (2024). Sixty years of global inflation: a post-GFC update. Working Paper 1189, Bank for International Settlements, Basel.

Balestra, C. and Oehler, F. (2023). Measuring the joint distribution of household income, consumption and wealth at the micro level. Methodological issues and experimental results. Working Paper, European Union & OECD, Luxembourg. Series: OECD Papers on Well-being and Inequalities Volume: 11.

Batey, P. W. J. (1985). Input-output models for regional demographic-economic analysis: some structural comparisons. *Environment and Planning A*, 17:73–99.

Batey, P. W. J. and Madden, M. (1983). The modelling of demographic-economic change within the context of regional decline: Analytical procedures and empirical results. *Socio-Economic Planning Sciences*, 17(5-6):315–328.

Batey, P. W. J., Madden, M., and Weeks, M. J. (1987). Household Income and Expenditure in Extended Input-Output Models: A Comparative Theoretical and Empirical Analysis. *Journal of Regional Science*, 27(3):341–356.

Batey, P. W. J. and Rose, A. Z. (1990). Extended Input-Output Models: Progress and Potential. *International Regional Science Review*, 13(1-2):27–49.

Bańbura, M., Bobeica, E., and Martínez Hernández, C. (2023). What drives core inflation?: the role of supply shocks. Working Paper 2875, European Central Bank, Frankfurt am Main.

Bernanke, B. and Blanchard, O. (2024). An analysis of pandemic-era inflation in 11 economies. Working Paper 24-11, Peterson Institute for International Economics.

Bettarelli, L., Estefania-Flores, J., Furceri, D., Loungani, P., and Pizzuto, P. (2023). Energy inflation and consumption inequality. *Energy Economics*, 124:106823.

Boissay, F., Collard, F., Toulouse School of Economics (CNRS), Manea, C., Bank for International Settlements, and Shapiro, A. (2023). Monetary Tightening, Inflation Drivers and Financial Stress. Working Paper 1155, Bank for International Settlements, Basel.

Capéau, B., Decoster, A., and Güner, D. (2022). Extension of the EUROMOD-ITT tool. Technical Report JRC/SVQ/2020/OP/1373, KU Leuven, Department of Economics.

Cazcarro, I., Amores, A. F., Arto, I., and Kratena, K. (2022). Linking Multisectoral Economic Models and Consumption Surveys for the European Union. *Economic Systems Research*, 34(1):22–40.

Chai, A. and Moneta, A. (2010). Retrospectives: Engel Curves. *Journal of Economic Perspectives*, 24(1):225–240.

Chai, A. and Moneta, A. (2013). Back to Engel? Some Evidence for the Hierarchy of Needs. In Pyka, A. and Andersen, E. S., editors, *Long Term Economic Development*, pages 33–59. Springer Berlin Heidelberg, Berlin, Heidelberg.

Chen, S., Tsang, E., and Zhang, L. S. Y. (2023). Global supply chain interdependence and shock amplification – evidence from Covid lockdowns. Working Paper 1123, Bank for International Settlements, Basel.

Claeys, G., Guetta-Jeanrenaud, L., McCaffrey, C., and Welsslau, L. (2024). Inflation inequality in the European Union and its drivers. Technical report, Bruegel.

Coli, A., Istatkov, R., Jayyousi, H., Oehler, F., and Tsigkas, O. (2022). *Distributional National Account Estimates for Household Income and Consumption: Methodological Issues and Experimental Results.* Publications Office of the European Union, Luxembourg, 2022 edition edition.

De Santis, R. A. (2024). Supply Chain Disruption and Energy Supply Shocks: Impact on Euro Area Output and Prices. Working Paper 2024/2884, Frankfurt am Main, European Central Bank.

Defourny, J. and Thorbecke, E. (1984). Structural Path Analysis and Multiplier Decomposition within a Social Accounting Matrix Framework. *The Economic Journal*, 94(373):111–136.

DeLong, B. (2023). The first inflation problem of the twenty-first century. *Review of Keynesian Economics*, 11(2):117–128.

D'Orazio, M., Di Zio, M., and Scanu, M. (2006). *Statistical matching: Theory and practice.* John Wiley & Sons, Chichester.

dos Santos, P. L. (2017). The Principle of Social Scaling. *Complexity*, 2017:1–9.

Dreoni, I., Serruys, H., Manso, L., Tudo Ramirez, J., and Amores, A. F. (2025). Statistical imputation and validation of consumption microdata for EUROMOD. EUROMOD Working Paper EM 05/25, Joint Research Centre, European Commission.

Eickmeier, S. and Hofmann, B. (2022). What drives inflation? Disentangling demand and supply factors. Working Paper 1047, Bank for International Settlements, Basel.

Eurostat (2008). *Eurostat Manual of Supply, Use and Input-Output Tables*. Publications Office of the European Union, Luxembourg.

Eurostat (2013). *European System of Accounts 2010*. Publications Office of the European Union, Luxembourg.

Fahr, S., Senner, R., and Vismara, A. (2024). The globalization of climate change: amplification of climate related physical risks through input output linkages. Working Paper 2942, European Central Bank, Frankfurt am Main.

Feenstra, R. C. and Sasahara, A. (2018). The 'China shock,' exports and U.S. employment: A global input–output analysis. *Review of International Economics*, 26(5):1053–1083. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/roie.12370.

Ferreira, C., Leiva, J. M., Nuño, G., Ortiz, ., Rodrigo, T., and Vazquez, S. (2024). The heterogeneous impact of inflation on households' balance sheets. Documentos de Trabajo 2403, Banco de España, Madrid. Series: Documentos de Trabajo.

Galbraith, J. K. (2023). The quasi-inflation of 2021–2022: a case of bad analysis and worse response. *Review of Keynesian Economics*, 11(2):172–182.

Golan, A. (2017). *Foundations of Info-Metrics: Modeling and Inference with Imperfect Information*. Oxford University Press, New York.

Greenacre, M., Groenen, P. J. F., Hastie, T., D'Enza, A. I., Markos, A., and Tuzhilina, E. (2022). Principal component analysis. *Nature Reviews Methods Primers*, 2(1):100. Publisher: Nature Publishing Group.

Gürer, E. and Weichenrieder, A. J. (2020). Multidimensional inequality across Europe: The role of differential inflation. *Review of Income and Wealth*, 66(4):869–896.

Honkkila, J. and Kavonius, I. K. (2013). Micro and Macro Analysis on Household Income, Wealth and Saving in the Euro Area. Working Paper 1619, European Central Bank, Frankfurt am Main.

Houthakker, H. S. (1957). An International Comparison of Household Expenditure Patterns, Commemorating the Centenary of Engel's Law. *Econometrica*, 25(4):532.

Hu, J., Pedauga, L., Rocchi, P., Reynès, F., and Rueda-Cantuche, J. M. (2020). FIDELIO: four modules linking input-output and general equilibrium modelling. Technical report, European Commission, Joint Research Centre.

International Monetary Fund (2022). *Countering the cost-of-living crisis*. Number October in World economic outlook / International Monetary Fund. International Monetary Fund, Washington, DC.

Jaravel, X. (2021). Inflation Inequality: Measurement, Causes, and Policy Implications. *Annual Review of Economics*, 13(1):599–629.

Katris, A., Figus, G., and Turner, K. (2017). Disaggregation of the 2010 UK Social Accounting Matrix to report household income quintiles. Working Paper 1708, University of Strathclyde Business School, Department of Economics.

Kaus, W. (2013). Beyond Engel's law - A cross-country analysis. *The Journal of Socio-Economics*, 47:118–134.

Kim, J. H. and Hewings, G. J. D. (2012). Integrating the fragmented regional and subregional socioeconomic forecasting and analysis: a spatial regional econometric input–output framework. *The Annals of Regional Science*, 49(2):485–513.

Kim, K., Kratena, K., and Hewings, G. J. (2015). The Extended Econometric Input–Output Model with Heterogeneous Household Demand System. *Economic Systems Research*, 27(2):257–285.

Kim, T.-J. and Hewings, G. J. D. (2015). Aging Population in a Regional Economy: Addressing Household Heterogeneity with a Focus on Migration Status and Investment in Human Capital. *International Regional Science Review*, 38(3):264–291.

Kimura, Y. and Kondo, H. (1999). Revisiting Miyazawa on Multisectoral Multipliers. In Hewings, G. J. D., Sonis, M., Madden, M., and Kimura, Y., editors, *Understanding and Interpreting Economic Structure*, pages 53–71. Springer Berlin Heidelberg, Berlin, Heidelberg. Series Title: Advances in Spatial Science.

Kirman, A. P. (1992). Whom or What Does the Representative Individual Represent? *Journal of Economic Perspectives*, 6(2):117–136.

Kiss, R. and Strasser, G. (2024). Different household – different inflation rate.

