Narrow Distributional Concerns in Households: How Common Are They?*

Sam Cosaert† Wenqi Lu‡

March 20, 2026

Abstract

This paper examines how money and time are allocated between spouses. Using evidence from Dutch couples, we show that within-household allocations and spousal preference alignment vary substantially across consumption, labor, and other time use decisions, with the strongest alignment in the labor domain. We propose an extended collective model that allows distributional preferences and preference alignment to hold narrowly within specific domains. Our structural nonparametric analysis then tests the homogeneity of household distributional concerns and links it with observable household characteristics, within each domain. Children and gender attitudes emerge as the key determinants of distributional concerns in time domains.

Keywords: intra-household allocations, narrow equity, preference alignment, revealed preference, LISS, consumption, non-market time, leisure, distributional types.

JEL classifications: D13, C14, D62, D63, D91, J22.

1 Introduction

Many family decisions involve distributing money and time within the household, such as setting consumption budgets and dividing work responsibilities between spouses. The household economics literature has significantly advanced our understanding of inequality within households, through influential frameworks such as the sharing rule and the collective model

*We thank Sebastiaan Maes, participants of the Household Economics 2023 Workshop in Leuven, and participants of the SEHO 2025 Conference in Zaragoza for their valuable input.

†Department of Economics, University of Antwerp. E-mail: sam.cosaert@uantwerpen.be. Sam Cosaert gratefully acknowledges FWO for financial support.

‡Department of Economics, University of Antwerp; ECARES, Université Libre de Bruxelles. E-mail: wenqi.lu@uantwerpen.be.

(Chiappori, 1992). However, analyses of the collective model often (implicitly or explicitly) adopt the egoistic specification, where an individual's willingness-to-pay for the spouse's consumption or leisure is zero. This assumption overlooks potential spillovers at the household level. In practice, such spillovers are likely widespread and policy relevant. For example, a husband may be willing to work more in order to reduce his wife's work hours, complicating government interventions to increase female labor supply. Furthermore, policies targeting vulnerable individuals within the household may have little or no impact in households with altruistic members or well-aligned preferences.

The current paper investigates the prevalence and heterogeneity of these within-household spillovers. We argue that individuals value the distribution of household resources rather than only their own share. This perspective accommodates a wide range of interdependencies, including those that stem from direct altruism or paternalism, maximin concerns, and fairness considerations. Our focus, however, is on *narrow* spillovers: between spouses within a specific domain of decision-making. In particular, we analyze equity separately in three domains: consumption, labor supply (or, equivalently, its complement: non-market time), and leisure. So, while Beckerian caring preferences imply broad distributional concerns, household members with narrow equity concerns may prioritize an equal distribution of resources within particular domains, such as the division of leisure between spouses. Some households may even be willing to accept inequality at the broader level in order to achieve what they perceive as a more suitable distribution within a particular domain. These domain-specific distributional concerns may further be shaped by gender norms in society, for instance, with respect to who engages in market work versus household chores (Bertrand et al., 2015).

The concept of distributional preferences is most useful when it is empirically tractable and capable of generating testable implications for existing data (Fehr et al., 2023). Building on this principle, the main objective of this paper is to examine not only the prevalence of narrow distributional concerns within households but also their consistency. Specifically, we test whether stable distributional preferences exist both within and across households. Within households, consistency requires that spouses share a common view on the desired distribution in a given domain. If they do not, changes in bargaining power could shift the household's equity concerns, making them dependent on contextual factors or budget constraint parameters. Across households, consistency requires that households with similar observable characteristics share common equity concerns. If variation instead reflects only idiosyncratic factors, then the concept of narrow distributional preferences offers little explanatory power for household allocation decisions. In this sense, our analysis echoes Fehr et al. (2023)'s call for parsimony in modeling heterogeneity in social preferences and extends it to the study of *narrow* distributional concerns *within households*.

This paper. We develop a conceptual framework for narrow[1] distributional considerations in families. In our framework, preferences are defined over the allocation of resources between the male and female spouse within each decision-making domain in the family. Each spouse derives utility from the allocation of consumption, non-market time (including leisure, chores and childcare) and leisure, represented by three distinct utility functions. We selected these three domains because they capture a range of equity concerns. For example, equal amounts of labor supply by the spouses does not necessarily imply equal leisure allocations, and vice versa. Given our focus on *narrow* equity, we assume that the distributional preferences within each domain are independent of allocations in other domains. At the same time, our framework imposes no prior structure on how the utilities from (equal) distributions within domains jointly contribute to overall welfare.

A distinguishing feature of our framework is that it also provides a more 'hybrid' perspective on preference heterogeneity within families. Spouses may display alignment of distributional preferences in some domains while exhibiting substantial divergence in others. For instance, the male and female spouse may share the same view regarding how much labor each should supply but diverge on how to allocate a given consumption budget. This contrasts with standard models of household decision-making, in which spouses either have identical preferences (the unitary model) or arbitrarily different preferences across all domains (the collective model). Our analysis provides a new framework to examine domain-specific preference alignment.

We apply our model to data from the Longitudinal Internet studies for the Social Sciences (LISS), a nationally representative survey of Dutch households comprising various studies. Cherchye et al. (2012b) initiated Assembled Study 34 ("*Time Use and Consumption*") to examine collective labor supply and household production—without consumption or leisure externalities—and used data from the first survey wave in their analysis. In contrast, we use data from three survey waves conducted in 2009, 2010, and 2012. The time use questions from this module offer separate measures of leisure and chores, enabling us to differentiate between distributional preferences regarding pure leisure and distributional views of total time away from paid work.

Additionally, our empirical set-up has two special features that distinguish it from existing applications. First, our analysis incorporates direct information on preference variation within the household, drawn from Core Study 5 ("*Family and Household*") of the LISS survey. This unique dataset contains self-reported differences of opinion between spouses,

[1]See Exley and Kessler (2024) for the concept of 'narrow equity' and Ellis and Freeman (2024); Fallucchi and Kaufmann (2021); Read, Loewenstein, and Rabin (2006) for the literature on 'narrow bracketing' in choice. Our use of the term aligns closest with Exley and Kessler (2024)'s.

elicited separately for each decision-making domain under consideration. The absence of reported differences of opinion is indicative of spousal preference alignment which, under the general collective model, is expected to attenuate the responsiveness of household choices to changes in bargaining power. Second, we increase cross-sectional price variation by constructing individual-specific Stone–Lewbel prices (Hoderlein and Mihaleva, 2008; Lewbel, 1989). Personalized consumption prices are computed as weighted averages of national price indices across product categories, using individuals' observed consumption shares as weights. Similarly, values of unpaid work (i.e., domestic productivity) are obtained as weighted averages of wages for comparable services, using individuals' time allocated to specific domestic activities as weights. Price and productivity data come from Statistics Netherlands (CBS).

We document substantial dispersion in within-household allocations across households within each domain, even though most households exhibit some form of spousal preference alignment. The main part of the paper examines this heterogeneity through the lens of our model with narrow distributional concerns. We show that the joint hypothesis of rationalizability with narrow distributional preferences and spousal alignment in a given domain yields testable conditions in the form of straightforward GARP restrictions applied to the observed allocations within that domain. The revealed preference conditions are robust to misspecification of the distributional utility functions and flexible enough to allow heterogeneous effects of distributional utilities on welfare. The results indicate that heterogeneity in prices, incomes, and household welfare indices is insufficient to explain the variation in observed allocations; accounting for the data additionally requires interhousehold heterogeneity in distributional concerns.

Common distributional concerns may nevertheless exist within certain subgroups of households. We thus compute the minimum number of distributional concerns required to eliminate all revealed preference violations. This yields four to seven distinct distributional types per domain. A natural next question is whether these different views on the desired allocation of (domain-specific) resources between the male and female spouse can be explained or predicted by observable factors. To address this question, we relate the identified heterogeneity to the observable variation in household characteristics following an empirical strategy recently proposed by Cherchye et al. (2024). The analysis considers demographic characteristics of individuals and households, their gender role attitudes, and broader views on diversity and inclusion. The approach allows us to examine whether narrow distributional concerns are driven by the same observed factors across the three decision-making domains. We find that the female spouse's view on gender roles is a powerful predictor of distributional concerns for labor, while concerns for other time uses are best explained by education. The presence of children is highly informative in both domains.

Related literature. There is extensive evidence of social preferences among individuals (Andreoni and Miller, 2002; Engelmann and Strobel, 2004; Fisman, Jakiela, and Kariv, 2015),[2] and it is reasonable to expect that distributional concerns are at least as significant between spouses in a couple. Yet few papers have explicitly studied distributional preferences within a family context. One challenge is that household-level decisions are the result of a complex interplay between spouses with (potentially) distinct preferences.[3] The collective model by Chiappori (1992) and Apps and Rees (1988) has become the workhorse to study the joint decisions of spouses. The current paper bridges both these literatures.

Two aspects in particular have received little attention in the household economics literature. First, while the collective model is consistent with broad caring preferences in the sense of Becker (1981), the literature typically shuts down direct spillovers of consumption/income, chores, or leisure between spouses. Exley and Kessler (2024) nonetheless show that in multi-dimensional outcome settings, equity considerations often apply narrowly within each outcome domain. In a household context, Berry, Dizon-Ross, and Jagnani (2025) study parental preferences for allocating investment across their children, and find a stronger preference for equalizing inputs (i.e., equality of opportunity) than for equalizing broader outcomes such as expected earnings. In this paper, we capture spillovers between spouses in the narrow sense by means of systematic distributional or social preferences per outcome domain, within the context of a collective model.

Second, the nature of household decision-making can vary strongly across domains of family outcomes.[4] Spouses' allocation decisions cover a wide range of qualitatively distinct domains, including consumption, chores, and leisure. Evidence from experiments shows that subjects are more willing to give up time rather than equivalent amounts of money (Brown et al., 2019; Ellingsen and Johannesson, 2009). Lilley and Slonim (2014) further find that the degree of substitutability between money and time donations depends on the intentions behind the donation. Thus, distributional preferences may also display substantial variation depending on the object of choice in the family. As a further consequence, spouses' distributional preferences may align better in some domains than in others. Given all this, we study the determinants of household distributional concerns for each domain separately.

Methodologically, our paper fits in the revealed preference literature, starting from Samuel-

[2] Fehr and Charness (2023) and Fisman et al. (2015) provide a recent discussion and overview of factors that may influence social preferences.

[3] Balafoutas, Kerschbamer, Kocher, and Sutter (2014) investigate the distributional preferences of individuals versus teams, albeit in a different allocation and group context than ours.

[4] Similarly, in the 'separate spheres' equilibrium proposed by Lundberg and Pollak (1993), the nature of decision-making can also vary across domains because each public good is provided exclusively by one spouse who acts as a dictator in that sphere. However, in Lundberg and Pollak (1993), this outcome is associated with a noncooperative equilibrium, while our analysis remains grounded in a collective approach.

son (1948) and developed further to test consistency of observed household behavior with unitary rationality (Afriat, 1967; Diewert, 1973; Varian, 1982) or collective rationality (Cherchye et al., 2012a, 2011) with gains from joint time (Cosaert et al., 2023). The main advantage of this approach in our setting is that it helps us test the consistency of observed behavior with our theoretical framework, without imposing ad-hoc structure on the distributional preferences. For this reason, the revealed preference toolkit has already been used to study the consistency of other-regarding preferences such as altruism and inequality aversion (Andreoni and Miller, 2002; Deb et al., 2014; Fisman et al., 2007). Our approach aligns moreover with other studies that compute the minimum number of 'types' necessary to rationalize all observed behavior in the sample (Castillo and Freer, 2018; Cherchye, Saelens, and Tuncer, 2024; Cosaert, 2017; Crawford and Pendakur, 2013). The latter papers focus on choices of consumption, time use or risky assets, rather than *distributions* of resources in a household context. Kerschbamer (2015) proposes a comprehensive classification of nine 'archetypes' of distributional preferences, as well as a nonparametric test to discriminate between them. However, this test focuses on individual types of prosocial attitudes rather than household types of distributional concerns for allocations between a male and female spouse.

Outline. The rest of the paper unfolds as follows. Section 2 describes the dataset and presents descriptive statistics for sample households and their within-household allocations. Section 3 lays out the theoretical framework with distributional concerns that hold narrowly in consumption, labor, and leisure. Section 4 outlines the implementation of the theoretical framework. Section 5 presents the results of our structural analysis, and Section 6 concludes.

2 Data and empirical evidence

This section documents empirical patterns in Dutch households that the theoretical framework aims to explain. We begin by describing the dataset, the sample selection procedure, and the characteristics of the households in our sample. We then study within-household allocations of consumption and time use, and conclude by examining the degree of spousal preference alignment per decision domain.

Survey data. This study draws on data from the Longitudinal Internet studies for the Social Sciences (LISS) panel, a comprehensive and nationally representative survey of Dutch households. The LISS panel has been conducted since 2007 and is based on a true probability sample drawn from the population register by Statistics Netherlands. It provides extensive

longitudinal data covering household characteristics, economic behavior, and individual preferences and opinions.

For this paper, we use data from three survey waves conducted in 2009, 2010, and 2012. The sample consists of married heterosexual couples aged 25 to 60, with or without children living at home. For households with children, we restrict the sample to those in which the youngest child is aged 18 or younger. We further limit the analysis to households in which both spouses participate in the labor market.[5] Although this restriction may introduce some selection bias, it is required for the structural analysis of Section 5, which relies on wage information for both partners.[6] Households in which at least one spouse does not participate in the labor market account for only 9.3% of the original sample, so this restriction does not substantially reduce the sample size. After applying all selection restrictions, the final sample consists of 661 households. A detailed description of the selection procedure is given in Appendix A.1.1. Our data cleaning process builds on the replication package provided by Cosaert et al. (2023), with extensions described below and in the 'Preference alignment' subsection.