Kröger, M., Longmuir, M., Neuhoff, K., and Schütze, F. (2023). The price of natural gas dependency: Price shocks, inequality, and public policy. *Energy Policy*, 175:113472.

Lara Jauregui, J., Weber, I. M., Nassif Pires, L., and Teixeira, L. (2025). Price shocks are redistribution shocks: Systemically significant prices for inequality in the united states. Working Paper, University of Massachusetts Amherst.

Leontief, W. (1986). *Input-output economics*. 2nd Edition, New York, NY, Oxford University Press.

Lequiller, F. and Blades, D. (2014). *Understanding National Accounts: Second Edition*. OECD Publishing.

Lever, J., Krzywinski, M., and Altman, N. (2017). Principal component analysis. *Nature Methods*, 14(7):641–642.

Lewbel, A. and Houthakker, H. S. (2018). Engel curve. In *The New Palgrave Dictionary of Economics*, pages 3717–3719. Palgrave Macmillan, London, third edition.

López-Laborda, J., Marín-González, C., and Onrubia-Fernández, J. (2021). Estimating Engel curves: a new way to improve the SILC-HBS matching process using GLM methods. *Journal of Applied Statistics*, 48(16):3233–3250.

Menyhért, B. (2023). Inflation and its diverse social consequences across the euro area. *Quarterly Report on the Euro Area (QREA)*, 22:7–16.

Milanović, B. (2019). *Capitalism, Alone. The Future of the System That Rules the World*. The Belknap Press of Harvard University Press, Cambridge.

Miller, R. E. and Blair, P. D. (2022). *Input-Output Analysis. Foundations and Extensions*. Cambridge University Press, New York, 3rd edition edition.

Miyazawa, K. (1976a). Input-Output Analysis and Interrelational Income Multiplier as a Matrix. In Beckmann, M. and Künzi, H. P., editors, *Input-Output Analysis and the Structure of Income Distribution*, volume 116, pages 22–42. Springer Berlin Heidelberg, Berlin, Heidelberg. Series Title: Lecture Notes in Economics and Mathematical Systems.

Miyazawa, K. (1976b). *Input-Output Analysis and the Structure of Income Distribution*, volume 116 of *Lecture Notes in Economics and Mathematical Systems*. Springer Berlin Heidelberg, Berlin, Heidelberg.

Mongelli, I., Neuwahl, F., and Rueda-Cantuche, J. M. (2010). Integrating a household demand system in the input–output framework. Methodological aspects and modelling implications. *Economic Systems Research*, 22(3):201–222.

Morais, L. T., Blatnik, N., and Kavonius, I. K. (2023). Cashing in on wealth: links between wealth and income inequality from the lens of Distributional Wealth Accounts.

OECD (2020). Distributional Information on Household Income, Consumption, and Saving in Line with National Accounts Guidelines.

OECD (2024). *OECD Handbook on the Compilation of Household Distributional Results on Income, Consumption and Saving in Line with National Accounts Totals*. OECD Publishing, Paris.

Okuyama, Y., Sonis, M., and Hewings, G. J. D. (1999). Economic Impacts of an Unscheduled, Disruptive Event: A Miyazawa Multiplier Analysis. In Hewings, G. J. D., Sonis, M., Madden, M., and Kimura, Y., editors, *Understanding and Interpreting Economic Structure*, pages 113–143. Springer Berlin Heidelberg, Berlin, Heidelberg. Series Title: Advances in Spatial Science.

Pallotti, F., Paz-Pardo, G., Slacalek, J., Tristani, O., and Violante, G. L. (2023). Who bears the costs of inflation?: euro area households and the 2021–2022 shock. Working Paper 2877, Publications Office, Frankfurt am Main.

Paul, S. K., Chowdhury, P., Moktadir, M. A., and Lau, K. H. (2021). Supply chain recovery challenges in the wake of COVID-19 pandemic. *Journal of Business Research*, 136:316–329.

Peichl, A. (2015). Linking Microsimulation and CGE models. *International Journal of Microsimulation*, 9(1):167–174.

Pichler, A. and Farmer, J. D. (2022). Simultaneous supply and demand constraints in input–output networks: the case of Covid-19 in Germany, Italy, and Spain. *Economic Systems Research*, 34(3):273–293. Publisher: Routledge _eprint: https://doi.org/10.1080/09535314.2021.1926934.

Pichler, A., Pangallo, M., Del Rio-Chanona, R. M., Lafond, F., and Farmer, J. D. (2022). Forecasting the propagation of pandemic shocks with a dynamic input-output model. *Journal of Economic Dynamics and Control*, 144:104527.

Piketty, T., Saez, E., and Zucman, G. (2018). Distributional National Accounts: Methods and Estimates for the United States*. *The Quarterly Journal of Economics*, 133(2):553–609.

Pyatt, G. (2001). Some Early Multiplier Models of the Relationship Between Income Distribution and Production Structure. *Economic Systems Research*, 13(2):139–164.

Pyatt, G. and Round, J. I. (1979). Accounting and Fixed Price Multipliers in a Social Accounting Matrix Framework. *The Economic Journal*, 89(356):850–873.

Pyatt, G. and Round, J. I. (1985). *Social accounting matrices: A basis for planning*. The World Bank, Washington, D.C.

Pyatt, G. and Round, J. I. (2006). Multiplier Effects and the Reduction of Poverty. In de Janvry, A. and Kanbur, R., editors, *Poverty, Inequality, and Development: Essays in Honor of Erik Thorbecke*, pages 233–259. Springer, New York.

Pyatt, G. and Thorbecke, E. (1976). *Planning techniques for a better future: a summary of a research project on planning for growth, redistribution and employment*. International Labour Office, Geneva.

Rässler, S. (2004). Data fusion: Identification Problems, Validity, and Multiple Imputation. *Austrian Journal of Statistics*, 33:153–171.

Remond-Tiedrez, I. and Rueda-Cantuche, J. M. (2019). *EU Inter-Country Supply, Use and Input-Output Tables Full International and Global Accounts for Research in Input-Output Analysis (FIGARO)*. Publications Office of the European Union, Luxembourg.

Rodrik, D. (2006). Industrial Development: Stylized Facts and Policies. In Ocampo, J. A., editor, *Industrial Development for the 21st Century*, pages 7–28. United Nations, New York.

Roland-Holst, D. W. and Sancho, F. (1992). Relative Income Determination in the United States: A Social Accounting Perspective. *Review of Income and Wealth*, 38(3):311–327.

Roland-Holst, D. W. and Sancho, F. (1995). Modeling Prices in a SAM Structure. *The Review of Economics and Statistics*, 77(2):361–371.

Rose, A. and Beaumont, P. (1988). Interrelational Income Distribution Multipliers for West Virginia. *Journal of Regional Science*, 28(4):461–475.

Rose, A., Li, D., and Schluter, G. (1999). Construction of an Input-Output Income Distribution Matrix for the U.S. In Hewings, G. J. D., Sonis, M., Madden, M., and Kimura, Y., editors, *Understanding and Interpreting Economic Structure*, pages 191–212. Springer, Berlin, Heidelberg.

Round, J. I. (2003). Social Accounting Matrices and SAM-based Multiplier Analysis. In Bourguignon, F. and Pereira da Silva, L. A., editors, *Techniques and Tools for Evaluating the Poverty Impact of Economic Policies*, chapter 14. World Bank and Oxford University Press, Washington, D.C.

Rubbo, E. (2024). What Drives Inflation? Lessons from Disaggregated Price Data.

S. Basso, H., Dimakou, O., and Pidkuyko, M. (2023). How inflation varies across Spanish households. Documentos Ocasionales 2307, Banco de España, Madrid. Series: Documentos Ocasionales.

Serafino, P. and Tonkin, R. (2017). Statistical matching of European Union statistics on income and living conditions (EU-SILC) and the household budget survey. Working Paper, Publications Office of the European Union, Luxembourg.

Shaikh, A. (2016). *Capitalism. Competition, Conflict, Crisis.* Oxford University Press, New York, first edition.

Sonis, M. and Hewings, G. J. D. (1999). Miyazawa's Contributions to Understanding Economic Structure: Interpretation, Evaluation and Extensions. In Hewings, G. J. D., Sonis, M., Madden, M., and Kimura, Y., editors, *Understanding and Interpreting Economic Structure*, pages 13–51. Springer Berlin Heidelberg, Berlin, Heidelberg. Series Title: Advances in Spatial Science.

Steenge, A. E., Incera, A. C., and Serrano, M. (2020). Income distributions in multi-sector analysis; Miyazawa's fundamental equation of income formation revisited. *Structural Change and Economic Dynamics*, 53:377–387.

Steenge, A. E. and Serrano, M. (2012). Income Distribution in Input-Output Models. *Economic Systems Research*, 24(4):391–412.