Beyond demographic characteristics, we incorporate spouses' gender role attitudes and views on ethnic diversity as additional sources of variation in household distributions. LISS Core Study 8, Politics and Values, includes questions on gender roles in childcare and household responsibilities. For example, one statement reads: "*A child that is not yet attending school is likely to suffer the consequences if their mother has a job.*" Responses are recorded on a five-point scale ranging from (1) Fully disagree to (5) Fully agree. The module includes seven such items related to gender role stereotypes, listed in Appendix A.1.2. We construct a gender value index by summing each respondent's scores across these items, where higher values indicate more conservative gender role views.

The module also contains questions capturing attitudes toward ethnic diversity and inclusion. For instance, one statement reads: "*It is good if society consists of people from different cultures.*" Responses are again recorded on a five-point scale. There are six such items related to attitudes toward foreigners and cultural diversity, listed in Appendix A.1.3. We construct an index by averaging each respondent's scores across these items, where higher values reflect more accepting views toward cultural diversity.

Descriptive statistics. Table 1 reports the descriptive statistics. Panel A presents demographic characteristics of the households. The average age is approximately 43 years for men and 41 years for women. Educational attainment is similar between men and women.

[5] We refer to Cherchye et al. (2026) for a recent analysis of the welfare effect of labor market participation.

[6] It is worth to note that households with a non-employed spouse are also more likely to exhibit unequal bargaining positions and resource allocations.

About 38.7% of men and 34.9% of women hold a university or postgraduate degree. At the household level, the average number of children is 1.54, and the youngest child has an average age of approximately 9 years.

Panel B reports individual consumption and time use levels, with details on variable construction provided in Appendix A.1.4. On average, men and women exhibit similar levels of monthly private consumption, with greater variability observed for men. Furthermore, men work about 46.4 hours per week on average, while women work around 29.6 hours per week. These measures include commuting time. This difference is consistent with the high prevalence of part-time employment among Dutch women. In contrast, women devote substantially more time to unpaid domestic work. On average, they spend about 31.8 hours per week on childcare and household chores, compared to 18.5 hours per week for men. As a result, men have an average of 17.7 hours of leisure per week, while women enjoy slightly less leisure, at around 16 hours per week. Given our focus on the (within-household) externalities associated with private goods, we only include 'private' leisure activities *without* the spouse.[7]

Table 1: Summary statistics of married households

	Male spouse		Female spouse	
	mean	sd	mean	sd
Panel A. Demographics				
Age	42.95	7.72	40.94	7.58
Primary and below	2.4%	0.15	3.8%	0.19
Secondary degree	58.9%	0.49	61.2%	0.49
University and above	38.7%	0.49	34.9%	0.48
Panel B. Intrahousehold Allocation				
Private consumption (euro/month)	390.65	587.50	386.11	269.54
Market work (hrs/week)	46.39	10.15	29.60	11.80
Domestic work (hrs/week)	18.53	12.86	31.79	17.32
Leisure (hrs/week)	17.66	13.68	15.97	12.05
Panel C. Households				
	mean		sd	
Number of kids	1.54		1.12	
Age youngest child	8.98		5.53	
Observations	661			

We further compare the distribution of consumption and time use between wives and husbands in Figure 1. Each dot represents a household, where the wife's consumption or time level is plotted on the x-axis and the husband's level on the y-axis. The solid line

[7] Joint leisure with a spouse can be treated as a public good; see Cosaert et al. (2023) for an analysis of preferences for togetherness, based on the LISS data.

depicts the OLS regression fit, while the dashed line represents the 45-degree reference line. A positive slope of the solid line indicates a positive association between spouses' goods. The figures separately report allocations of private expenditure, leisure, and non-market time, where non-market time captures all the time away from paid work (including private leisure). Blue dots indicate that the husband has more resources or time than his wife, while red dots indicate the opposite.

The correlation in private expenditure is positive but modest, with a substantial share of women exhibiting higher private consumption shares. Husbands generally devote more time to leisure, and the within-couple correlation of private leisure is weak. By contrast, wives devote substantially more time to non-market activities than husbands, reflecting a greater share of time devoted to housework and childcare. Overall, the time use figures reveal a clear pattern of labor specialization in Dutch households, where women bear a larger share of housework and childcare responsibilities, while men allocate more time to market work.

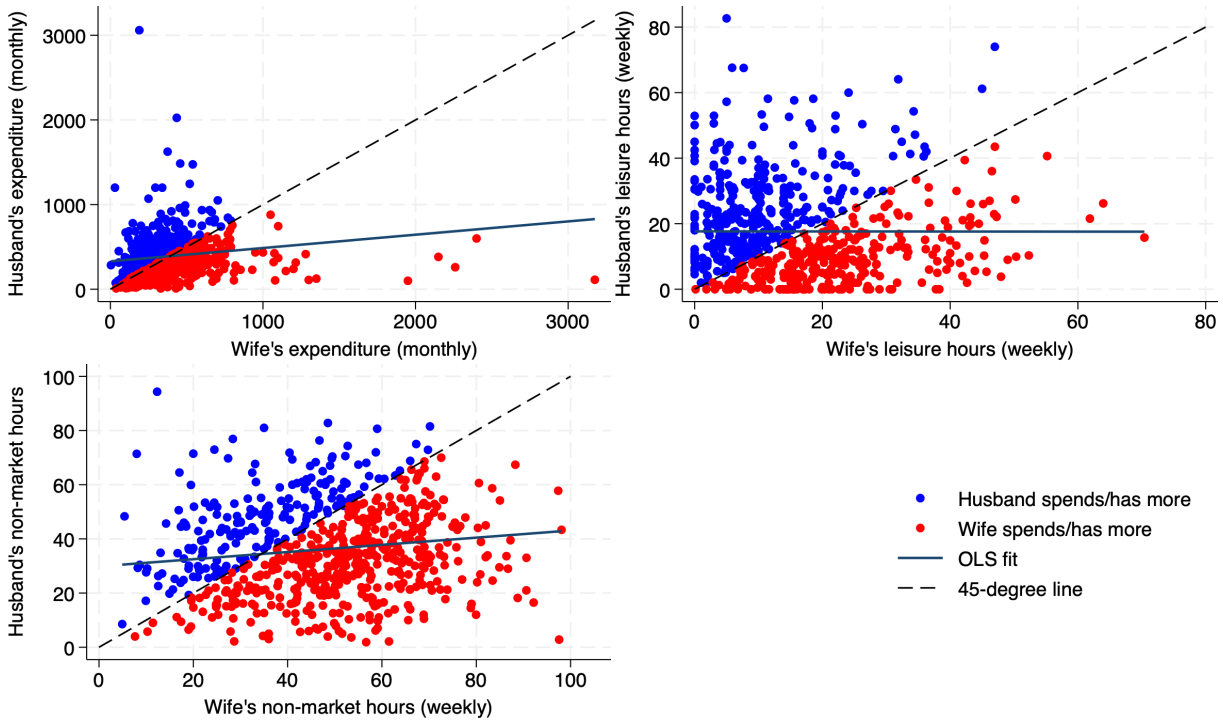

Figure 1: Expenditure and time use differences between spouses

In sum, we document two key findings on intra-household distributions. First, allocation patterns differ markedly across decision domains. Non-market time allocations exhibit greater inequality, whereas leisure and especially consumption allocations are more equalized between partners. Second, the wide dispersion of observed allocations within each domain points to substantial heterogeneity across households. In principle, this is not inconsistent

with egoistic versions of the collective labor supply model. Selfish spouses with distinct preferences, combined with substantial heterogeneity in bargaining power, can account for both findings. We next examine this explanation by focusing on within-household preference heterogeneity.

Preference alignment. To measure the degree of spousal preference alignment, we use survey questions from LISS Core Study 5: Family and Household. This module elicits direct information on within-household discussions through questions such as: "*Can you indicate whether you and your partner had any differences of opinion regarding the following issues, over the past year?*". The issues covered are *money expenditure*, *working (too much)*, and *who does what in terms of household work*. We map these responses to our three decision domains. Specifically, we infer alignment in consumption preferences (C) from responses to the *money expenditure* question. Differences of opinion regarding *working (too much)* reflect concerns about (insufficient) time away from paid work by either spouse (T), as time devoted to market and non-market activities sum up to total available time. Finally, we argue that responses to the *who does what in terms of household work* question are closely linked with the allocation of leisure (L) between spouses.[8]

Each household provides responses from both spouses, who answer the questions on a three-point scale: (1) Practically never, (2) Occasionally, and (3) Often. If both spouses respond with *Practically never* in a given domain, the household is categorized as having no differences of opinion in that domain. If at least one spouse responds with *Occasionally* or *Often*, the household is classified as having different views. This classification rule is applied uniformly across all three decision domains. We interpret the absence of reported differences of opinion as evidence that spouses' preferences, especially their distributional concerns, are aligned in that domain. While under the collective model members with different preferences ultimately do 'agree' on household decisions, the decision process arguably still brings the differences of opinion to the surface. This is why we view reported differences of opinion as consistent with—and even supportive of—bargaining under the collective model. This need for discussion and negotiation is expected to diminish when preferences are sufficiently aligned.[9]

We divide the sample based on whether spouses report aligned preferences in separate

[8] Conditional on labor supply, the division of household chores directly determines the amount of leisure available to each spouse. Moreover, household tasks may differ in their (dis)utility, and more enjoyable tasks may partially offset a lack of (pure) leisure.

[9] One alternative interpretation is that members have no explicit differences of opinion because one member behaves as a dictator. While we cannot rule this out, such an explanation is observationally equivalent to preference alignment. The underlying allocations could also be rationalized by a representative (unitary) distributional objective.

decision domains. In the sample, 12.1% of households exhibit fully aligned preferences across all domains, while 27.8% report complete misalignment. This pattern indicates that full spousal preference alignment is rare, consistent with evidence against unitary household models. Interestingly, complete misalignment is also uncommon, as most couples share similar preferences in at least one domain. This motivates a closer examination of the decision domains in which spouses exhibit aligned preferences.

Table 2 reports the share of households with spousal preference alignment in each decision domain.[10] Spousal preference alignment is least pronounced in leisure and most pronounced in the domain of labor, indicating a higher likelihood of shared views within the household on working (too much) by either spouse. Specifically, in the division of time between non-market activities and market labor, about 48% of households report aligned preferences, whereas 52% express differing concerns. For the further division of non-market time between leisure and housework, only 33% of households report aligned preferences and 67% report differences of opinion. Finally, around 41% of households exhibit aligned preferences in consumption, while the remaining 59% report differences of opinion over expenditure.

Table 2: Spousal preference alignment across decision domains

Domain	Number of Households	Percentage Aligned
Consumption	270	40.85%
Non-market Time	318	48.11%
Leisure	220	33.28%
Total	661	100.00%

To examine the link between household allocations and spousal preference alignment, Table 3 reports allocations between spouses for households with aligned and unaligned preferences, within each decision domain. The levels of private expenditure do not differ significantly between spouses, regardless of preference alignment in the household. In contrast, wives have substantially more non-market time than husbands in all households. For leisure, husbands generally have more time than wives. Differences in leisure are statistically insignificant only in households with aligned preferences in leisure and in households that report differences of opinion in working (too much).

We further connect spousal preference alignment to household demographics and observed allocations using a linear probability model. The estimation results indicate which types of couples are more likely to exhibit aligned preferences in each decision domain, as well as which aspects of within-household distributions are least prone to friction. The results

[10] These classifications are not mutually exclusive, as a couple may be aligned in one domain but misaligned in another.

Table 3: Spousal distribution by preference (mis)alignment

	Avg Expenditure			Avg Non-market Time			Avg Leisure		
	Male	Female	Diff	Male	Female	Diff	Male	Female	Diff
C-aligned	425.70	389.31	.	36.25	46.77	***	18.67	16.61	*
C-unaligned	366.45	383.89	.	36.14	48.44	***	16.96	15.52	*
T-aligned	362.37	392.01	.	37.01	47.39	***	18.78	16.52	**
T-unaligned	416.87	380.63	.	35.42	48.10	***	16.62	15.46	.
L-aligned	423.82	408.57	.	35.70	47.99	***	17.09	16.21	.
L-unaligned	374.10	374.90	.	36.43	47.65	***	17.94	15.85	**

Notes: *** 1%, ** 5%, * 10%. A dot indicates a p-value $> 10\%$. Non-market time and leisure are measured in hours per week. Expenditures are spouses' private consumption measured in euros per month. Spousal preferences are aligned or unaligned in consumption (C), non-market time (T), and leisure (L).

should be interpreted as correlations rather than causal effects. The dependent variable equals 1 if spousal preferences are aligned in a given domain and 0 otherwise. The regressors include spouses' educational attainment and age, as well as the observed allocation in the same domain. Educational attainment in LISS is coded on a six-point scale: 1–primary or below; 2–intermediate secondary; 3–higher secondary; 4–vocational education; 5–university; and 6–postgraduate. Higher values indicate higher levels of education. Time use variables, including market work and household chores, are expressed in monthly hours for consistency with expenditure measures.

Table 4 shows that households in which the female spouse is at a later life-cycle stage are more likely to report aligned preferences across all decision domains, consistent with preference convergence over longer marriage durations. By contrast, husbands' age is negatively related to alignment in non-market time. This finding is consistent with a stronger attachment to the labor market among husbands with longer work experience. The number of children is negatively associated with the likelihood of preference alignment in both consumption and non-market time. This pattern suggests that having more children increases pressure on household spending and market work intensity, which in turn amplifies discussion between spouses. A different pattern emerges for leisure. Spousal alignment in leisure is significantly associated with wives' stated gender role attitudes. In particular, wives with more traditional views are more likely to report alignment with their spouse in the division between leisure and household chores. Husbands' reported attitudes are not statistically significant. One possible explanation is that wives with more traditional gender role views are more inclined to defer to their husbands' preferences in decisions over housework. Finally, we find that the observed levels of consumption and housework are not significantly associated with spousal preference alignment. By contrast, longer market work hours for both spouses

are associated with reduced alignment in the labor domain, possibly reflecting discussion over excessive market work.