Stiglitz, J. E., Sen, A., and Fitoussi, J.-P. (2009). Report by the Commission on the Measurement of Economic Performance and Social Progress. Technical report, Commission on the Measurement of Economic Performance and Social Progress, Paris.

Stone, R. (1985). The Disaggregation of the Household Sector in the National Accounts. In Pyatt, G. and Round, J. I., editors, *Social Accounting Matrices: A Basis for Planning*, pages 145–185. The World Bank, Washington, D.C.

Tunalı Çalışkan, E. and Aydoğuş, O. (2007). Türkiye ve Seçilmiş Avrupa Ülkelerinde Enerji Fiyatlarının Enflasyonist Etkileri. *Ege Academic Review*. Inflationary Effects of Energy Prices in Turkey and Selected European Countries.

UNDESA (2013). Guidelines on Integrated Economic Statistics. Technical Report 108, United Nations, New York. OCLC: 892919254.

UNDESA (2018). Handbook on Supply and Use Tables and Input-Output Tables with Extensions and Applications. Technical report, United Nations Publication, New York.

Uxó, J., Febrero, E., and Álvarez, I. (2025). Prices, markups and wages: inflation and its distributive consequences in Spain, 2021-2023. *Structural Change and Economic Dynamics*, 72:179–192.

Valadkhani, A. and Mitchell, W. F. (2002). Assessing the Impact of Changes in Petroleum Prices on Inflation and Household Expenditures in Australia. *Australian Economic Review*, 35(2):122–132.

Van'T Klooster, J. and Weber, I. M. (2024). Closing the EU's inflation governance gap. The limits of monetary policy and the case for a new policy framework for shockflation. Technical report, Economic Governance and EMU Scrutiny Unit (EGOV), European Comission.

Weber, I., Jauregui, J. L., Teixeira, L., and Pires, L. N. (2024a). Inflation in times of overlapping emergencies: Systemically significant prices from an input–output perspective. *Industrial and Corporate Change*, 33(2):297–341.

Weber, I. M., Thie, J.-E., Jauregui, J. L., and Teixeira, L. (2024b). Carbon Prices and Inflation in a World of Shocks: Systemically significant prices and industrial policy targeting in Germany. Technical report, Bertelsmann Stiftung, Sustainable Social Market Economies, Gütersloh. Publisher: Bertelsmann Stiftung.

Weber, I. M. and Wasner, E. (2023). Sellers' inflation, profits and conflict: why can large firms hike prices in an emergency?: Online appendix. *Review of Keynesian Economics*, 11(2):1–21.

Wimer, C., Collyer, S., and Jaravel, X. (2022). The costs of being poor: inflation inequality leads to three million more people in poverty. Technical report, Center on Poverty & Social Policy at Columbia University.

Appendix A: Valuation adjustment: from purchaser's prices to basic prices

As noted in Section 3.3, while the IOTs are compiled in basic prices, the valuation of $\mathbf{C}^{pp}_{cpa}$ corresponds to purchaser's prices (Lequiller and Blades, 2014). Due to indirect taxes and distribution costs—which include transport costs, trade margins, and taxes less subsidies on products (Eurostat, 2013, p. 17)—consumers and producers face different prices. The fundamental accounting identity linking the two valuations is:

$$p_i^{pp} = p_i^{bp} + T_i + M_i \tag{14}$$

where p_i^{pp} denotes household consumption of product i in purchaser's prices, p_i^{bp} denotes the same flow in basic prices, T_i represents net taxes on products (taxes less subsidies), and M_i captures trade and transport margins.

Deriving implicit rates

Moving from purchaser's prices to basic prices requires knowledge of net tax rates and trade and transport margin rates, which are confidential for several countries including Spain and Germany. Following Mongelli et al. (2010) and Cazcarro et al. (2022), we derive implicit rates from available supply-use table components. Let τ_i^{tls} denote the implicit net tax rate and τ_i^{ttm} the implicit trade and transport margin rate for product i. These are computed as:

$$\tau_i^{tls} = \frac{T_i^{ref}}{p_i^{bp,ref}} \qquad \text{and} \qquad \tau_i^{ttm} = \frac{M_i^{ref}}{p_i^{bp,ref}} \tag{15}$$

where the superscript ref indicates reference values from the supply-use tables or estimated from comparable countries.

Transformation algorithm

Given expenditure in purchaser's prices p_i^{pp} and the implicit rates τ_i^{tls} and τ_i^{ttm}, the transformation to basic prices proceeds as follows.

Step 1. Compute the tax component by inverting the markup relationship:

$$T_i = p_i^{pp} - \frac{p_i^{pp}}{1+\tau_i^{tls}} = p_i^{pp} \cdot \frac{\tau_i^{tls}}{1+\tau_i^{tls}} \tag{16}$$

Step 2. Compute trade and transport margins for products with *non-negative* margin rates. Define the pre-margin value as $\tilde{p}_i = p_i^{pp}/(1+\tau_i^{tls})$, then:

$$M_i^{+} = \begin{cases} \dfrac{\tilde{p}_i}{1+\tau_i^{ttm}} \cdot \tau_i^{ttm} & \text{if } \tau_i^{ttm} \geq 0 \\ 0 & \text{if } \tau_i^{ttm} < 0 \end{cases} \tag{17}$$

Step 3. Aggregate positive margins to form the redistribution base. Since trade and transport margins represent a reallocation across products (the margin industries receive what other products "give up"), they must sum to zero in aggregate. The positive margins provide the pool to be redistributed:

$$\bar{M} = \sum_{i:\,\tau_i^{ttm} \geq 0} M_i^+ \tag{18}$$

Step 4. Compute margins for products with *negative* margin rates. These represent products where the margin industries effectively subsidize distribution (e.g., certain agricultural or energy products):[4]

$$M_i = \begin{cases} M_i^+ & \text{if } \tau_i^{ttm} \geq 0 \\ \tau_i^{ttm} \cdot \bar{M} & \text{if } \tau_i^{ttm} < 0 \end{cases} \tag{19}$$

Step 5. Derive basic prices from the accounting identity:

$$p_i^{bp} = p_i^{pp} - T_i - M_i \tag{20}$$

In matrix form, let $\mathbf{c}^{pp}$ be the row vector of household consumption in purchaser's prices. Define the diagonal matrices $\hat{\alpha} = \mathrm{diag}(\tau_i^{tls}/(1+\tau_i^{tls}))$ and $\hat{\beta} = \mathrm{diag}(\tau_i^{ttm}/(1+\tau_i^{ttm}))$, with appropriate adjustments for negative margin rates. The transformation can be written compactly as:

$$\underset{cpa}{\mathbf{c}^{bp}} = F(\underset{cpa}{\mathbf{c}^{pp}}) = \mathbf{c}^{pp}(\mathbf{I} - \hat{\alpha} - \hat{\gamma}) \tag{21}$$

where $\hat{\gamma}$ is the diagonal matrix of effective margin shares incorporating the redistribution logic from Steps 2–4. We verify the transformation by confirming that $\mathbf{c}^{pp} = \mathbf{c}^{bp} + \mathbf{t} + \mathbf{m}$, where $\mathbf{t}$ and $\mathbf{m}$ are the vectors of tax and margin components.

The reference vectors used in the calculation are matched to the corresponding entries in the FIGARO IOTs, even where this implies some deviation from other intermediate annual national accounts figures.

[4]For products with negative margin rates, we compute margins as a share of the aggregate positive margin pool ($\bar{M}$) rather than using each product's own expenditure base. This treatment preserves the redistributive interpretation of trade and transport margins and ensures internal consistency when aggregating across products.

Appendix B - Robustness

This appendix documents the quantitative evaluation of the two central statistical procedures described in Section 3.1 and Section 3.2: (i) the statistical matching of ES-SILC and ES-HBS, and (ii) the grossing-up adjustment to national accounts benchmarks. Following the validation hierarchy established by Rässler (2004) and operationalised in recent Eurostat-JRC work (Dreoni et al., 2025), the diagnostics are organised in decreasing order of stringency: preservation of marginal distributions (Level 4), correlation structures (Level 3), and joint distributional patterns (Level 2). Level 1 validation—comparison of imputed versus true individual values—requires an administratively merged dataset and is beyond the scope of this exercise. All diagnostics cover the 2010–2023 series.

B.1 Statistical matching diagnostics

B.1.1 Donor usage and fallback levels

Figure 15 plots the distribution of donor uses and the number of matching variables retained across the series. A well-behaved matching concentrates donor uses at or below the constrained maximum (≤ 3) and requires minimal variable fallback. Figure 16 shows the annual share of unmatched observations remaining after all fallback steps.