What do we learn from this? To conclude, the empirical evidence in this section shows distinct distributional patterns in consumption, labor, and leisure, alongside substantial heterogeneity across households. While some of this variation could be rationalized by an egoistic model with strong spousal preference variation and heterogeneous bargaining power, the data on spouses' self-reported differences of opinion are less supportive of this explanation. In particular, we show that the degree of spousal preference alignment varies across decision domains, and that overall alignment is substantial: 72.2% of households exhibit within-household preference alignment in at least one domain. Taken together, these findings call for a theoretical framework that accommodates both domain-specific distributional preferences and domain-specific preference alignment. In the next section, we introduce a novel extension of the collective model that incorporates these features. Allowing (narrow) interdependent preferences, the framework captures a wide range of behaviors in a reduced form way, including altruistic behavior, gendered distributions driven by social norms, and other forms of within-household externalities that depend on the decision domain.

3 Theoretical framework

We now present a theoretical framework that incorporates distributional preferences and (partial) preference alignment into a non-unitary model of household decision-making, moving beyond the typical egoistic version of collective labor supply models. In our model, social or distributional preferences capture household members' views on the desired allocation of money and time between the male and female spouse. We focus on narrowly defined distributional preferences in three domains: consumption, labor, and leisure. This allows us to address potentially distinct distributional concerns in each domain, complementing existing research that focuses on the overall division of non-labor income (the sharing rule).

We first introduce three subutility functions per spouse to represent their domain-specific distributional preferences. The utility from consumption $U_C^i(c^1, c^2)$ describes spouse i's preferred allocations of private consumption between spouse 1 (the husband) and spouse 2 (the wife).[11] The utility function from leisure $U_L^i(l^1, l^2)$ represents i's preferences for the distribution of private leisure between the husband and wife. Finally, the utility from non-market time $U_T^i(t^1, t^2)$ describes i's preferred combinations of time away from paid labor by

[11] Other forms of (household-level) expenditure x that cannot be assigned to a specific spouse will be captured in a household commodity q, discussed below.

Table 4: Regressions of spousal preference alignment

	(1) Consumption	(2) Non-market Time	(3) Leisure
Age of husband	-0.002	-0.014**	-0.003
	(0.006)	(0.006)	(0.006)
Age of wife	0.012**	0.011*	0.012**
	(0.006)	(0.006)	(0.006)
Education of husband	0.015	-0.016	-0.007
	(0.015)	(0.016)	(0.015)
Education of wife	0.011	0.002	0.013
	(0.016)	(0.017)	(0.016)
Gender role attitude of husband	0.004	0.004	-0.004
	(0.005)	(0.005)	(0.005)
Gender role attitude of wife	0.002	-0.007	0.016***
	(0.005)	(0.005)	(0.005)
Number of children	-0.045**	-0.039**	-0.007
	(0.018)	(0.019)	(0.020)
Expenditure of husband	0.000*		
	(0.000)		
Expenditure of wife	-0.000		
	(0.000)		
Market work of husband		-0.001**	
		(0.000)	
Market work of wife		-0.001**	
		(0.000)	
Chores time of husband			0.000
			(0.000)
Chores time of wife			-0.000
			(0.000)
Constant	-0.106	1.131***	-0.226
	(0.183)	(0.222)	(0.175)
Observations	661	661	661

Notes: Standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Time-use regressors are measured in hours per month to align with the monthly consumption data. Using weekly hours yields the same significance levels. Estimation using a Probit model also yields the same significance levels.

both spouses.[12] This reflects a disutility of market work—beyond intrinsic preferences for leisure—that may stem from individual (e.g., cost of effort, lack of control) or distributional (e.g., spouse's work stress, gender norms discouraging female labor supply) considerations. Seen together, these utility functions can capture a wide range of domain-specific externalities within the household, including altruism and fairness considerations as well as desired distributions driven by gender roles.

Moreover, men's and women's distributional preferences may sometimes align, limiting the role of intra-household bargaining. Our framework explicitly accounts for this by allowing spousal preferences to coincide in some domains—such as how to allocate money—while differing in others—like who does what in terms of housework. We make this more explicit in Definition 2 below. The main assumption in our framework is that each subutility function U_d^i, with $d \in \{C, T, L\}$, is monotone and concave in its arguments,[13] and independent of the allocations in other domains $o \neq d$.

The formal household objective function is

$$\begin{aligned}&\delta_C^1 U_C^1(c^1, c^2) + \delta_T^1 U_T^1(t^1, t^2) + \delta_L^1 U_L^1(l^1, l^2)\\ &+\delta_C^2 U_C^2(c^1, c^2) + \delta_T^2 U_T^2(t^1, t^2) + \delta_L^2 U_L^2(l^1, l^2)\end{aligned}$$

where each coefficient δ_d^i determines the relative weight in the decision process of spouse i's preference for allocations in the domain d. Notice that this includes the possibility of dictatorship in specific outcome domains. Furthermore, the linear combination of distributional utilities is less restrictive than it seems. The maximization of this objective is equivalent to the maximization of $W(U_C^1, U_T^1, U_L^1, U_C^2, U_T^2, U_L^2)$ as long as the weights δ can vary adequately, with $\frac{\delta_d^i}{\delta_o^j} = \frac{\partial W/\partial U_d^i}{\partial W/\partial U_o^j}$. This also shows that consumption utilities U_C are not necessarily separable from non-market time or leisure utilities U_T, U_L in the overall welfare assessment. Finally, the welfare index (or equivalently, the weights δ) can depend on household characteristics, contextual factors, as well as budget constraint parameters. This distinguishes the framework from a standard unitary model. A special case obtains when $W = \delta^1 W^1 + \delta^2 W^2$ with $W^i = \alpha_C^i U_C^i + \alpha_T^i U_T^i + \alpha_L^i U_L^i$, and so that δ^1, δ^2 can depend on budget constraint parameters while $\alpha_C^i, \alpha_T^i, \alpha_L^i$ cannot. In this sense, the framework nests a classical collective model with non-selfish preferences in each domain. In our following application, we leave

[12] For couples with children, t^1 and t^2 contain not only private leisure and household chores but also time devoted to children.

[13] While the monotonicity assumption rules out some classic types of distributional preferences such as inequality/difference aversion, envy, or spite, it is consistent with the concerns of both selfish agents and altruistic agents, Rawlsian/Leontief/Maximin preferences, as well as benevolence towards agents who are better off ('kiss-up'). We refer to Kerschbamer (2015) for a detailed discussion of these distributional preference archetypes.

$\delta_C^1, \delta_T^1, \delta_L^1, \delta_C^2, \delta_T^2, \delta_L^2$ completely free to avoid unnecessary homogeneity assumptions across households.

We now construct the household's budget constraint in terms of consumption (c), non-market time (t), and leisure (l). Let p_C^1 and p_C^2 denote the prices of consumption by husbands and wives, respectively,[14] and let w_M^1 and w_M^2 denote their individual wages in the labor market. We assume that a domestic commodity q can be produced by the household (with members' domestic work time h^1 and h^2 as inputs) as well as purchased in the market with expenditure x, so that $q = x + f(h^1, h^2)$.[15] Each household has an exogenous lower bound on the provision of this (marketable) commodity, $q \geq q^*$. For simplicity, we further specify that $f(h^1, h^2) = w_H^1 h^1 + w_H^2 h^2$ is a linear function where w_H^1 and w_H^2 denote the husband's and wife's individual average productivity of domestic work. The non-labor income is ω and the husband and wife's market work is m^1 and m^2, respectively. The budget constraint is derived as follows,

$$\begin{aligned} & p_C^1 c^1 + p_C^2 c^2 + x \leq \omega + w_M^1 m^1 + w_M^2 m^2 \\ \Leftrightarrow\ & p_C^1 c^1 + p_C^2 c^2 + q^* - w_H^1 h^1 - w_H^2 h^2 \leq \omega + w_M^1 m^1 + w_M^2 m^2 \\ \Leftrightarrow\ & p_C^1 c^1 + p_C^2 c^2 - w_H^1 h^1 - w_H^2 h^2 - w_M^1 m^1 - w_M^2 m^2 \leq \omega^* \end{aligned}$$

where ω^* in the last line captures the household's disposable non-labor budget. This is the difference between the true non-labor income and the lower bound on the provision of the household commodity: $\omega^* = \omega - q^*$.

We also denote $t^i = \mathcal{T} - m^i$ and subsequently $l^i = t^i - h^i$ to substitute for m^i and h^i (with $\mathcal{T}$ the total time available):[16]

$$\begin{aligned} & p_C^1 c^1 + p_C^2 c^2 - w_H^1 (t^1 - l^1) - w_H^2 (t^2 - l^2) - w_M^1 (\mathcal{T} - t^1) - w_M^2 (\mathcal{T} - t^2) \leq \omega^* \\ \Leftrightarrow\ & p_C^1 c^1 + p_C^2 c^2 + w_H^1 l^1 + w_H^2 l^2 + (w_M^1 - w_H^1) t^1 + (w_M^2 - w_H^2) t^2 \leq \omega^* + (w_M^1 + w_M^2)\mathcal{T} \end{aligned}$$

Finally replacing $p_T^i = (w_M^i - w_H^i)$ and $p_L^i = w_H^i$, this gives the following budget constraint:

$$p_C^1 c^1 + p_C^2 c^2 + p_T^1 t^1 + p_T^2 t^2 + p_L^1 l^1 + p_L^2 l^2 \leq \omega^* + (w_M^1 + w_M^2)\mathcal{T}$$

[14] The consumption prices are individual-specific to account for differences in the composition (hence cost) of each member's personal consumption bundle. We refer to Section 4.1 for details.

[15] This formulation fits in the literature on marketable home production starting from Gronau (1977).

[16] In our application, $\mathcal{T}$ is the sum of private leisure, household chores, and market work.

The household then solves,

$$\begin{aligned} &\max \delta_C^1 U_C^1(c^1, c^2) + \delta_T^1 U_T^1(t^1, t^2) + \delta_L^1 U_L^1(l^1, l^2) \\ &+\delta_C^2 U_C^2(c^1, c^2) + \delta_T^2 U_T^2(t^1, t^2) + \delta_L^2 U_L^2(l^1, l^2) \\ &s.t. \\ &p_C^1 c^1 + p_C^2 c^2 + p_T^1 t^1 + p_T^2 t^2 + p_L^1 l^1 + p_L^2 l^2 \leq \omega^* + (w_M^1 + w_M^2)\mathcal{T} \end{aligned} \tag{1}$$

To sum up, our model offers a theoretical framework to incorporate distributional preferences into the analysis of (non-unitary) household decisions. By construction, these distributional preferences introduce within-household externalities with respect to consumption, non-market time, and leisure. A decrease in consumption for one spouse can have negative spillovers for the other spouse. Similarly, the individual may be willing to pay to increase private leisure and alleviate time stress of the spouse.

For our following structural analyses, we adopt an empirical set-up with just one observation per household s. One observation includes information on the distribution of consumption, non-market time, and leisure within that household. We also obtain household- and individual-specific prices following a procedure described in Section 4.1. Let D denote the dataset with the observed allocations and constructed prices for households in the sample. Definition 1 formalizes the concept of rationalizability with narrow distributional concerns. Specifically, D is rationalizable if there exist coefficients $\delta_{C,s}^i, \delta_{T,s}^i, \delta_{L,s}^i$ (for each household s in the data, and for each gender i) and subutility functions U_C^i, U_T^i, U_L^i (for each gender i) so that each household's choices solve program (1). A special feature of this definition is that it imposes minimal prior restrictions on the functional form of the distributional utility functions. This is particularly appealing in our context, as so far little is known about the types of narrow distributional concerns within the household. Moreover, full flexibility of the δ coefficients across households accounts for different contributions of a given distributional assessment U_d^i to the households' (situation-specific) objectives. Differences in household allocations can thus be explained by heterogeneity in prices, income, or the household-specific aggregation of the distributional utilities. At this point, the main consistency requirement is that the households $s \leq S$ share a common set of utility functions $U_C^1, U_C^2, U_T^1, U_T^2, U_L^1, U_L^2$ representing preferences for within-household distributions. This then indicates that specific distributional preferences are prevalent and stable across (groups of) households.

Definition 1 Consider a dataset $D = \{(p_{C,s}^1, p_{C,s}^2, p_{T,s}^1, p_{T,s}^2, p_{L,s}^1, p_{L,s}^2), (c_s^1, c_s^2, t_s^1, t_s^2, l_s^1, l_s^2)\}_{s \leq S}$. The data is rationalizable by a collective model with narrow distributional preferences if there exist household- and gender-specific coefficients $\delta_{C,s}^1, \delta_{C,s}^2, \delta_{T,s}^1, \delta_{T,s}^2, \delta_{L,s}^1, \delta_{L,s}^2$ and gender-specific distributional preferences $U_C^1, U_C^2, U_T^1, U_T^2, U_L^1, U_L^2$ so that the data solve (1).

Definition 1 is still general in the sense that it does not impose preference alignment between the spouses in a household. So, even if the distributional preferences of men U_C^1 and women U_C^2 are *not* household-specific, the overall distributional concern can be household-specific because the aggregate objective $\delta_{C,s}^1 U_C^1 + \delta_{C,s}^2 U_C^2$ depends on s. This then implies that the household's distributional concerns become sensitive to changes in policy parameters, such as wages or non-labor income, that may influence $\delta_{C,s}$. The main aim of the current paper is to test, in each of the three decision-making domains, if there exists a common *household-level* distributional concern that is independent of these policy parameters or other contextual factors. Consistency requires not only stability across households but also preference alignment between spouses. Definition 2 formalizes this notion of a representative (unitary) distributional objective at the household level, for each domain. By this definition, alignment refers to the phenomenon where the marginal rates of substitution between male and female outcomes coincide, rather than the requirement that both spouses have the same prosocial attitudes. For instance, the household's views on the desired distribution between the male and female spouse may be well aligned if one individual is fully selfish and the other is fully altruistic (since in this special case: $U_C^1(c^1) = U_C^2(c^1) = U_C(c^1)$). By contrast, in households where each individual derives their utility exclusively from the spouse's consumption, preferences are not aligned according to Definition 2 although both members are very altruistic ($U_C^1(c^2) \neq U_C^2(c^1)$).

It is worth to note that the spousal preferences may align more closely in some domains than in others. For instance, spouses may share the same view on the amount of labor supply by each spouse ($U_T^1 = U_T^2$) but diverge on the within-household distribution of chores/leisure time ($U_L^1 \neq U_L^2$). This could arise because labor supply is more publicly observable and therefore more tightly constrained by prevailing gender norms and institutional expectations. From a theoretical perspective, this framework explores a middle ground between the classical unitary and collective models, characterizing various classes of family decision-making based on the degree of spousal preference alignment in each of the three domains.

Definition 2 Consider a dataset $D = \{(p_{C,s}^1, p_{C,s}^2, p_{T,s}^1, p_{T,s}^2, p_{L,s}^1, p_{L,s}^2), (c_s^1, c_s^2, t_s^1, t_s^2, l_s^1, l_s^2)\}_{s \leq S}$.

1. There is a common household distributional concern in the consumption domain if

   (a) Spousal preferences for consumption align: $U_C^1(c_s^1, c_s^2) = U_C^2(c_s^1, c_s^2)$, *and*

   (b) D is rationalizable by a collective model with narrow distributional preferences.

2. There is a common household distributional concern in the labor domain if

   (a) Spousal preferences for non-market time align: $U_T^1(t_s^1, t_s^2) = U_T^2(t_s^1, t_s^2)$, *and*

(b) D is rationalizable by a collective model with narrow distributional preferences.

3. There is a common household distributional concern in the leisure domain if

(a) Spousal preferences for leisure align: $U_L^1(l_s^1, l_s^2) = U_L^2(l_s^1, l_s^2)$, *and*

(b) D is rationalizable by a collective model with narrow distributional preferences.

We use revealed preference methods to bring the model to the data and test for common household distributional concerns. An appealing feature of our framework is that the existence of common household distributional concerns translates into convenient and fairly standard testable conditions. We summarize the revealed preference characterizations in Proposition 1. The first statement of the proposition shows that any dataset D can be rationalized in the sense of Definition 1. This is a particular application of a result in Cherchye et al. (2012a), namely that data always pass the collective model when there are (not more than) two goods in the utility function and both are purely public. Note that each consumption or time use variable in our set-up acts as a public good because of its corresponding spillovers or externalities. The second statement of the proposition shows that the presence of common household distributional concerns in the sense of Definition 2 is equivalent to the data passing the Generalized Axiom of Revealed Preference (GARP) in a given domain. The proof is shown in Appendix A.2. The tests are both necessary and sufficient for the existence of common household distributional concerns.

Proposition 1 Consider a dataset $D = \{(p_{C,s}^1, p_{C,s}^2, p_{T,s}^1, p_{T,s}^2, p_{L,s}^1, p_{L,s}^2), (c_s^1, c_s^2, t_s^1, t_s^2, l_s^1, l_s^2)\}_{s \leq S}$.

1. D is rationalizable by a collective model with narrow distributional preferences.

2. Moreover,

(a) there is a common household distributional concern in the consumption domain if and only if $D_C = \{(p_{C,s}^1, p_{C,s}^2); (c_s^1, c_s^2)\}_{s \leq S}$ satisfies GARP.

(b) there is a common household distributional concern in the labor domain if and only if $D_T = \{(p_{T,s}^1, p_{T,s}^2); (t_s^1, t_s^2)\}_{s \leq S}$ satisfies GARP.

(c) there is a common household distributional concern in the leisure domain if and only if $D_L = \{(p_{L,s}^1, p_{L,s}^2); (l_s^1, l_s^2)\}_{s \leq S}$ satisfies GARP.

4 Operationalization of the theoretical framework

This section describes how we combine several state-of-the-art methods to implement our framework. We first construct spouses' personalized prices for private consumption and the opportunity costs of their non-market time and leisure—key inputs for applying the Generalized Axiom of Revealed Preference (GARP). Our pricing procedure builds on the logic of Stone–Lewbel prices (Hoderlein and Mihaleva, 2008; Lewbel, 1989). Next, we outline the revealed preference approach of Crawford and Pendakur (2013) and Cosaert (2017), which we use to quantify fundamental but unobserved heterogeneity. In our context, heterogeneity is summarized by the minimum number of distinct distributional types needed to rationalize observed behavior—among couples reporting aligned preferences. Finally, we examine the source of this heterogeneity using the method proposed by Cherchye et al. (2024). This identifies the household characteristics that best predict variation in household distributional concerns.

4.1 Price construction

A key input in the revealed preference procedure described in Section 3 is the set of prices for consumption (p_C^i), non-market time (p_T^i), and leisure (p_L^i). These prices are individual-specific and reflect the cost of allocating additional consumption, non-market time, or leisure to the male ($i = 1$) or female ($i = 2$) spouse. In our framework, prices may differ between spouses for several reasons, including differences in the composition of their consumption bundles, the mix of underlying household activities, and wages in the labor market. In addition, prices vary across households, and this cross-household variation helps identify heterogeneity in distributional concerns.

We construct personalized consumption prices (p_C^i) based on national price indices and the shares of the different goods within that person's consumption bundle. Our procedure connects data on consumer price indices obtained from Statistics Netherlands (CBS) with detailed information on the composition of personal bundles from the LISS. Let P^g denote the CPI of product category g in the year under consideration. We define the quantity of product g consumed by household member i as the ratio of their observed expenditure on g to the corresponding CPI, $c^{ig} = e^{ig}/P^g$. We then construct individual-specific consumption prices p_C^i as follows,

$$p_C^i = \sum_g \frac{c^{ig}}{c^i} P^g$$

where $c^i = \sum_g c^{ig}$, and we assume that the individual-specific consumption shares c^{ig}/c^i are

fixed proportions.[17] This also guarantees that the demands are homothetically separable, which is necessary to construct these cross section prices of the Stone-Lewbel type (Hoderlein and Mihaleva, 2008; Lewbel, 1989).

We further calculate w_M^i by dividing each individual's labor income by their hours worked (shown in Appendix A.1.4). To obtain the remaining prices p_T^i and p_L^i, we first derive a measure for the average productivity of each individual in domestic tasks (w_H^i). To this end, our procedure now combines nominal hourly wages from CBS with the time allocation across domestic activities reported in the LISS survey. Let W^g represent the nominal hourly wages in various industries g, such as food preparation and financial services, reported in CBS. Furthermore, let h^{ig} denote the time spent by individual i on domestic task g. We then construct individual-specific prices of domestic time as follows:

$$\bar{w}_H^i = \sum_g \frac{h^{ig}}{h^i} W^g$$

This valuation of unpaid work fits in the replacement cost approach, which values household production using the market wages demanded for equivalent services. The method is widely used in labor and family economics; one recent example is Jokubauskaitė and Schneebaum (2022). We also take into account that a regular household member cooking or doing administrative tasks is less productive than their specialist counterparts (e.g., a trained cook or accountant) in the market. We therefore scale domestic productivity $\tilde{w}_H^i = \alpha \bar{w}_H^i$ by a factor α, with α set to 0.5 in our baseline analysis.[18]

Table 5 summarizes the individual-specific consumption prices p_C^i, wages w_M^i, opportunity costs of leisure $p_L^i = w_H^i$ and opportunity costs of time away from work $p_T^i = (w_M^i - w_H^i)$. The index 1 corresponds to male household members, while the index 2 refers to female household members. Details on the CPI (with base year 2015 = 1), nominal wage rates, and the corresponding commodity bundles and activity lists are provided in Appendix A.3.

[17] In other words, each individual consumes their goods in constant ratios. This arises when individuals have Leontief utility functions with different weight vectors across product categories. Although this imposes strong simplifying restrictions on lower-level budgeting behavior, the higher-level distributional preferences in the consumption domain can be arbitrarily complex.

[18] By setting $\alpha = 0.5$, we simultaneously correct for specialist labor and for income taxes included in the nominal hourly wages from CBS. For comparison, Landefeld et al. (2009) use $\alpha = 0.75$ as a correction for specialist labor only. Moreover, as shown in Appendix A.5.2, lower or higher levels of α do not seem to affect our main conclusions. Finally, we impose an upper bound of $0.9 \times w_M^i$ on domestic productivity ($w_H^i = \min(\tilde{w}_H^i, 0.9 w_M^i)$) so that $w_H^i \leq w_M^i$.

Table 5: Constructed prices and wages

	Mean	SD	Min	Max
p_C^1	0.95	0.032	0.76	1.03
p_C^2	0.96	0.024	0.84	1.01
w_M^1	11.97	4.947	0.72	51.28
w_M^2	11.22	7.213	1.80	161.54
w_H^1	8.27	1.540	0.65	12.40
w_H^2	7.95	1.680	1.62	11.35
p_T^1	3.70	4.424	0.07	40.84
p_T^2	3.27	6.880	0.18	153.90

4.2 Capturing unobserved heterogeneity of distributional concerns

In the following section, we test the homogeneity of (narrow) distributional concerns across households, focusing on those that report spousal preference alignment within a given decision domain. Beyond within-household alignment, these households must satisfy the collective model with narrow distributional concerns that are common across households. Proposition 1 in Section 3 shows that this model implies nonparametric, testable restrictions. Consider two households, s and v, whose spouses' distributional preferences are aligned in consumption. The proposition requires that their observed consumption allocations pass the Generalized Axiom of Revealed Preference (GARP). A violation occurs if household s reveals a preference for its own allocation when that of v was also affordable, $p^1_{C,s}(c^1_s - c^1_v) + p^2_{C,s}(c^2_s - c^2_v) \geq 0$, and household v simultaneously reveals the reverse preference, $p^1_{C,v}(c^1_v - c^1_s) + p^2_{C,v}(c^2_v - c^2_s) \geq 0$, with at least one strict inequality. In this case, the revealed preference test rejects common distributional concerns for consumption across the two households, and we set a dissimilarity indicator $I_{sv} = 1$.

The natural next question is how many distinct distributional concerns are needed to rationalize the observed within-household allocations in any given decision domain. To address this question, we follow an approach laid out by Crawford and Pendakur (2013) and Cosaert (2017). These authors show that the minimum number of utility functions is equivalent to the minimum number of data partitions necessary to eliminate all revealed preference violations. In our setting, each partition of households can then be rationalized by a distinct distributional concern, which we term a distributional type. Implementation proceeds in two steps. First, we construct the matrix I of pairwise dissimilarity indicators. Second, we solve the linear program in Proposition 2, which identifies the minimum number of distributional types, $\hat{\tau}$, in the subsample. We refer to Cherchye et al. (2024) for a detailed discussion of this program.

Proposition 2 Consider the difference matrix I. The minimum number of types $\hat{\tau}$ is the solution to the following program, with variables $X_{sn} \in \{0,1\}$ and $Y_n \in \{0,1\}$:

$$\hat{\tau} = \min \sum_n Y_n$$

$$s.t.$$

$$\forall s : \sum_n X_{sn} = 1$$

$$\forall (s, v : s \neq v), \forall n : \text{if } I_{sv} = 1 \text{ then } X_{sn} + X_{vn} \leq 1$$

$$\forall s, \forall n : X_{sn} \leq Y_n$$

Technically, $X_{sn} = 1$ indicates that household s is added to type n, and $Y_n = 1$ indicates that type n is active. Each household is assigned to exactly one type (first condition), households with revealed distributional differences cannot belong to the same type (second condition), and households can only be assigned to 'active' types (third condition). The estimator $\hat{\tau}$ is then defined as the minimum cardinality of the set of active types. Intuitively, greater heterogeneity in distributional concerns typically requires a larger number of distinct types.

4.3 Explaining heterogeneous distributional concerns with observables

In the following section, we will also try to explain the identified heterogeneity of distributional concerns based on observable household characteristics. To investigate the heterogeneity between households, we follow a recent proposal by Cherchye et al. (2024). The authors develop an intuitive measure of how well one or several observed characteristics describe preference variation in the sample. We adjust their methodology to our set-up with distributional concerns. Specifically, consider an observed household characteristic k that splits the data in Z^k partitions. Let $\tau^{k,z}$ denote the minimum number of types in the zth partition of household characteristic k. Then index κ^k in Definition 3 measures how well characteristic k separates households with different distributional concerns.

Definition 3 (κ^k-ratio)

$$\kappa^k = \frac{\tau^k}{\hat{\tau}},$$

$$\tau^k = \sum_{z=1}^{Z^k} \tau^{k,z}.$$

According to this definition, lower values of κ^k indicate that characteristic k is a better predictor of household distributional concerns. To see this, suppose that a partitioning in Z^k subgroups eliminates all the revealed preference violations in the sample. If this is possible with the smallest number of partitions (i.e., $Z^k = \hat{\tau}$) then $\kappa^k = 1$. Suppose on the other hand that after partitioning based on k, the identified heterogeneity within each partition is as large as the heterogeneity of distributional concerns originally identified in the full sample. This creates a total of $\tau^k = Z^k\hat{\tau}$ types, resulting in $\kappa^k = Z^k$. In theory, κ^k is situated between 1 and Z^k, and lower levels indicate that observed characteristic k is a better criterion to split the sample in homogeneous subsets of households.