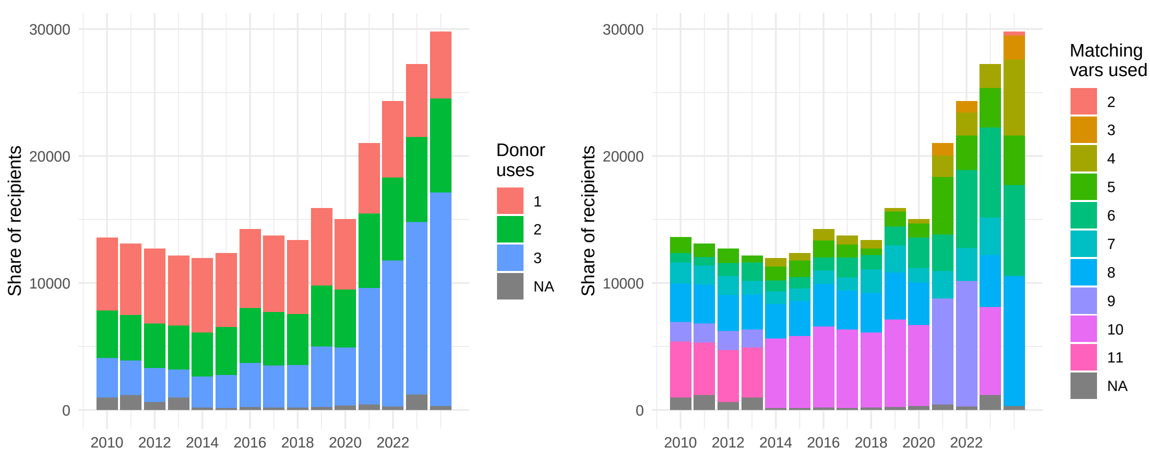


Figure 15: Distribution of donor uses (left) and matching variable fallback levels (right) across the 2010–2023 series. *Source*: Author's elaboration.

B.1.2 Preservation of marginal distributions (Level 4)

The minimum validation requirement is that the matched dataset preserves the marginal distributions of the donor variables (Rässler, 2004). Figure 17 reports two summary measures: the weighted mean absolute deviation (WMAD) of expenditure

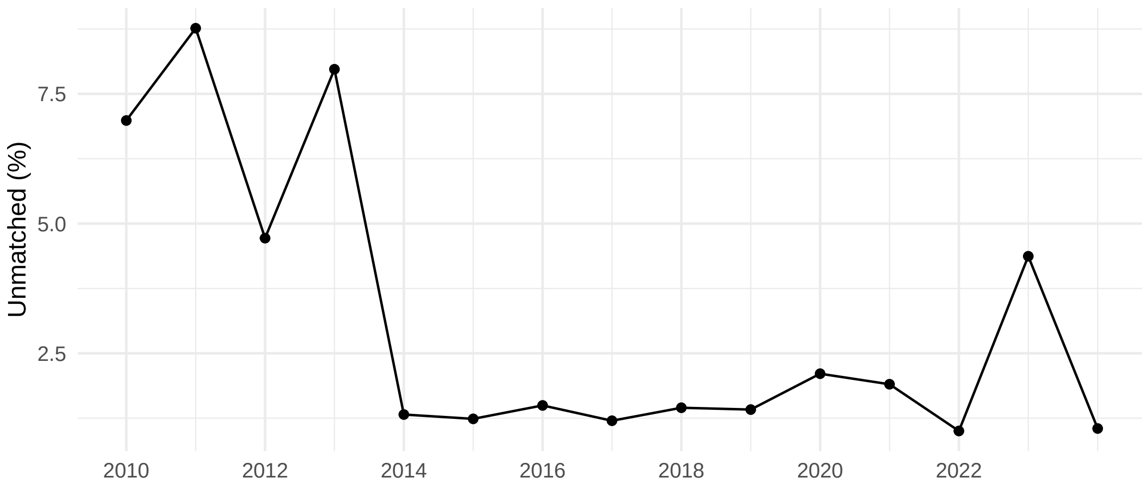


Figure 16: Share of unmatched observations by year. *Source*: Author's elaboration.

shares between the matched and original distributions, and the Hellinger distance over the full expenditure distribution. Both are computed at the 2-digit COICOP level.

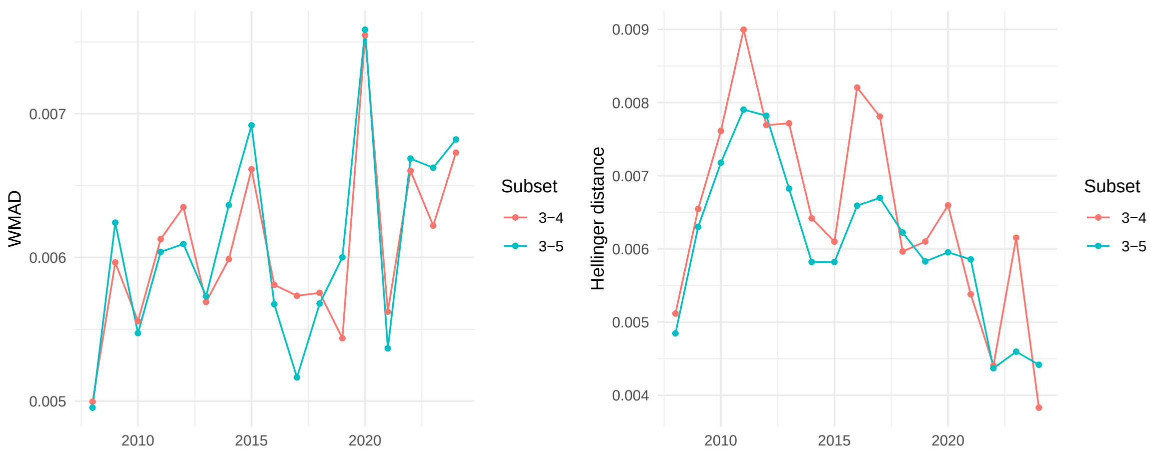


Figure 17: WMAD of expenditure shares (left) and Hellinger distance (right) between the matched synthetic dataset and the original ES-HBS donor distributions, 2010–2023. *Source*: Author's elaboration.

We can further assess the robustness of the statistical matching by visual inspection of the imputed and the observed distribution of income shares of expenditures across income ventiles. Figures 18 to 20 compare the two for each 3-digit COICOP heading (see Dreoni et al., 2025; Akoğuz et al., 2020; Capéau et al., 2022). The closest the two curves are the better the matching peserves the original distirbution.

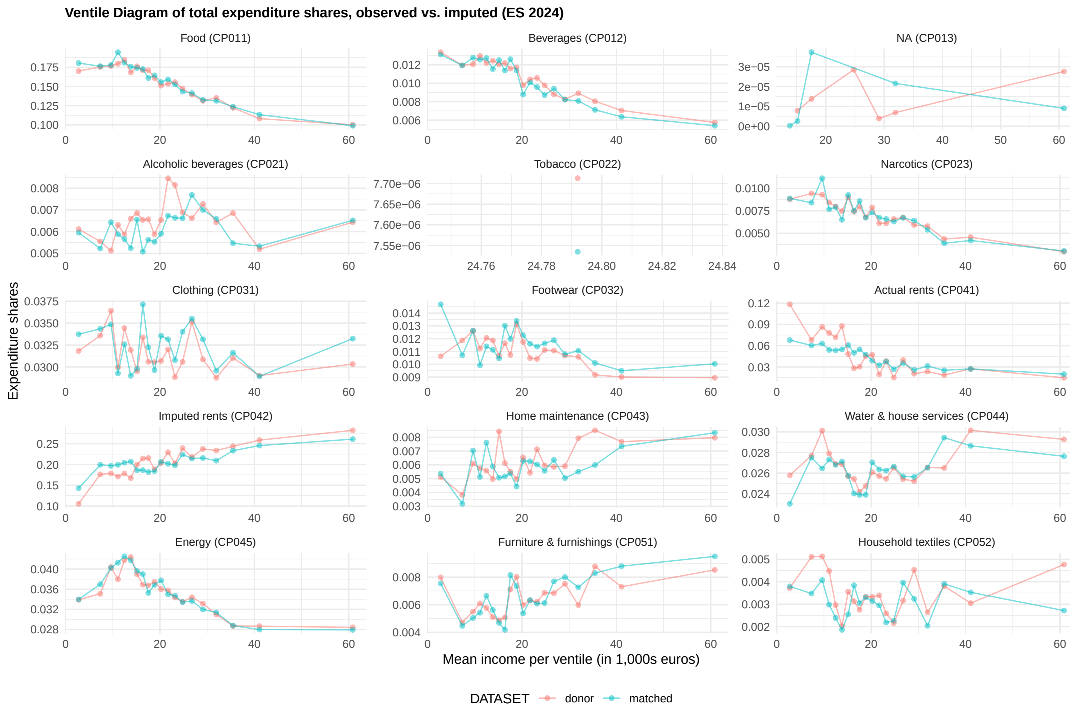


Figure 18: Ventile Diagram of expenditure shares by COICOP heading (CP011 to CP052) in 2024, observed vs. imputed. *Source*: Author's elaboration.