5 Structural analysis

This section addresses two main questions using a structural nonparametric approach. First, we assess how many types of distributional concerns are required to rationalize household behavior within each decision domain. Second, we investigate whether, and to what extent, variation in distributional concerns can be explained by observable household characteristics.

Identified heterogeneity. We begin by testing, for each decision domain, whether observed household allocations can be rationalized by a single distributional concern that is common across households. We focus on households in which spouses report no differences of opinion within the relevant domain. Proposition 1 shows that this household behavior is consistent with a common (and unitary) distributional objective in a given domain if and only if the observed within-household allocations pass GARP. This then indicates the existence of a strong, common view on within-household allocations in that domain (e.g., stemming from universal fairness considerations and/or systematic gender norms) independent of households' specific characteristics or budget constraint parameters. Applying the GARP test to each subsample, we find violations across all three decision domains. This suggests that a single household distributional concern per domain is insufficient to rationalize the observed allocations, pointing instead to interhousehold heterogeneity in views about what constitutes a desirable allocation between the male and female spouse.

To quantify the extent of this heterogeneity, we apply the algorithm introduced in Section 4.2. For each domain, we compute the minimum number of distinct distributional types required to rationalize observed behavior among households with aligned preferences, using the procedure in Proposition 2. Table 6 reports the results. The second column presents the minimum number of types, with each type corresponding to a distinct household distributional concern. The third column reports the number of households exhibiting spousal

preference alignment in each domain, and the fourth column shows the ratio of these households to the minimum number of types. We interpret this ratio as a measure of distributional homogeneity, with higher values indicating that larger groups of households share common views on the desired allocation between the male and female spouse, in the decision domain under consideration.

Households exhibit the highest level of homogeneity in their distributional concerns for consumption, while they display the greatest heterogeneity in leisure. Specifically, among 270 households with spousal preference alignment in consumption, we identify four distinct types. Approximately every 68 households, a new distributional utility function must be introduced to account for all the within-household consumption allocations in this subsample. Next, the 220 households with spousal preference alignment in the leisure domain belong to (at least) six distinct types. In this case only about 37 households can share a common distributional concern for the allocation of leisure. The labor domain falls between these two cases, with approximately every 45 households exhibiting different distributional concerns. This corresponds to a total of seven distinct types required to eliminate revealed preference violations among 318 households with spousal preference alignment in the labor domain.

Although this exploratory analysis reveals different numbers of types across decision domains, we are careful not to attribute these differences solely to distributional views being more universal in some domain (e.g., consumption). The strength of our nonparametric identification procedure may also differ across domains depending on the available price variation, so we avoid drawing strong conclusions from direct comparisons. What is clear, however, is that multiple household distributional concerns are required within each domain to rationalize the observed allocations. This finding motivates a closer examination of household characteristics to better understand the source of variation in distributional concerns.

Table 6: Minimum number of distributional types

Domains	Number of Types	Number of Households	Ratio
Consumption	4	270	67.50
Non-market Time	7	318	45.43
Leisure	6	220	36.67

Explained heterogeneity. The second part of the analysis relates unobserved heterogeneity in household distributional concerns to observable household characteristics, using the methodology described in Section 4.3. We consider as potential predictors: household-level aggregates of spouses' characteristics (age, education, number of children, gender role attitudes, and ethnic diversity views), intra-household differences (measured by age gaps and

education differences), and spouse-specific measures of gender role attitudes and ethnic diversity views. This analysis sheds light on how household distributional concerns are shaped and which factors are most closely associated with them.

For each characteristic, we first compute the 33rd and 67th percentiles of its empirical distribution. The cutoff values for each characteristic are reported in Appendix A.4. We then use these cutoffs to partition households with alignment in a given domain in the low, middle, and upper groups associated with this characteristic. The number of observations in each partition is reported in Table A.11.[19] We finally compute the corresponding κ^k-ratio per decision domain. A lower κ^k indicates that characteristic k better captures heterogeneity in household distributional concerns, whereas a higher κ^k reflects weaker prediction. Table 7 reports the κ^k-ratio for each characteristic across decision domains. In the consumption domain, nearly all characteristics produce the same κ^k-ratio of 2.25, with the age gap yielding a higher value of 2.50. Overall, the characteristics show comparable performance in predicting household distributional concerns for consumption.

We find more diverse patterns in the labor domain. Wives' gender role attitudes and the number of children emerge as the most informative characteristics, both achieving the lowest ratio of 2.00 and outperforming other characteristics such as education. Childbearing interacts with women's career trajectories and generates long-lasting effects on their labor market prospects, widely documented as the child penalty (Kleven et al., 2025, 2019). Wives' gender role attitudes also play an important role, as they are correlated with perceptions of financial autonomy and career expectations. Together, these factors shape women's views on market work and, in turn, influence joint labor supply. By contrast, men typically exhibit a stronger attachment to the labor market and are more likely to assume a breadwinner role anyway. This helps explain why husbands' gender role attitudes and household-level aggregates are somewhat weaker predictors of heterogeneity in market versus non-market time. Views on ethnic diversity, at both the household and individual levels, perform poorly in this domain, as reflected in their higher κ^k-ratios.

In the leisure domain, education and the number of children are the strongest predictors of household distributional concerns, both attaining the lowest ratio of 1.83. Children increase

[19]The three-partition specification yields approximately 80 households per subgroup and serves as our main specification, as it balances conceptual appeal with empirical performance. Conceptually, too few groups may obscure theoretically meaningful distinctions and fail to capture non-monotonic patterns (e.g., with respect to education or the presence of children). Consistent with this concern, robustness tests reported in Appendix A.5.1 show a less differentiated predictive performance of observables under two-group partitions. Conversely, increasing the number of groups further reduces overall predictive performance due to overly fine stratification, a pattern reflected in the higher κ^k-ratios observed for four-group partitions. Nevertheless, both alternative specifications broadly confirm our main findings regarding the most successful predictors of heterogeneity.

childcare responsibilities, while education is strongly associated with preferences over the division of housework. Views on ethnic diversity—at both the household and individual levels—also perform well, likely reflecting their correlation with education and broader social attitudes. Age and the spousal age gap yield similar ratios. Age reflects life-cycle dynamics such as changes in productivity and household composition that shape time use. By contrast, aggregate gender values and wives' own views display weaker explanatory power, whereas husbands' gender attitudes are more strongly associated with housework-leisure allocations. This pattern is consistent with women's continued primary responsibility for housework in Dutch households, regardless of their stated views, while husbands' attitudes more directly influence their contribution to housework (see also Table A.23). Finally, spousal differences in educational attainment are not a key source of heterogeneity in distributional concerns.

Table 7: κ^k-ratios across characteristics

Characteristics	Consumption	Non-market Time	Leisure
Panel A. Household characteristics			
Age	2.25	2.14	2.00
Education	2.25	2.14	1.83
Children	2.25	2.00	1.83
Gender values	2.25	2.14	2.17
Ethnic views	2.25	2.43	2.00
Age gap	2.50	2.14	2.00
Education difference	2.25	2.29	2.33
Panel B. Individual characteristics			
Husband's gender values	2.25	2.14	2.00
Wife's gender values	2.25	2.00	2.17
Husband's ethnic views	2.25	2.29	2.00
Wife's ethnic views	2.25	2.14	2.00

We conduct two sets of robustness checks to verify the validity of our results, in Appendix A.5. First, we vary the number of sub-partitions used to classify households based on their characteristics. Second, we consider alternative values for the scaling factors used to construct domestic productivity. Both exercises yield qualitatively similar results.

Permutations-based sensitivity test. We further assess the robustness of our results using a permutations-based analysis following Cherchye et al. (2024). A low κ^k-ratio can arise by chance even if variation in characteristic k is unrelated to true preference heterogeneity. This raises the concern that partitioning based on variable k may appear informative due to sampling noise rather than systematic differences in household distributional views.

To address this concern, we evaluate the explanatory power of each characteristic by comparing the actual κ^k-ratio to the distribution of $\hat{\kappa}^k$-ratios obtained from repeated random permutations of variable k across households. This procedure allows us to assess statistical power and rule out spurious findings.

Specifically, we generate counterfactual samples by randomly reassigning characteristic k across households, while preserving its overall distribution through sampling without replacement. This procedure ensures that the reshuffled characteristic is independent of the underlying household distributional concern. For each reassignment, we compute a new $\hat{\kappa}^k$-ratio. Repeating this process 200 times yields an empirical distribution of ratios under the null hypothesis of independence. Finally, we compare the actual κ^k-ratio to this empirical distribution. As supporting evidence, Table A.12 reports the average $\hat{\kappa}^k$-ratio across the 200 permuted datasets. For most characteristics, the average simulated ratio exceeds the actual ratio, indicating that the actual characteristics are more informative about common household distributional concerns than a random assignment.

Table 8: Percentage of simulated $\hat{\kappa}^k$ strictly larger than (equal to) the real κ^k

Characteristics	Consumption	Non-market Time	Leisure
Panel A. Household characteristics			
Age	10% (88%)	47% (44%)	37% (52%)
Education	9% (84%)	43% (40%)	83% (17%)
Children	14% (61%)	78% (21%)	80% (20%)
Gender values	9% (88%)	42% (47%)	5% (36%)
Ethnic views	14% (86%)	0% (8%)	43% (49%)
Age gap	0% (12%)	44% (42%)	37% (47%)
Education difference	10% (87%)	7% (31%)	1% (6%)
Panel B. Individual characteristics			
Husband's gender values	10% (84%)	44% (43%)	51% (41%)
Wife's gender values	9% (90%)	93% (7%)	9% (35%)
Husband's ethnic views	13% (81%)	8% (27%)	41% (43%)
Wife's ethnic views	9% (88%)	45% (47%)	46% (42%)

Table 8 presents the percentage of simulated ratios that are *strictly* larger than the actual κ^k-ratio, with the percentage of simulated ratios equal to the actual ratio reported between brackets. A high percentage indicates that the actual κ^k-ratio is unusually low, signaling stronger explanatory power of characteristic k. Distinct patterns emerge when comparing the three decision domains, in line with our previous findings. Regarding consumption decisions, approximately only 10% of the simulated ratios lie strictly above the actual ratio, suggesting limited predictive power of observed characteristics in the consumption domain.

In the labor domain, most characteristics provide meaningful predictive content for household distributional concerns, with the main exceptions being the husband's and household-level ethnic inclusion attitudes as well as differences in spouses' educational attainment. Wives' gender values and the number of children are especially salient, with 93% and 78% of simulated ratios strictly exceeding the observed ratio, respectively. The husband's and household-level gender values also display meaningful explanatory power, with shares around 42–44%. Household age, potentially capturing seniority and associated work patterns, further retains explanatory power.

For the division of time between leisure and household chores, education and the number of children exhibit the strongest explanatory power, with 83% and 80% of permuted samples generating ratios strictly above the empirical benchmark, respectively. Husbands' gender values also show relatively strong predictive content, whereas wives' and household-level gender values fall below 10%. This asymmetry aligns with the pattern that women's housework levels tend to be relatively stable regardless of their stated attitudes, while men with more progressive views take on substantially more housework. Ethnic inclusion views, which correlate with education and broader social progressiveness, also exert meaningful influence on the desired allocations of leisure between the male and female spouse.

6 Conclusion

We examine within-household allocations between spouses in three distinct domains of family decision-making: consumption, labor, and other time uses. Our analysis draws on the Longitudinal Internet studies for the Social Sciences panel, which provides detailed information on demographics, personal values, time use, and consumption for Dutch couples. We find that households' allocation patterns vary substantially across decision domains, with considerable dispersion in distributions across households within each domain.

A distinctive feature of the data is that it includes reported differences of opinion between spouses in each decision domain. Households most frequently report aligned preferences in labor decisions, and least frequently in decisions involving other uses of time. Importantly, about 72.2% of households exhibit within-household preference homogeneity in at least one decision domain. Motivated by these findings, we develop an extension of the collective model that allows distributional preferences and spousal preference alignment to hold 'narrowly' within specific domains.

We then turn to the question of what explains the substantial dispersion in observed allocations, particularly among households that report spousal preference alignment in the domain under consideration. According to our model, the remaining variation in alloca-

tions can be attributed to differences in prices, income, the household welfare index, and distributional concerns across households. Our structural nonparametric analysis recovers the minimum number of distributional types required to rationalize the observed allocation patterns—each type corresponding to a different household-level view on the desired allocation of (domain-specific) resources between the male and female spouse.

To implement this revealed preference analysis, we supplement the LISS data with Consumer Price Indices and industry-specific nominal wage rates from Statistics Netherlands (CBS), which we use to construct personalized consumption prices and measures of domestic productivity. This construction relies on a particular structure of the *lower-level* composition of spouse-specific bundles per domain: each individual sets their personalized quantities (i.e., across product categories or domestic tasks) in constant ratios, independent of quantities chosen by the spouse. However, the household welfare index remains unrestricted, as do the distributional utility functions, which are the central objects of our analysis.

The results indicate that four or more distributional types are required in each domain to rationalize the observed allocations. We then examine a range of observable characteristics to identify the source of heterogeneity in household distributional concerns. Distinct patterns emerge across decision domains. For labor, variation in distributional concerns across households is best explained by wives' gender values and the number of children. For the division between leisure and housework, the household's education and the number of children play the most prominent roles, with husbands' gender values also showing some predictive power. In line with a persistent attachment of each spouse to their traditionally gendered domain, the allocations of labor supply, respectively housework, are less responsive to husbands' (wives') reported gender values.