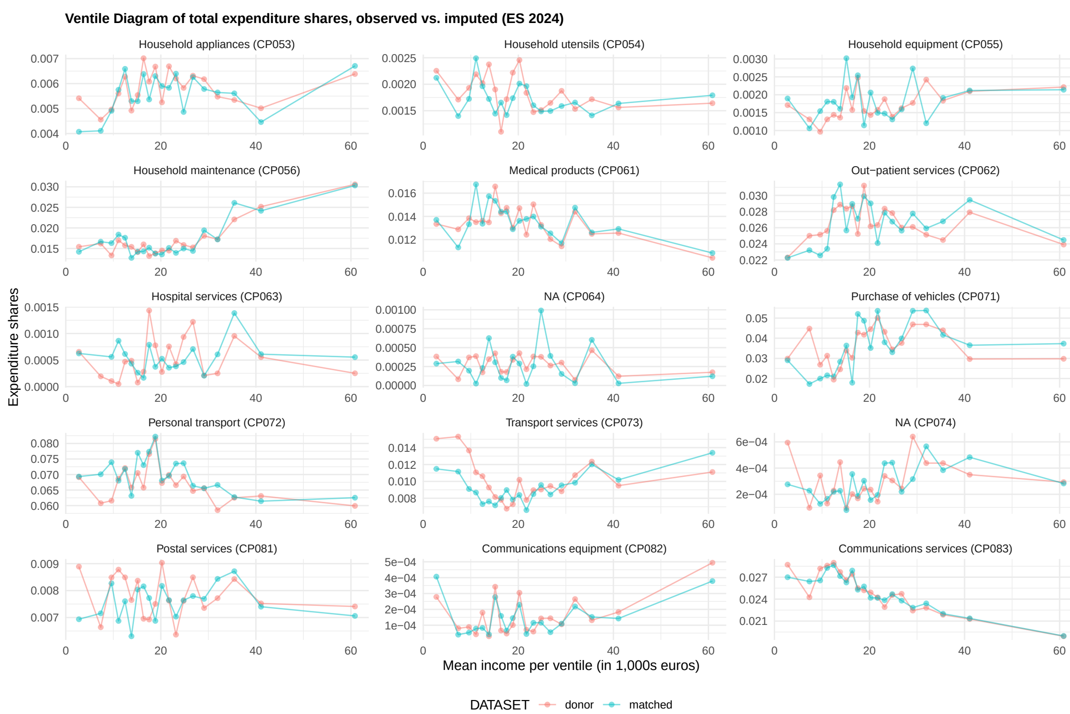


Figure 19: Ventile Diagram of expenditure shares by COICOP heading (CP053 to CP083) in 2024, observed vs. imputed. *Source*: Author's elaboration.

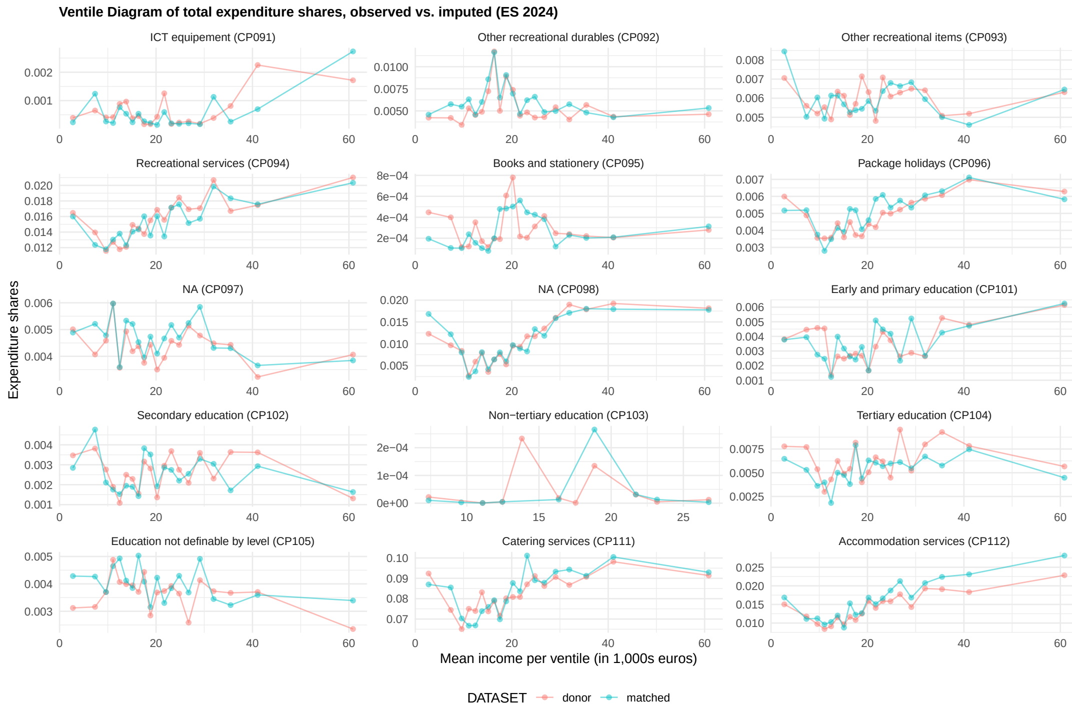


Figure 20: Ventile Diagram of expenditure shares by COICOP heading (CP091 to CP0112) in 2024, observed vs. imputed. *Source*: Author's elaboration.

B.1.3 Preservation of correlation structures (Level 3)

Beyond marginal distributions, the matching should preserve within-group expenditure patterns conditional on sociodemographic covariates, approximating the correlation structure criterion (Rässler, 2004; Dreoni et al., 2025). Figure 21 reports the WMAD of expenditure shares within strata defined by each matching covariate, averaged across 2-digit COICOP categories. Smaller deviations indicate that the imputed expenditure-covariate association faithfully mirrors the donor dataset.

Figure 22 provides the time-series evolution of this within-covariate WMAD for each matching variable, allowing us to identify whether specific covariates experience deterioration in matching quality in particular years—for instance, following changes in survey design or sample composition.

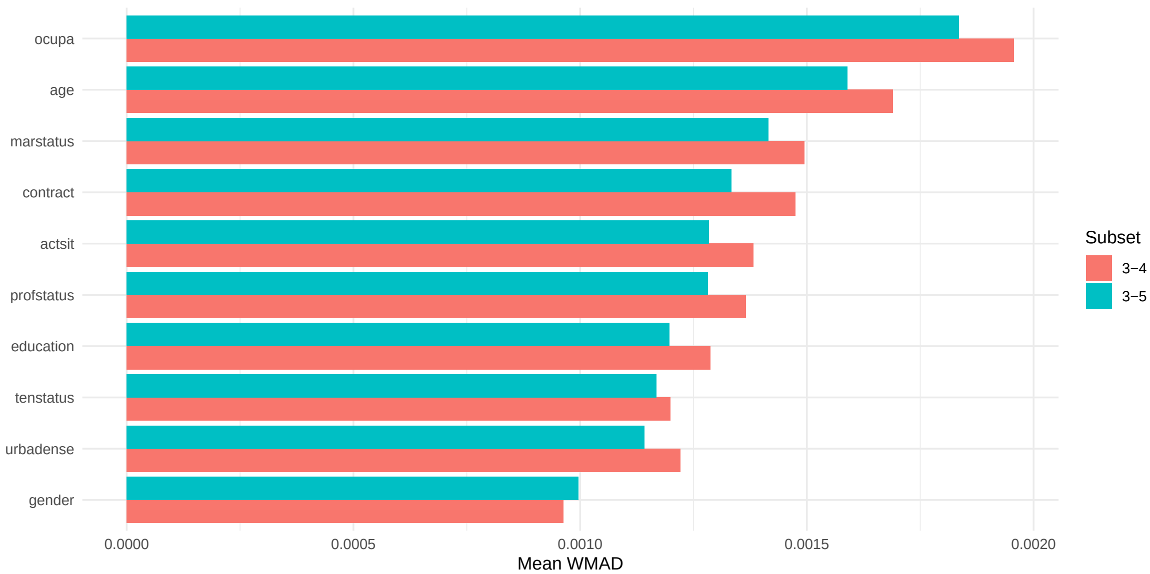


Figure 21: Within-covariate WMAD of 2-digit COICOP expenditure shares, averaged across years. Lower values indicate better preservation of the covariate-expenditure association. *Source*: Author's elaboration.

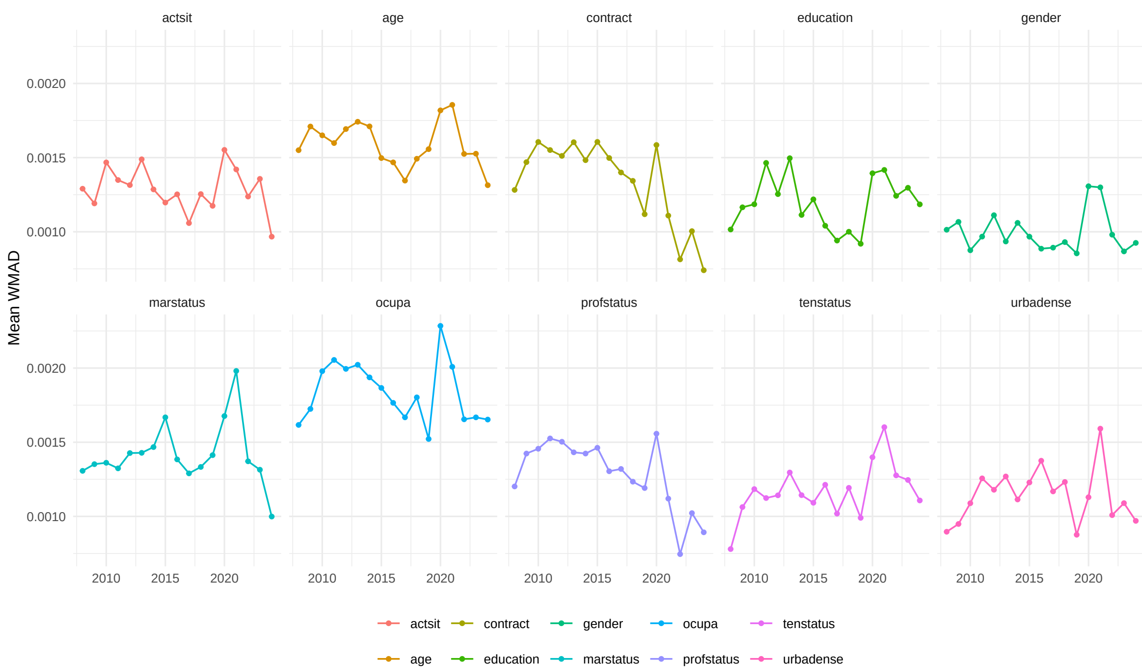


Figure 22: Within-covariate WMAD over time by matching variable (baseline subset). *Source*: Author's elaboration.

B.2 National accounts linkage evaluation

B.2.1 Coverage ratios

Figure 23 display the pre-adjustment coverage ratios (survey total / national accounts target) for income and expenditure items, respectively. Values below unity

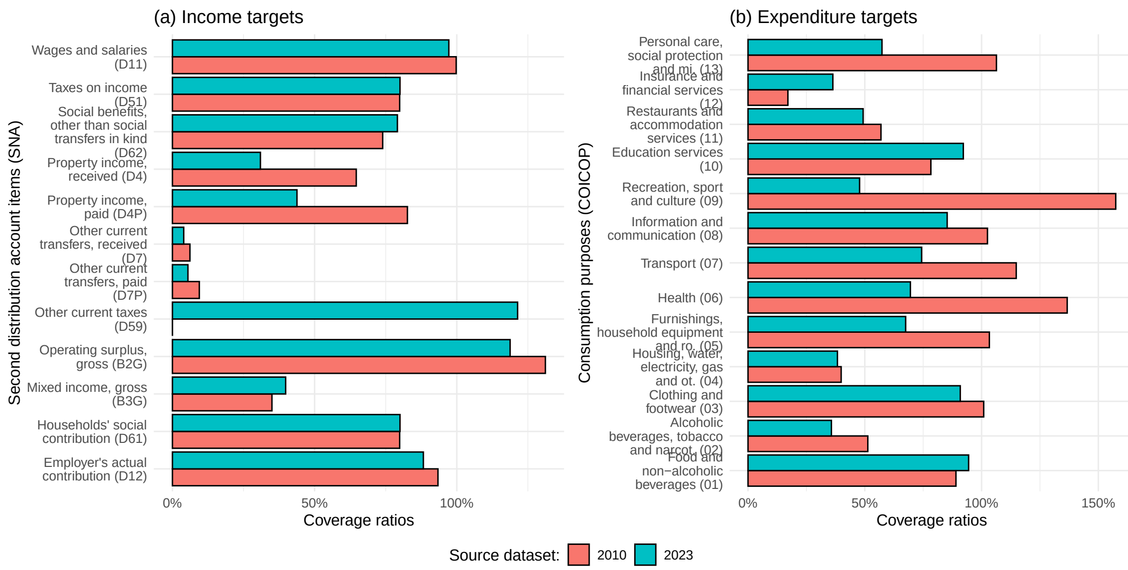


Figure 23: Initial SILC-HBS to national accounts coverage ratios, 2023. *Source*: Author's elaboration.

indicate survey under-reporting; values above unity arise occasionally for specific transfers or from differences in scope.

B.2.4 Post-adjustment consistency

Figure 24 confirms that the grossing-up procedure brings micro totals into alignment with national accounts targets. Each point represents an income item–year combination; proximity to the 45-degree line indicates successful adjustment.

B.2.5 Distributional impact of the grossing-up

Figure 25 plots the Gini coefficient of equivalized disposable income after the grossing-up, confirming that the progressive gap allocation preserves the broad inequality profile while bringing aggregates into consistency with the national accounts.

Figure 26a and Figure 26b displays the decile composition of selected national accounts income items, illustrating how the gap allocation distributes the micro-to-macro adjustment across the income distribution. The progressive gradient concentrates adjustment in upper deciles for most items, consistent with the well-documented pattern of top-income under-reporting in household surveys (Morais et al., 2023; Coli et al., 2022).

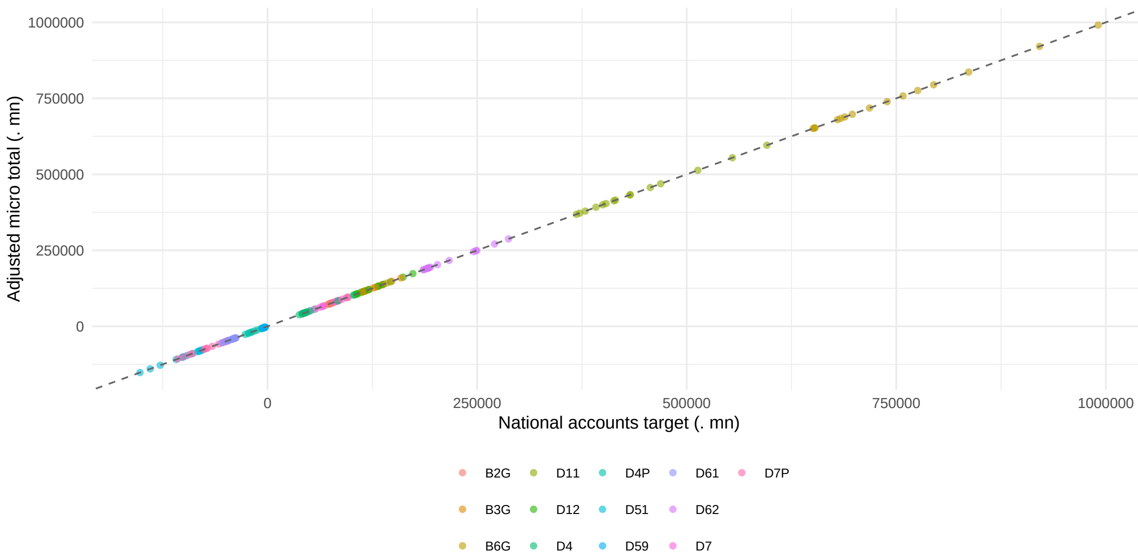


Figure 24: Post-grossup micro totals versus national accounts targets by income item, 2010–2023. *Source*: Author's elaboration.