References

Achard, P., Albrecht, S., Ghidoni, R., Cettolin, E., Suetens, S., 2025. Local exposure to refugees changed attitudes to ethnic minorities in the Netherlands. The Economic Journal 135 (667), 808–837.

Afriat, S. N., 1967. The construction of utility functions from expenditure data. International Economic Review 8 (1), 67–77.

Andreoni, J., Miller, J., 2002. Giving according to GARP: An experimental test of the consistency of preferences for altruism. Econometrica 70 (2), 737–753.

Apps, P. F., Rees, R., 1988. Taxation and the household. Journal of Public Economics 35 (3), 355–369.

Balafoutas, L., Kerschbamer, R., Kocher, M., Sutter, M., 2014. Revealed distributional preferences: Individuals vs. teams. Journal of Economic Behavior & Organization 108, 319–330.

Becker, G. S., 1981. Altruism in the family and selfishness in the market place. Economica 48 (189), 1.

Berry, J., Dizon-Ross, R., Jagnani, M., 2025. Not playing favorites: Parents and the value of equal opportunity. American Economic Journal: Applied Economics 17 (3), 117–160.

Bertrand, M., Kamenica, E., Pan, J., 2015. Gender identity and relative income within households. The Quarterly Journal of Economics 130 (2), 571–614.

Brown, A. L., Meer, J., Williams, J. F., 2019. Why do people volunteer? An experimental analysis of preferences for time donations. Management Science 65 (4), 1455–1468.

Castillo, M., Freer, M., 2018. Revealed differences. Journal of Economic Behavior & Organization 145, 202–217.

Cherchye, L., De Rock, B., Platino, V., 2012a. Private versus public consumption within groups: Testing the nature of goods from aggregate data. Economic Theory 54 (3), 485–500.

Cherchye, L., De Rock, B., Surana, K., 2026. Labor market participation, marriage and individual welfare. Journal of Political Economy Microeconomics.

Cherchye, L., De Rock, B., Vermeulen, F., 2011. The revealed preference approach to collective consumption behaviour: Testing and sharing rule recovery. The Review of Economic Studies 78 (1), 176–198.

Cherchye, L., Rock, B. D., Vermeulen, F., 2012b. Married with children: A collective labor supply model with detailed time use and intrahousehold expenditure information. American Economic Review 102 (7), 3377–3405.

Cherchye, L., Saelens, D., Tuncer, R., 2024. From unobserved to observed preference heterogeneity: A revealed preference methodology. Economica 91, 996–1022.

Chiappori, P.-A., 1992. Collective labor supply and welfare. Journal of Political Economy 100 (3), 437–467.

Cosaert, S., 2017. What types are there? Computational Economics 53 (2), 533–554.

Cosaert, S., Theloudis, A., Verheyden, B., 2023. Togetherness in the household. American Economic Journal: Microeconomics 15 (1), 529–579.

Crawford, I., Pendakur, K., 2013. How many types are there? The Economic Journal 123 (567), 77–95.

Deb, R., Gazzale, R. S., Kotchen, M. J., 2014. Testing motives for charitable giving: A revealed-preference methodology with experimental evidence. Journal of Public Economics 120, 181–192.

Diewert, W. E., 1973. Afriat and revealed preference theory. The Review of Economic Studies 40 (3), 419–425.

Ellingsen, T., Johannesson, M., 2009. Time is not money. Journal of Economic Behavior & Organization 72 (1), 96–102.

Ellis, A., Freeman, D. J., 2024. Revealing choice bracketing. American Economic Review 114 (9), 2668–2700.

Engelmann, D., Strobel, M., 2004. Inequality aversion, efficiency, and maximin preferences in simple distribution experiments. American Economic Review 94 (4), 857–869.

Exley, C. L., Kessler, J. B., 2024. Equity concerns are narrowly framed. American Economic Journal: Microeconomics 16 (2), 147–179.

Fallucchi, F., Kaufmann, M., 2021. Narrow bracketing in work choices. Working paper.

Fehr, E., Charness, G., 2023. Social preferences: Fundamental characteristics and economic consequences. SSRN Electronic Journal.

Fehr, E., Epper, T., Senn, J., 2023. The fundamental properties, stability and predictive power of distributional preferences. SSRN Electronic Journal.

Fisman, R., Jakiela, P., Kariv, S., 2015. How did distributional preferences change during the great recession? Journal of Public Economics 128, 84–95.

Fisman, R., Kariv, S., Markovits, D., 2007. Individual preferences for giving. American Economic Review 97 (5), 1858–1876.

Fostel, A., Scarf, H. E., Todd, M. J., 2004. Two new proofs of afriat's theorem. Economic Theory 24 (1), 211–219.

Gronau, R., 1977. Leisure, home production, and work–the theory of the allocation of time revisited. Journal of Political Economy 85 (6), 1099–1123.

Hoderlein, S., Mihaleva, S., 2008. Increasing the price variation in a repeated cross section. Journal of Econometrics 147 (2), 316–325.

Jokubauskaitė, S., Schneebaum, A., 2022. Assessing the value of household work based on wages demanded on online platforms for substitutes. Review of Economics of the Household 20 (1), 153–160.

Kerschbamer, R., 2015. The geometry of distributional preferences and a non-parametric identification approach: The equality equivalence test. European Economic Review 76, 85–103.

Kleven, H., Landais, C., Leite-Mariante, G., 2025. The child penalty atlas. Review of Economic Studies 92 (5), 3174–3207.

Kleven, H., Landais, C., Søgaard, J. E., 2019. Children and gender inequality: Evidence from Denmark. American Economic Journal: Applied Economics 11 (4), 181–209.

Landefeld, J. S., Fraumeni, B. M., Vojtech, C. M., 2009. Accounting for household production: A prototype satellite account using the american time use survey. Review of Income and Wealth 55 (2), 205–225.

Lewbel, A., 1989. Identification and estimation of equivalence scales under weak separability. The Review of Economic Studies 56 (2), 311.

Lilley, A., Slonim, R., 2014. The price of warm glow. Journal of Public Economics 114, 58–74.

Lundberg, S., Pollak, R. A., 1993. Separate spheres bargaining and the marriage market. Journal of Political Economy 101 (6), 988–1010.

Read, D., Loewenstein, G., Rabin, M., 2006. Choice Bracketing. Cambridge University Press, pp. 372–396.

Samuelson, P. A., 1948. Consumption theory in terms of revealed preference. Economica 15 (60), 243–253.

Varian, H. R., 1982. The nonparametric approach to demand analysis. Econometrica 50 (4), 945–973.

A Appendix

A.1 Sample selection and variable construction

A.1.1 Sample selection

We focus on households in which both spouses are aged between 25 and 60. After excluding observations with missing values for key demographic variables and reported views on gender roles and ethnic diversity, the sample contains 1,193 households. We retain households with and without children. For households with children, we further restrict the sample to those whose youngest child is aged 18 or younger. This restriction removes 31.2 percent of households and leaves 821 households.

To ensure the availability of wage information for both spouses, which is required for the GARP test, we drop households in which the hourly wage rate of either spouse is missing. This includes households in which at least one spouse does not participate in the labor market or does not report income. We further exclude households in which either spouse reports an hourly wage above 300 euros, due to concerns about measurement error. Households with missing wage information account for 7.7% among wives and 2.9% among husbands, while households in which neither spouse works represent only a very small share of the sample, at around 0.7%. After applying these restrictions, the sample consists of 740 households.

In a final step, we exclude households with inconsistent reports of consumption and time use. This includes households reporting full-time market work with no remaining time allocation, as well as households reporting zero or missing values for all time use categories. We apply the same criteria to consumption data and drop households with zero or missing values for all consumption bundles. After these final exclusions, the sample consists of 661 households, which form the basis of our empirical analysis.

A.1.2 Gender values

From LISS Module "Politics and Values"

cv09b109 A working mother's relationship with her children can be just as close and warm as that of a non-working mother

cv09b110 A child that is not yet attending school is likely to suffer the consequences if his or her mother has a job

cv09b111 Overall, family life suffers the consequences if the mother has a full-time job

cv09b112 Both father and mother should contribute to the family income

cv09b113 The father should earn money, while the mother takes care of the household and the family

cv09b114 Fathers ought to do more in terms of household work than they do at present

cv09b115 Fathers ought to do more in terms of childcare than they do at present

The responses are recorded on a five-point scale: 1. Fully disagree; 2. Disagree; 3. Neither agree nor disagree; 4. Agree; 5. Fully agree. We reverse the coding of the variables "cv12e109" "cv12e112" "cv12e114" and "cv12e115" so that all recoded responses measure the respondent's level of conservativeness, with higher values indicating more traditional views.

A.1.3 Ethnic diversity views

From LISS Module "Politics and Values"

cv09b116 It is good if society consists of people from different cultures

cv09b118 It should be made easier to obtain asylum in the Netherlands

cv09b119 Legally residing foreigners should be entitled to the same social security as Dutch citizens

cv09b120 There are too many people of foreign origin or descent in the Netherlands

cv09b122 Some sectors of the economy can only continue to function because people of foreign origin or descent work there

cv09b123 It does not help a neighborhood if many people of foreign origin or descent move in

The responses are recorded on a five-point scale: 1. Fully disagree; 2. Disagree; 3. Neither agree nor disagree; 4. Agree; 5. Fully agree. We reverse the coding of the variables "cv09b120" and "cv09b123" so that all recoded responses measure the respondent's degree of cultural inclusion. A similar use of these variables can be found in Achard et al. (2025).

A.1.4 Consumption, time use and wages

Spouses' private consumption is derived from the total monthly expenditure on items intended for personal use, based on information from the LISS Core Study 34–Time Use and Consumption. Taking the 2009 wave as an example, the relevant categories include eating at home by themselves ($bf09a091$), food and drinks outside the house ($bf09a095$), cigarettes and other tobacco products ($bf09a096$), clothing ($bf09a097$), personal care products and services ($bf09a098$), medical care and health costs not covered by insurance ($bf09a099$), leisure time expenditure ($bf09a100$), schooling ($bf09a101$), donations and gifts ($bf09a102$),

and others ($bf09a103$).

Time allocation is based on information from the LISS Core Study 6–Work and Schooling and Core Study 34–Time Use and Consumption. The former provides data on working and schooling activities, while the latter includes a detailed set of questions covering various aspects of daily time use, such as household chores, personal care, commuting, and other activities.

Market work is constructed as the time each spouse spends on paid work and commuting. Paid work includes hours devoted to the primary job ($cw09b127$) and, if applicable, a secondary job ($cw09b144$), both drawn from Core Study 6. When information on paid work hours is missing in this source, alternative measures are taken from Core Study 34 ($bf09a005$, $bf09a006$). Commuting time is obtained from the same study ($bf09a007$, $bf09a008$). To avoid implausibly high values, total market hours are capped at 84 per week, corresponding to 12 hours per day over seven days.

Domestic work measures the time each spouse devotes to household production activities, including household chores ($bf09a009$–$bf09a010$), administrative tasks and financial management ($bf09a025$–$bf09a026$), and childcare if child(ren) are present ($bf09a013$–$bf09a014$). Leisure captures each spouse's private leisure time, constructed from the individual leisure measures ($bf09a021$, $bf09a022$) net of joint leisure time spent with the spouse ($bf09a064$, $bf09a065$).

Market wages w_M^i are computed as monthly earnings divided by monthly hours of paid work, excluding commuting time. Monthly earnings correspond to personal income after taxes (variable $nettoinkf$), which, when not directly reported by the respondent, are imputed by the LISS based on reported gross income.

A.2 Revealed preference characterizations

- We first derive the nonparametric testable conditions from the collective model with narrow distributional preferences described in Definition 1. The first-order optimality conditions are,

$$\begin{aligned}
\delta^1_{C,v}\frac{\partial U^1_C}{\partial c^i_v} + \delta^2_{C,v}\frac{\partial U^2_C}{\partial c^i_v} - \lambda_v p^i_{C,v} &= 0 \\
\delta^1_{T,v}\frac{\partial U^1_T}{\partial t^i_v} + \delta^2_{T,v}\frac{\partial U^2_T}{\partial t^i_v} - \lambda_v p^i_{T,v} &= 0 \\
\delta^1_{L,v}\frac{\partial U^1_L}{\partial l^i_v} + \delta^2_{L,v}\frac{\partial U^2_L}{\partial l^i_v} - \lambda_v p^i_{L,v} &= 0
\end{aligned}$$

Concavity of the distributional utility functions further requires,

$$\begin{aligned} u^i_{C,s} - u^i_{C,v} &\leq \left(\frac{\partial U^i_C}{\partial c^1_v}(c^1_s - c^1_v) + \frac{\partial U^i_C}{\partial c^2_v}(c^2_s - c^2_v) \right) \\ u^i_{T,s} - u^i_{T,v} &\leq \left(\frac{\partial U^i_T}{\partial t^1_v}(t^1_s - t^1_v) + \frac{\partial U^i_T}{\partial t^2_v}(t^2_s - t^2_v) \right) \\ u^i_{L,s} - u^i_{L,v} &\leq \left(\frac{\partial U^i_L}{\partial l^1_v}(l^1_s - l^1_v) + \frac{\partial U^i_L}{\partial l^2_v}(l^2_s - l^2_v) \right) \end{aligned}$$

Next, we define shadow prices $\pi^{ij}_{C,v} = \frac{\delta^i_{C,v}}{\lambda_v}\frac{\partial U^i_C}{\partial c^j_v}$, $\pi^{ij}_{T,v} = \frac{\delta^i_{T,v}}{\lambda_v}\frac{\partial U^i_T}{\partial t^j_v}$ and $\pi^{ij}_{L,v} = \frac{\delta^i_{L,v}}{\lambda_v}\frac{\partial U^i_L}{\partial l^j_v}$ based on the marginal utilities and decision weights. This gives:

$$\begin{aligned} u^i_{C,s} - u^i_{C,v} &\leq \frac{\lambda_v}{\delta^i_{C,v}} \left(\pi^{i1}_{C,v}(c^1_s - c^1_v) + \pi^{i2}_{C,v}(c^2_s - c^2_v) \right) \\ &\leq \lambda^i_{C,v} \left(\pi^{i1}_{C,v}(c^1_s - c^1_v) + \pi^{i2}_{C,v}(c^2_s - c^2_v) \right) \\ u^i_{T,s} - u^i_{T,v} &\leq \frac{\lambda_v}{\delta^i_{T,v}} \left(\pi^{i1}_{T,v}(t^1_s - t^1_v) + \pi^{i2}_{T,v}(t^2_s - t^2_v) \right) \\ &\leq \lambda^i_{T,v} \left(\pi^{i1}_{T,v}(t^1_s - t^1_v) + \pi^{i2}_{T,v}(t^2_s - t^2_v) \right) \\ u^i_{L,s} - u^i_{L,v} &\leq \frac{\lambda_v}{\delta^i_{L,v}} \left(\pi^{i1}_{L,v}(l^1_s - l^1_v) + \pi^{i2}_{L,v}(l^2_s - l^2_v) \right) \\ &\leq \lambda^i_{L,v} \left(\pi^{i1}_{L,v}(l^1_s - l^1_v) + \pi^{i2}_{L,v}(l^2_s - l^2_v) \right) \end{aligned}$$

with $\pi^{1j}_{C,s} + \pi^{2j}_{C,s} = p^j_{C,s}$, $\pi^{1j}_{T,s} + \pi^{2j}_{T,s} = p^j_{T,s}$, and $\pi^{1j}_{L,s} + \pi^{2j}_{L,s} = p^j_{L,s}$.