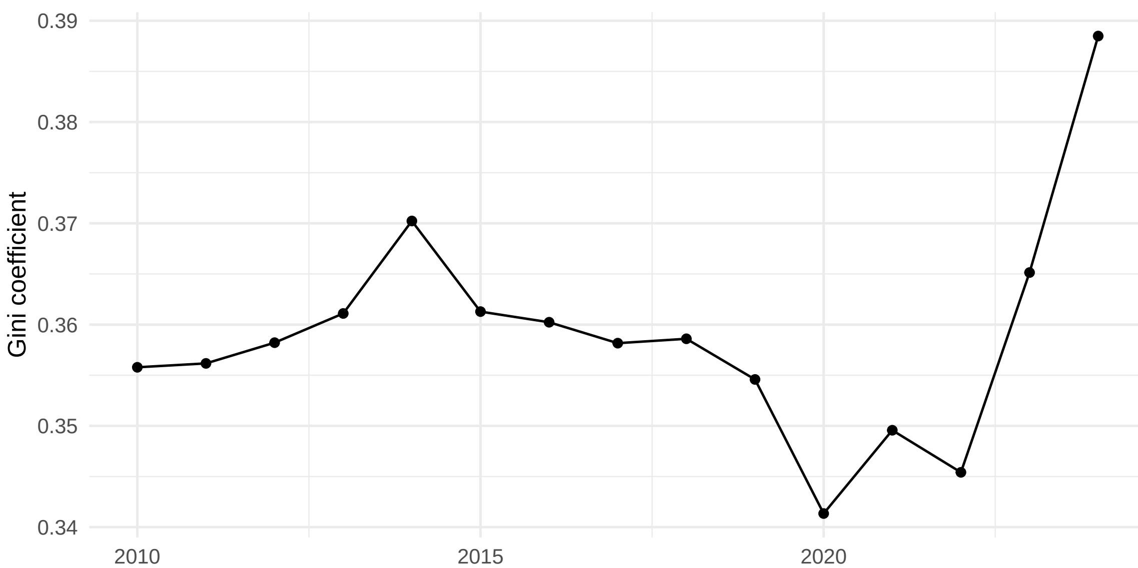
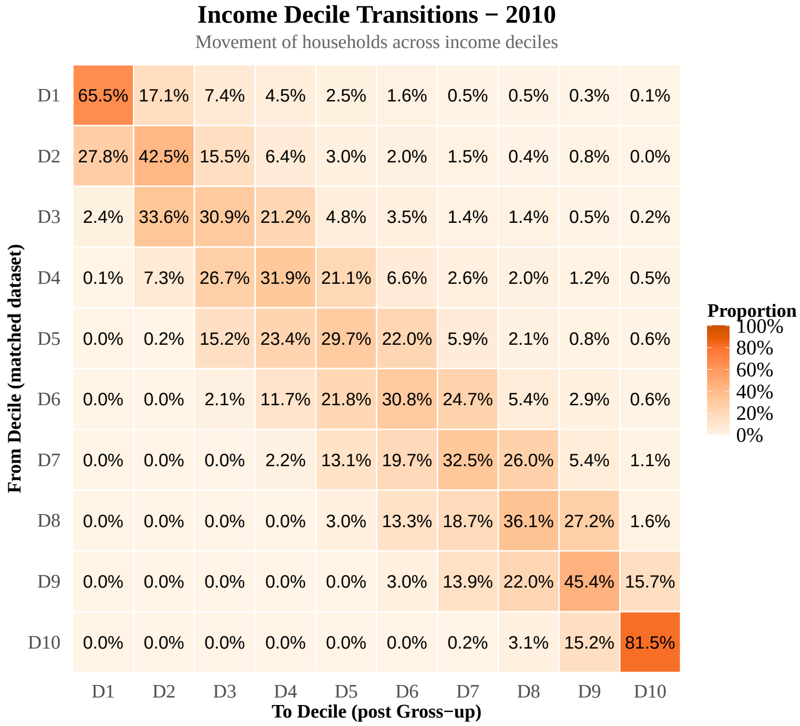


Figure 25: Gini coefficient of equivalized household disposable income, post-grossup, 2010–2023. *Source*: Author's elaboration.

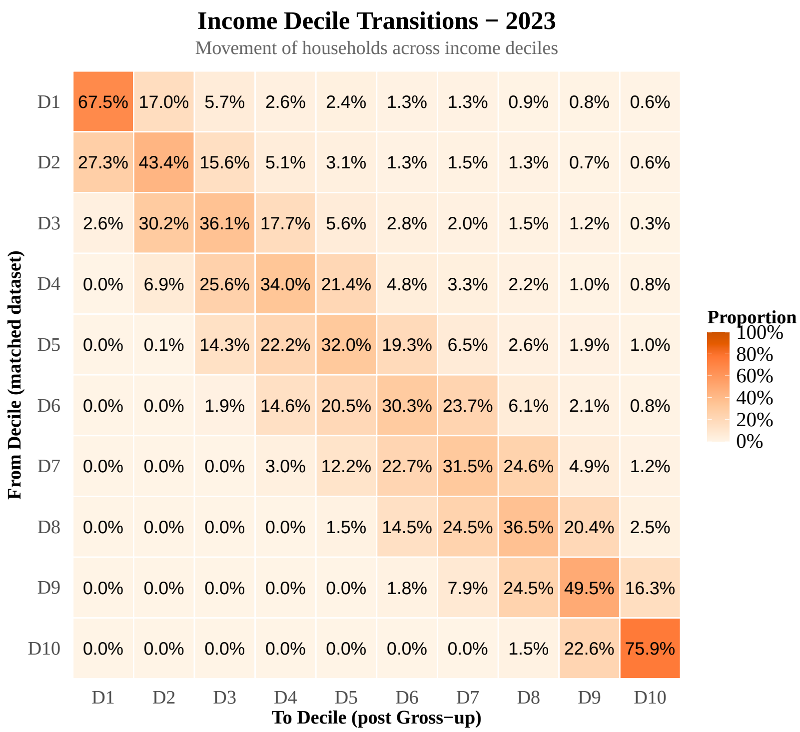
B.2.6 Saving rates by quintile

Finally, Figure 27 compares the implicit household saving rate by quintile across the original HBS, the matched synthetic dataset, and the post-grossup distributional accounts. Consistency of these profiles across processing steps provides additional confidence that the joint income–expenditure distribution is internally coherent.



	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
D1	65.5%	17.1%	7.4%	4.5%	2.5%	1.6%	0.5%	0.5%	0.3%	0.1%
D2	27.8%	42.5%	15.5%	6.4%	3.0%	2.0%	1.5%	0.4%	0.8%	0.0%
D3	2.4%	33.6%	30.9%	21.2%	4.8%	3.5%	1.4%	1.4%	0.5%	0.2%
D4	0.1%	7.3%	26.7%	31.9%	21.1%	6.6%	2.6%	2.0%	1.2%	0.5%
D5	0.0%	0.2%	15.2%	23.4%	29.7%	22.0%	5.9%	2.1%	0.8%	0.6%
D6	0.0%	0.0%	2.1%	11.7%	21.8%	30.8%	24.7%	5.4%	2.9%	0.6%
D7	0.0%	0.0%	0.0%	2.2%	13.1%	19.7%	32.5%	26.0%	5.4%	1.1%
D8	0.0%	0.0%	0.0%	0.0%	3.0%	13.3%	18.7%	36.1%	27.2%	1.6%
D9	0.0%	0.0%	0.0%	0.0%	0.0%	3.0%	13.9%	22.0%	45.4%	15.7%
D10	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.2%	3.1%	15.2%	81.5%

(a) 2010



	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
D1	67.5%	17.0%	5.7%	2.6%	2.4%	1.3%	1.3%	0.9%	0.8%	0.6%
D2	27.3%	43.4%	15.6%	5.1%	3.1%	1.3%	1.5%	1.3%	0.7%	0.6%
D3	2.6%	30.2%	36.1%	17.7%	5.6%	2.8%	2.0%	1.5%	1.2%	0.3%
D4	0.0%	6.9%	25.6%	34.0%	21.4%	4.8%	3.3%	2.2%	1.0%	0.8%
D5	0.0%	0.1%	14.3%	22.2%	32.0%	19.3%	6.5%	2.6%	1.9%	1.0%
D6	0.0%	0.0%	1.9%	14.6%	20.5%	30.3%	23.7%	6.1%	2.1%	0.8%
D7	0.0%	0.0%	0.0%	3.0%	12.2%	22.7%	31.5%	24.6%	4.9%	1.2%
D8	0.0%	0.0%	0.0%	0.0%	1.5%	14.5%	24.5%	36.5%	20.4%	2.5%
D9	0.0%	0.0%	0.0%	0.0%	0.0%	1.8%	7.9%	24.5%	49.5%	16.3%
D10	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	1.5%	22.6%	75.9%

(b) 2024

Figure 26: Share of households changing equivalized income decile after the grossing-up adjustment to national accounts. *Source*: Author's elaboration.