Finally applying the insights of Fostel et al. (2004), an equivalent expression of the above inequalities is:

$$\begin{aligned} &\{\pi^{11}_{C,s}, \pi^{12}_{C,s}, c^1_s, c^2_s\}_{s\leq S} \text{ satisfies GARP} \\ &\{\pi^{21}_{C,s}, \pi^{22}_{C,s}, c^1_s, c^2_s\}_{s\leq S} \text{ satisfies GARP} \end{aligned} \quad \text{with } \pi^{1j}_{C,s} + \pi^{2j}_{C,s} = p^j_{C,s}$$

$$\begin{aligned} &\{\pi^{11}_{T,s}, \pi^{12}_{T,s}, t^1_s, t^2_s\}_{s\leq S} \text{ satisfies GARP} \\ &\{\pi^{21}_{T,s}, \pi^{22}_{T,s}, t^1_s, t^2_s\}_{s\leq S} \text{ satisfies GARP} \end{aligned} \quad \text{with } \pi^{1j}_{T,s} + \pi^{2j}_{T,s} = p^j_{T,s}$$

$$\begin{aligned} &\{\pi^{11}_{L,s}, \pi^{12}_{L,s}, l^1_s, l^2_s\}_{s\leq S} \text{ satisfies GARP} \\ &\{\pi^{21}_{L,s}, \pi^{22}_{L,s}, l^1_s, l^2_s\}_{s\leq S} \text{ satisfies GARP} \end{aligned} \quad \text{with } \pi^{1j}_{L,s} + \pi^{2j}_{L,s} = p^j_{L,s}$$

This is the revealed preference characterization of the baseline model with narrow distributional preferences. It is equivalent to nonparametric tests of a collective model with two purely public goods, applied separately to each domain.

- We then note that this revealed preference test of the model with narrow distributional preferences is in fact trivial. Regardless of the number of households in D, any data set will pass these conditions. One possible rationalization is easily obtained by setting all $\pi^{12}_{C,s}, \pi^{21}_{C,s}, \pi^{12}_{T,s}, \pi^{21}_{T,s}, \pi^{12}_{L,s}, \pi^{21}_{L,s} = 0$.

  This conclusion is a specific application of the result in Cherchye et al. (2012a) (more specifically, see their Footnote 9), namely that the collective model with purely public goods cannot be rejected when the number of goods in the utility function is less than three.

- In this paper, empirical tractability is restored by the additional requirement of spousal preference alignment presented in Definition 2. We now show how we integrate this requirement within the collective model with narrow distributional preferences. We focus on spousal preference alignment in the consumption domain. The reasoning for labor and leisure is analogous so we will not repeat it here.

  First, alignment in the consumption domain imposes intra-household preference homogeneity with respect to the distribution of consumtion. In other words, there exist numbers $u_{C,s}$ and $\partial U_C/\partial c^j_v$ such that we can replace $u^1_{C,s}$ and $u^2_{C,s}$ with $u_{C,s}$, and $\frac{\partial U^1_C}{\partial c^j_v}, \frac{\partial U^2_C}{\partial c^j_v}$ with $\frac{\partial U_C}{\partial c^j_v}$ (for all s, v).

  Filling in the final equalities in $\pi^{1j}_{C,v} + \pi^{2j}_{C,v} = p^j_{C,v}$, the latter becomes $\frac{\delta^1_{C,v}+\delta^2_{C,v}}{\lambda_v}\frac{\partial U_C}{\partial c^j_v} = p^j_{C,v}$. The Afriat inequalities of both spouses thus coincide and reduce to,

$$
\begin{aligned}
u_{C,s} - u_{C,v} &\leq \left(\frac{\partial U_C}{\partial c^1_v}(c^1_s - c^1_v) + \frac{\partial U_C}{\partial c^2_v}(c^2_s - c^2_v)\right) \\
\Leftrightarrow u_{C,s} - u_{C,v} &\leq \frac{\lambda_v}{\delta^1_{C,v} + \delta^2_{C,v}}\left(p^1_{C,v}(c^1_s - c^1_v) + p^2_{C,v}(c^2_s - c^2_v)\right) \\
\Leftrightarrow u_{C,s} - u_{C,v} &\leq \lambda_{v,C}\left(p^1_{C,v}(c^1_s - c^1_v) + p^2_{C,v}(c^2_s - c^2_v)\right)
\end{aligned}
$$

  Again applying Fostel et al. (2004), the final set of inequalities is equivalent to the condition that $D_C = \{p^1_{C,s}, p^2_{C,s}, c^1_s, c^2_s\}_{s \leq S}$ satisfies GARP.

  So far, we have focused on deriving necessary conditions for consistency with our model. Sufficiency of the conditions may also be established. Applied to our set-up, one can

construct the following utility functions:

$$\begin{aligned} U_C(c^1, c^2) &= \min_v \left\{ u_{C,v} + \lambda_{C,v} \left(p^1_{C,v}(c^1 - c^1_v) + p^2_{C,v}(c^2 - c^2_v) \right) \right\} \\ U^i_T(t^1, t^2) &= \min_v \left\{ u^i_{T,v} + \lambda^i_{T,v} \left(\pi^{i1}_{T,v}(t^1 - t^1_v) + \pi^{i2}_{T,v}(t^2 - t^2_v) \right) \right\} \\ U^i_L(l^1, l^2) &= \min_v \left\{ u^i_{L,v} + \lambda^i_{L,v} \left(\pi^{i1}_{L,v}(l^1 - l^1_v) + \pi^{i2}_{L,v}(l^2 - l^2_v) \right) \right\} \end{aligned}$$

and decision weights $\delta^1_{C,v} = \delta^2_{C,v} = 0.5/\lambda_{C,v}$, $\delta^i_{T,v} = 1/\lambda^i_{T,v}$, and $\delta^i_{L,v} = 1/\lambda^i_{L,v}$. It can then be shown along the lines of Cherchye et al. (2011) that the constructed utility functions provide a rationalization (with narrow distributional concerns) of D, with spousal preference alignment in consumption.

A.3 CPI, nominal wages, and corresponding bundles

We construct each spouse's personal consumption price and domestic productivity following the methodology in Section 4.1. The personalized price for private consumption is constructed from a bundle of 10 commodities, with item-level prices taken from the CPI published by Statistics Netherlands (CBS). The personalized level of domestic productivity is constructed from a bundle of three household activities, where the value of each activity is (a scaled measure of) the nominal wage rate in the corresponding sector, using data from CBS.

The bundle of commodities is presented in Table A.9. The first column lists the items in the private consumption bundle,[20] and the second column gives the corresponding variable names from LISS, using the 2009 wave as an example. The third column reports the matching CBS expenditure categories, and the remaining columns report the Consumer Price Index for the sample years 2009, 2010, and 2012. The CPI uses 2015 as the base year with an index of 100. For the calculation of personal prices, we rescale each series by dividing by 100 so that 2015 equals 1. This rescaling preserves relative price variation and does not affect the results.

The household activities are listed in Table A.10. The first column lists the domestic work activities included, and the second column provides the corresponding LISS variable names. For each activity, the first variable records hours spent and the second variable records additional minutes spent. The third column reports the matching CBS sectors, and the remaining columns report wages per hour worked for 2009, 2010, and 2012.

[20] The "medical care products" refer specifically to items and services not covered by insurance.

Table A.9: Bundle of private commodity and CPI

Commodities	Variable (LISS)	Expenditure category (CBS)	Y2009	Y2010	Y2012
Eating at home	bf09a091	010000 Food and non-alco drinks	93.17	93.09	97.07
Food outside the house	bf09a095	111110 Restaurants and cafe	88.23	90.25	94.61
Tobacco products	bf09a096	022000 Tobacco	72.55	75.78	84.15
Clothing	bf09a097	030000 Clothing and footwear	101.59	101.07	102.62
Personal care products	bf09a098	121000 Personal care	93.69	94.35	96.84
Medical care products	bf09a099	061000 Medical products	99.48	100.06	100.35
Leisure expenditure	bf09a100	090000 Recreation and culture	94.74	94.98	96.96
Schooling	bf09a101	100000 Education	86.20	87.96	92.63
Donations and gifts	bf09a102	093120 Toys and celebration articles	100.43	97.55	97.87
Others	bf09a103	000000 All items	90.44	91.59	96.04

Table A.10: Bundle of domestic activity and wages per hour worked

Activities	Variable (LISS)	Sector (CBS)	Y2009	Y2010	Y2012
Household chores	bf09a009, bf09a010	I. Accommodation and food serving	14.4	14.4	14.6
Administrative chores and family finances	bf09a025, bf09a026	N. Renting and other business support	16.7	16.5	17.2
Childcare	bf09a013, bf09a014	Q. Health and social work activities	23.4	23.9	24.8

A.4 Supplementary materials for heterogeneity analyses

We sum the ages of both spouses and group households into three age categories: 51–77 years (younger), 78–91 years (middle-aged), and 92–118 years (older). Educational attainment is coded on a six-point scale, where (1) denotes primary or below, (2) intermediate secondary, (3) higher secondary, (4) vocational education, (5) university, and (6) postgraduate. We then sum spouses' education scores and classify households as low educated for a total score of 2–7, intermediately educated for a score of 8–9, and highly educated for a score of 10–12. To account for the presence of children, we group households by the number of children currently living at home: households without children, households with 1–2 children, and households with three or more children.

We also include two sets of variables that capture spouses' personal values. The first set concerns gender role attitudes within households. Based on respondents' answers to survey questions about the division of market work and domestic work at home (see Appendix A.1.2), we construct for each spouse a composite index of gender role attitudes, where a higher score reflects a more conservative view. We then sum the spouses' indices to obtain a household-level measure of gender norms. Based on this measure, we classify households as liberal when the total score is between 14 and 31, moderate when it is between 32 and 37, and conservative when it is between 38 and 59. The second set of variables captures views on ethnic diversity (see Appendix A.1.3). We compute the mean of the spouses' responses, where a higher value reflects a more inclusive attitude toward immigrants and ethnic minorities. Using this measure, we classify households as restrictive for scores from 1.17 up to and including 2.67, ambivalent for scores above 2.67 and below 3.08, and inclusive for scores from 3.08 to 4.5.

Further characteristics include the spouses' age gap, differences in educational attainment, each spouse's gender value, and their views on ethnic diversity. We measure the age gap as the husband's age minus the wife's age and group households into three ranges: -11 to 1 years, 2 to 3 years, and 4 to 19 years. The education difference is defined as the husband's education score minus the wife's score. Households are classified as having a negative gap from -4 to -1, no gap at 0, and a positive gap from 1 to 5.

For gender values, households are classified using each spouse's index. Men's scores of 7–16 are labeled liberal, 17–19 moderate, and 20–34 conservative; women's scores of 7–14 are labeled liberal, 15–18 moderate, and 19–29 conservative. For ethnic diversity views, each spouse reports an average on a 0–5 scale. Using the husband's score, we classify households as restrictive for values from 1 to 2.67, ambivalent for values above 2.67 and up to 3, and inclusive for values above 3 and up to 5. Using the wife's score, we classify households as restrictive for values from 1.33 to 2.67, ambivalent for values above 2.67 and below 3.17, and

inclusive for values above 3.17 and up to 4.67.

These classification rules are designed to ensure comparable sample sizes across sub-partitions, and the exact number of observations in each partition is reported in Table A.11. Although we aim to divide household samples as evenly as possible across partitions, the division based on the number of children is less balanced. This reflects the highly concentrated distribution of this variable, with households without children and those with two children accounting for about 25% and 42% of the sample, respectively, which makes it difficult to create three evenly sized groups. As a robustness check, we also collapse these characteristics into two and four sub-partitions instead of three, and the results are broadly consistent with the three-group specification. This reduces concerns that our conclusions are driven by the number of sub-partitions or by sample-size imbalances in this variable.

Table A.11: Number of observations across demographics (three partitions)

	Consumption			Non-market Time			Leisure		
	Low	Mid	Upper	Low	Mid	Upper	Low	Mid	Upper
Panel A. Household characteristics									
Age	83	78	109	115	103	100	56	85	79
Education	94	111	65	122	132	64	84	84	52
Children	76	157	37	97	174	47	52	136	32
Gender values	106	78	86	125	97	96	80	65	75
Ethnic views	94	90	86	117	100	101	65	87	68
Age gap	130	74	66	154	76	88	106	53	61
Education difference	80	101	89	93	126	99	65	81	74
Panel B. Individual characteristics									
Husband's gender values	104	80	86	125	84	109	83	66	71
Wife's gender values	92	89	89	120	102	96	68	69	83
Husband's ethnic views	124	56	90	147	72	99	88	60	72
Wife's ethnic views	97	93	80	109	127	82	68	94	58

Notes: Each column reports the number of households in the low, middle, and upper partitions for the corresponding characteristic and decision domain.

Table A.12: Power test: simulated $\hat{\kappa}^k$

Characteristics	Consumption	Non-market Time	Leisure
Panel A. Household characteristics			
Age	2.27	2.21	2.05
Education	2.25	2.19	2.06
Children	2.22	2.15	2.03
Gender values	2.26	2.19	2.06
Ethnic views	2.28	2.21	2.07
Age gap	2.25	2.19	2.04
Education difference	2.27	2.19	2.06
Panel B. Individual characteristics			
Husband's gender values	2.26	2.20	2.08
Wife's gender values	2.27	2.20	2.07
Husband's ethnic views	2.26	2.17	2.05
Wife's ethnic views	2.26	2.21	2.08

A.5 Robustness checks

A.5.1 Different sub-partition number

To assess the robustness of our results, we perform a sensitivity analysis by varying the number of household sub-partitions. Specifically, we recompute the κ^k-ratios after dividing households into two and four sub-partitions based on the same characteristics. In the two-partition case, the sample is split at the median (50th percentile) of each characteristic to obtain two groups of comparable size. In the four-partition case, we use the 25th, 50th, and 75th percentiles as cutoffs to form four evenly sized partitions. The exact numbers of observations and the cutoff rules are reported in Tables A.13 and A.15, respectively.

We recompute the κ^k-ratios based on these alternative partitions. Table A.14 reports the κ^k-ratios for the two-partition specification, and Table A.16 presents those for the four-partition specification. While the coarser splits under the two-partition specification seem to dilute detectable heterogeneity (i.e., with more uniform κ^k-ratios across characteristics), the overall patterns remain consistent. In the consumption domain, all characteristics under the two- and four-partition specifications yield highly stable ratios, in line with the main analysis. This confirms that all characteristics show comparable performance in predicting household concerns for consumption distributions. In labor, we observe a pattern similar to the baseline results: gender values and the presence of children remain among the strongest predictors of heterogeneity in distributional concerns. Ethnic views perform less well. In the leisure domain, education and the presence of children continue to perform strongly

in the robustness checks, although the age gap becomes the most relevant factor in the four-partition specification.

Table A.13: Number of observations across characteristics (two partitions)

	Consumption		Non-market Time		Leisure	
	Low	Upper	Low	Upper	Low	Upper
Panel A. Household characteristics						
Age	122	148	174	144	98	122
Education	148	122	191	127	125	95
Children	76	194	97	221	52	168
Gender values	142	128	171	147	114	106
Ethnic views	140	130	168	150	109	111
Age gap	161	109	185	133	134	86
Education difference	181	89	219	99	146	74
Panel B. Individual characteristics						
Husband's gender values	160	110	177	141	130	90
Wife's gender values	143	127	173	145	110	110
Husband's ethnic views	151	119	176	142	111	109
Wife's ethnic views	158	112	191	127	136	84

Notes: Each column reports the number of households in the lower and upper partitions for the corresponding characteristic and decision domain. Aggregate age is grouped into 51–84 years and 85 years and above. Spouses' aggregate educational attainment is classified into total scores of 2–8 and 9 and above. The number of children divides households into those without children and those with children. Spouses' aggregate gender values are split into 14–34 and 35–59. Household ethnic diversity views are split into 1.167–2.917 and above 2.917–4.5. The age gap, defined as the husband's age minus the wife's age, is coded as -11 to 2 for the lower partition and 3 to 19 for the upper partition. The education difference divides households into those in which the husband's education is less than or equal to the wife's and those in which the husband's education exceeds the wife's. Husbands' gender values split households into progressive (scores 7–18) and conservative (scores 19–34), while wives' gender values are split into less conservative (scores 7–16) and more conservative (scores 17–29). Regarding ethnic diversity views, men's value indices are divided into 1–2.833 and above 2.833–5, while women's indices are divided into 1.333–3 and above 3–4.667.

Table A.14: κ^k-ratios across characteristics (two partitions)

Characteristics	Consumption	Non-market Time	Leisure
Panel A. Household characteristics			
Age	1.75	1.57	1.83
Education	1.50	1.57	1.50
Children	1.50	1.57	1.50
Gender values	1.50	1.57	1.50
Ethnic views	1.50	1.71	1.50
Age gap	1.75	1.71	1.67
Education difference	1.75	1.57	1.83
Panel B. Individual characteristics			
Husband's gender values	1.50	1.43	1.50
Wife's gender values	1.50	1.43	1.67
Husband's ethnic views	1.50	1.57	1.50
Wife's ethnic views	1.50	1.57	1.50

Table A.15: Number of observations across characteristics (four partitions)

	Consumption				Non-market Time				Leisure			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Panel A. Household characteristics												
Age	64	58	66	82	92	82	71	73	41	57	66	56
Education	66	82	57	65	95	96	63	64	67	58	43	52
Children	76	50	107	37	97	39	135	47	52	32	104	32
Gender values	72	70	65	63	93	78	71	76	58	56	50	56
Ethnic views	76	64	68	62	92	76	82	68	49	60	65	46
Age gap	99	62	61	48	111	74	76	57	80	54	47	39
Education difference	80	101	40	49	93	126	46	53	65	81	35	39
Panel B. Individual characteristics												
Husband's gender values	76	84	48	62	88	89	58	83	63	67	35	55
Wife's gender values	78	65	63	64	98	75	66	79	57	53	43	67
Husband's ethnic views	91	60	68	51	112	64	83	59	61	50	71	38
Wife's ethnic views	71	87	51	61	81	110	66	61	48	88	39	45

Notes: Each column reports the number of households in four partitions Q1–Q4 for the corresponding characteristic and decision domain. Aggregate age is grouped into 51-73 years, 74-84 years, 85-96 years, and 97-118 years. Spouses' aggregate educational attainment is classified into total scores of 2-6, 7-8, 9, and 10-12. The number of children divides households into those without children, households with one child, households with two children, and households with three or more children. Spouses' aggregate gender value index is split into 14-29, 30-34, 35-39, and 40-59. Household ethnic diversity views are split into 1.167-2.5, above 2.5-2.92, above 2.92-3.25, and above 3.25-4.5. The age gap is coded as -11 to 0, 1-2 years, 3-4 years, and 5-19 years. The education difference divides households into four groups: -4 to -1, equal educational attainment (0), husbands with one score higher than their wives, and husbands with scores 2-5 higher. Husbands' gender values split households into scores 7-15, 16-18, 19-20, and 21-34, while wives' gender values are split into scores 7-13, 14-16, 17-19, and 20-29. Regarding ethnic diversity views, men's value indices are divided into 1-2.5, above 2.5-2.83, above 2.83-3.33, and above 3.33-5, while women's value indices are divided into 1.33-2.5, above 2.5-3, above 3-3.33, and above 3.33-4.67.

Table A.16: κ^k-ratios across characteristics (four partitions)

Characteristics	Consumption	Non-market Time	Leisure
Panel A. Household characteristics			
Age	2.50	2.71	2.50
Education	3.00	2.57	2.50
Children	2.75	2.57	2.50
Gender values	3.00	2.57	2.67
Ethnic views	3.00	2.86	2.83
Age gap	3.00	3.00	2.33
Education difference	3.00	2.71	2.83
Panel B. Individual characteristics			
Husband's gender values	3.00	3.00	2.83
Wife's gender values	3.00	2.43	2.67
Husband's ethnic views	3.00	2.71	2.67
Wife's ethnic views	3.00	2.86	2.50

A.5.2 Different productivity scaling factor

As a further robustness check, we recompute personalized domestic productivity using alternative scaling factors. Specifically, we use a lower scaling factor of $\alpha = 0.4$ and a higher scaling factor of $\alpha = 0.6$. The constructed prices and wages for both spouses are reported in Tables A.17 and A.18. The minimum number of types identified among households with aligned spousal preferences in each decision domain are reported in Tables A.19 and A.20. The two scaling values yield highly similar results. The only deviation occurs for non-market time when $\alpha = 0.4$, which yields six types rather than the seven types found in the main analysis.

The results of the heterogeneity analyses are reported in Tables A.21 and A.22. Since changes in the productivity scaling do not affect personalized consumption prices, the κ^k-ratios for the consumption domain remain unchanged. The discussion therefore focuses on labor and leisure. In the labor domain, wives' gender values continue to be among the strongest predictors of household distributional concerns, whereas views on ethnic diversity remain among the least informative factors. The findings for leisure are also consistent with the main analysis, with the number of children providing the strongest explanation for differences in household distributional concerns over leisure (and housework).

Table A.17: Constructed prices and wages ($\alpha = 0.4$)

	Mean	SD	Min	Max
p_C^1	0.95	0.032	0.76	1.03
p_C^2	0.96	0.024	0.84	1.01
w_M^1	11.97	4.947	0.72	51.28
w_M^2	11.22	7.213	1.80	161.54
w_H^1	6.95	1.207	0.65	9.92
w_H^2	6.68	1.293	1.62	9.88
p_T^1	5.02	4.743	0.07	42.93
p_T^2	4.54	7.057	0.18	155.43

Table A.18: Constructed prices and wages ($\alpha = 0.6$)

	Mean	SD	Min	Max
p_C^1	0.95	0.032	0.76	1.03
p_C^2	0.96	0.024	0.84	1.01
w_M^1	11.97	4.947	0.72	51.28
w_M^2	11.22	7.213	1.80	161.54
w_H^1	9.23	1.951	0.65	14.43
w_H^2	8.78	2.057	1.62	13.62
p_T^1	2.73	4.019	0.07	38.75
p_T^2	2.44	6.652	0.18	152.38

Table A.19: Minimum number of distributional types when $\alpha = 0.4$

Domains	Number of Types	Number of households	Ratio
Consumption	4	270	67.50
Non-market Time	6	318	53.00
Leisure	6	220	36.67

Table A.20: Minimum number of distributional types when $\alpha = 0.6$

Domains	Number of Types	Number of households	Ratio
Consumption	4	270	67.50
Non-market Time	7	318	45.43
Leisure	6	220	36.67

Table A.21: κ^k-ratios across characteristics ($\alpha = 0.4$)

Characteristics	Consumption	Non-market Time	Leisure
Panel A. Household characteristics			
Age	2.25	2.33	2.00
Education	2.25	2.50	1.83
Children	2.25	2.33	1.83
Gender values	2.25	2.50	2.00
Ethnic views	2.25	2.50	2.00
Age gap	2.50	2.33	1.83
Education difference	2.25	2.50	2.17
Panel B. Individual characteristics			
Husband's gender values	2.25	2.50	2.00
Wife's gender values	2.25	2.33	1.83
Husband's ethnic views	2.25	2.33	1.83
Wife's ethnic views	2.25	2.50	2.00

Table A.22: κ^k-ratios across characteristics ($\alpha = 0.6$)

Characteristics	Consumption	Non-market Time	Leisure
Panel A. Household characteristics			
Age	2.25	2.14	2.00
Education	2.25	2.00	2.17
Children	2.25	2.14	1.83
Gender values	2.25	2.00	2.33
Ethnic views	2.25	2.29	2.00
Age gap	2.50	2.14	2.17
Education difference	2.25	2.14	2.17
Panel B. Individual characteristics			
Husband's gender values	2.25	2.14	2.17
Wife's gender values	2.25	2.00	2.33
Husband's ethnic views	2.25	2.14	2.17
Wife's ethnic views	2.25	2.14	2.17

A.6 Correlation of housework and gender norms

Table A.23: Housework hours and gender role attitudes

	(1) Share	(2) Husband housework hours	(3) Wife housework hours
Age of husband	-0.006**	2.316***	-0.330
	(0.003)	(0.783)	(1.039)
Age of wife	0.002	-1.039	-0.345
	(0.003)	(0.874)	(1.148)
Education of husband	0.005	-3.152	1.181
	(0.006)	(2.046)	(2.556)
Education of wife	-0.018***	3.662*	-3.732
	(0.006)	(2.154)	(2.376)
Number of children	0.013	5.246**	15.512***
	(0.009)	(2.668)	(3.991)
Age of the youngest child	0.005**	-4.241***	-3.482***
	(0.002)	(0.704)	(0.933)
Gender role attitude of husband	0.007***	-1.851**	0.593
	(0.002)	(0.756)	(0.750)
Gender role attitude of wife	-0.002	0.664	0.829
	(0.002)	(0.687)	(0.744)
Constant	0.725***	80.283***	172.053***
	(0.077)	(25.672)	(32.103)
Observations	515	515	515

Notes: Standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Time-use regressors are measured in hours per month to align with the monthly consumption data. Using weekly hours yields the same significance levels. The housework share is defined as the wife's share of total housework hours. Higher values of the gender role attitude index indicate more conservative views of the individual. The number of observations (515) is smaller than in the main text because the age of the youngest child is included as a regressor, which excludes childless households from the sample. When this variable is excluded, the results remain largely consistent.