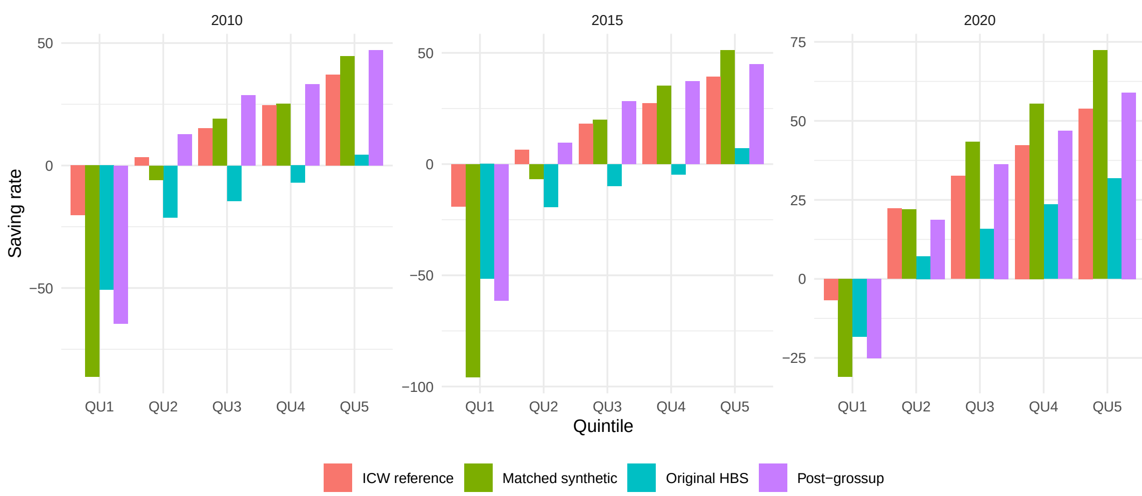


Figure 27: Saving rates by income quintile: original HBS, matched synthetic dataset, and post-grossup distributional accounts. *Source*: Author's elaboration.

Appendix C: Other intermediate and final results

National Accounts		ES-SILC	
Definition	**Code**	**Definition**	**Code**
Operating surplus, gross	B2G	Imputed rent	HY030N
Mixed income, gross	B3G	Income from self-employment Value of own production	PY050G HY170G
Wages and salaries	D11	Employee cash or near cash income Non-cash employee income Company car	PY010G PY020G PY021G
Employer's actual contribution	D12	Employer's actual contribution	PY030G
Property income, received Property income, receive Property income, paid	D4	Interest, dividends, profits Income from property rents Interest repayments on mortgage	HY090G HY040G HY100G
Taxes on income	D51	Taxes on income and social contr.	HY140G
Other current taxes	D59	Regular taxes on wealth	HY120G
Households' actual social contr. Households' social contr.	D613 D614	Taxes on income and social contr.	HY140G
Social benefits and transfers	D62	Unemployment benefits Old-age benefits Survivor benefits Sickness benefits Disability benefits Education-related allowances Housing allowances Family/children related allowances Other social exclusion	PY090G PY100G PY110G PY120G PY130G PY140G HY070G HY050G HY060G
Other current transfers, received Other current transfers, paid	D7	Regular inter-household transfer	HY080G HY110G HY130G

Table 5: Correspondence between ES-SILC and National Accounts income items

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
Factors										
Food	-0.24	-0.37	0.04	0.00	0.03	-0.03	0.05	-0.05	0.08	-0.03
Beverages	-0.24	-0.37	0.04	0.00	0.03	-0.03	0.05	-0.06	0.03	-0.03
Alcoholic beverages	-0.24	-0.37	0.04	0.00	0.03	-0.03	0.05	-0.06	0.03	-0.03
Tobacco	-0.24	-0.37	0.04	-0.01	0.03	-0.04	0.05	-0.05	0.04	-0.03
Clothing	-0.03	0.03	-0.37	0.10	0.07	0.18	0.25	-0.31	-0.04	0.10
Footwear	-0.03	0.02	-0.38	0.09	0.06	0.13	0.15	-0.38	-0.01	0.15
Actual rents	0.00	0.05	0.07	-0.05	0.56	-0.05	-0.05	-0.03	-0.03	-0.01
Imputed rents	0.00	0.05	0.07	-0.05	0.56	-0.05	-0.05	-0.03	-0.03	-0.01
Home maintenance	0.00	0.02	0.01	-0.01	-0.01	0.02	0.00	-0.01	0.03	0.09
Water & house services	0.00	0.05	0.07	-0.05	0.56	-0.05	-0.05	-0.03	-0.04	-0.01
Energy	-0.03	0.02	0.02	-0.02	-0.02	0.12	-0.05	0.09	-0.30	0.06
Furniture & furnishings	-0.03	0.00	-0.33	0.05	0.03	-0.16	-0.23	0.36	0.01	-0.18
Household textiles	-0.03	0.02	-0.44	0.10	0.06	0.07	0.06	-0.20	0.01	0.12
Household appliances	-0.01	0.01	0.00	-0.06	-0.05	0.10	-0.18	-0.06	-0.42	-0.03
Household utensils	-0.02	0.00	-0.03	-0.03	-0.03	0.05	-0.06	0.03	-0.08	0.20
Household equipment	-0.11	-0.13	-0.02	-0.07	-0.05	0.06	-0.13	0.01	-0.45	0.13
Household maintenance	-0.16	-0.20	-0.08	-0.01	0.01	0.01	0.00	0.09	-0.10	-0.01
Medical products	-0.01	0.00	-0.09	-0.01	-0.02	-0.22	-0.01	0.16	-0.04	-0.07
Out-patient services	0.06	0.04	-0.02	-0.08	-0.06	-0.54	0.30	-0.09	-0.24	0.00
Hospital services	0.06	0.04	-0.02	-0.08	-0.06	-0.54	0.30	-0.09	-0.24	0.00
Purchase of vehicles	0.00	0.02	0.01	-0.03	-0.02	0.09	-0.12	0.00	-0.21	0.19
Personal transport	-0.05	0.03	0.02	-0.01	-0.02	0.15	-0.12	0.12	-0.27	0.25
Transport services	0.07	0.03	-0.02	-0.18	-0.04	0.02	-0.03	-0.06	0.21	0.03
Postal services	0.21	0.01	-0.06	-0.35	-0.05	-0.01	0.00	-0.08	0.22	-0.04
Communications equipment	0.01	0.02	0.00	-0.04	-0.05	0.05	-0.21	-0.29	-0.17	-0.52
Communications services	0.00	0.02	0.03	0.02	-0.02	0.01	-0.02	-0.06	0.06	-0.04
ICT equipement	0.01	0.04	0.02	-0.03	-0.07	0.05	-0.20	-0.36	-0.09	-0.52
Other recreational durables	-0.04	0.00	-0.38	0.07	0.04	-0.12	-0.19	0.27	0.03	-0.12
Other recreational items	-0.18	-0.25	-0.04	0.01	0.02	-0.05	-0.02	0.06	0.10	-0.03
Recreational services	-0.05	0.04	0.07	0.13	-0.02	-0.02	-0.02	-0.08	0.15	-0.03
Books and stationery	0.01	0.01	-0.08	-0.02	-0.02	0.00	-0.04	0.12	0.10	-0.06
Package holidays	-0.11	0.13	0.12	0.45	-0.01	-0.08	-0.03	-0.04	0.03	0.00
Early and primary education	0.34	-0.23	0.00	0.16	0.03	0.01	-0.01	0.01	-0.04	0.00
Secondary education	0.34	-0.23	0.00	0.16	0.03	0.01	-0.01	0.01	-0.04	0.00
Non-tertiary education	0.34	-0.23	0.00	0.16	0.03	0.01	-0.01	0.01	-0.04	0.00
Tertiary education	0.34	-0.23	0.00	0.16	0.03	0.01	-0.01	0.01	-0.04	0.00
Education not definable by level	0.34	-0.23	0.00	0.16	0.03	0.01	-0.01	0.01	-0.04	0.00
Catering services	-0.13	0.14	0.12	0.45	-0.01	-0.08	-0.02	-0.03	0.01	0.00
Accommodation services	-0.13	0.14	0.12	0.45	-0.01	-0.09	-0.02	-0.03	0.01	0.00
Personal care	-0.02	0.03	-0.01	0.04	0.04	0.26	0.46	0.25	-0.14	-0.28
Personal effects n.e.c.	-0.03	0.01	-0.42	0.07	0.04	-0.10	-0.10	0.16	0.02	-0.13
Social protection	0.06	0.02	-0.01	-0.09	-0.02	-0.04	0.06	-0.02	0.16	0.04
Insurance	-0.01	0.03	0.02	-0.01	-0.01	0.06	0.08	0.07	0.10	-0.01
Financial services n.e.c.	0.02	0.02	0.02	-0.04	-0.02	0.00	0.04	-0.01	0.10	-0.03
Other services n.e.c.	-0.01	0.06	0.02	0.02	0.07	0.28	0.47	0.26	-0.12	-0.28
Parameters										
Eigenvalue	5.79	4.36	4.03	3.30	3.00	2.12	1.96	1.72	1.64	1.57
Variance (%)	12.88	9.68	8.96	7.33	6.67	4.71	4.37	3.83	3.66	3.50
Cumulative (%)	12.88	22.56	31.52	38.85	45.52	50.23	54.59	58.42	62.07	65.57

Table 6: Principal components analysis of unique industry shocks on 3-digit COICOP items. *Source*: Author's elaboration.